

А.Н.Колмогоров

ОСНОВНЫЕ ПОНЯТИЯ ТЕОРИИ ВЕРОЯТНОСТЕЙ

Книга, изданная в 1933 г. на немецком языке и в 1936 г. на русском, несколько раз переиздавалась в английском переводе. Хотя значительная часть содержания включена в учебники, она сохраняет интерес для лиц, занимающихся обстоятельно теорией вероятностей. Основной текст переиздается лишь с небольшой редакционной правкой.

СОДЕРЖАНИЕ

Предисловие к первому изданию	5
Предисловие ко второму изданию	7
I. Элементарная теория вероятностей	
§ 1. Аксиомы	10
§ 2. Отношение к данным опыта	12
§ 3. Терминологические замечания	14
§ 4. Непосредственные следствия из аксиом, условные вероятности, теорема Байеса	15
§ 5. Независимость	17
§ 6. Условные вероятности как случайные величины; цепи Маркова	23
II. Бесконечные поля вероятностей	
§ 1. Аксиома непрерывности	26
§ 2. Борелевские поля вероятностей	29
§ 3. Примеры бесконечных полей вероятностей	31
III. Случайные величины	
§ 1. Вероятностные функции	36
§ 2. Определение случайных величин, функции распределения	38
§ 3. Многомерные функции распределения	41
§ 4. Вероятности в бесконечномерных пространствах	44
§ 5. Эквивалентные случайные величины, разные виды сходимости	52
IV. Математические ожидания	
§ 1. Абстрактные интегралы Лебега	57
§ 2. Абсолютные и условные математические ожидания	60
§ 3. Неравенство Чебышева	63
§ 4. Некоторые признаки сходимости	65
§ 5. Дифференцирование и интегрирование математических ожиданий по параметру	66
V. Условные вероятности и математические ожидания	
§ 1. Условные вероятности	70
§ 2. Объяснение одного парадокса Бореля	75
§ 3. Условные вероятности относительно случайной величины	76
§ 4. Условные математические ожидания	78
VI. Независимость. Закон больших чисел	
§ 1. Независимость	83
§ 2. Независимые случайные величины	85
§ 3. Закон больших чисел	88
§ 4. Замечания к понятию математического ожидания	100
§ 5. Усиленный закон больших чисел, сходимость рядов	104
Дополнение. Одна замечательная теорема теории вероятностей	116
Литература	418

ПРЕДИСЛОВИЕ К ПЕРВОМУ ИЗДАНИЮ

Целью предлагаемой работы является аксиоматическое обоснование теории вероятностей. Ведущей мыслью автора было при этом естественное включение основ теории вероятностей, считавшихся еще недавно совершенно своеобразными, в ряд общих понятий современной математики. До возникновения лебеговой теории меры и интеграла эта задача была почти безнадежна. После исследований Лебега стала ясной аналогия между мерой множества и вероятностью события, а также между интегралом от функции и математическим ожиданием случайной величины. Эта аналогия допускает и дальнейшее продолжение: так, например, многие свойства независимых случайных величин вполне аналогичны соответствующим свойствам ортогональных функций. Для того чтобы, исходя из этой аналогии, обосновать теорию вероятностей, следовало еще освободить теорию меры и теорию интегрирования от геометрических элементов, которые еще имелись у Лебега. Это освобождение было осуществлено Фреше.

Попытки построения основ теории вероятностей, исходящие из этой общей точки зрения, уже имеются, и весь круг идей, излагаемых здесь, уже успел приобрести известную популярность в узком кругу специалистов; однако отсутствовало полное и свободное от излишних усложнений изложение всей системы (подготавливается, впрочем, к печати книга Фреше, см. Fréchet [2]).

Я хотел бы еще указать здесь на те места в дальнейшем изложении, которые выходят за пределы упомянутого выше круга идей, уже достаточно знакомого

в общих чертах специалистам. Эти места следующие: распределения вероятностей в бесконечномерных пространствах (глава третья, § 4), дифференцирование и интегрирование математических ожиданий по параметру (глава четвертая, § 5) и особенно теория условных вероятностей и математических ожиданий (глава пятая). Следует при этом отметить, что все эти новые понятия и проблемы с необходимостью возникают при рассмотрении вполне конкретных физических задач ¹⁾.

Шестая глава содержит обзор отдельных результатов А. Я. Хинчина и автора, касающихся условий применимости простого и усиленного закона больших чисел. В списке литературы приведены некоторые новые работы, представляющие интерес с точки зрения вопросов обоснования теории вероятностей.

Приношу свою сердечную благодарность А. Я. Хинчину, внимательно прочитавшему всю рукопись и предложившему целый ряд улучшений.

Клязьма близ Москвы,
1 мая 1933 г.

А. Колмогоров

¹⁾ Ср., например, цитированную в списке ¹⁾ к стр. 69 работу М. А. Леонтовича и автора, а также M. L e o n t o w i t s c h, Zur Statistik der kontinuierlichen Systeme und des zeitlichen Verlaufes der physikalischen Vorgänge, Physik Zeitschr. d. Sowjetunion, т. 3, 1933, стр. 35—63.

ПРЕДИСЛОВИЕ КО ВТОРОМУ ИЗДАНИЮ

С первого немецкого издания этой книжки прошло сорок лет. Было решено, тем не менее, не подвергать ее существенной переработке. А. Н. Ширяевым и мною внесены небольшие усовершенствования изложения. Модернизированы некоторые обозначения. Для некоторых теорем § 3 — 5 главы VI даны доказательства, отредактированные А. Н. Ширяевым по моим работам 1925—1930 годов. В современных учебниках эти теоремы обычно доказываются с помощью аппарата характеристических функций. Мои первоначальные доказательства прямыми, элементарными средствами, может быть, сохраняют некоторый интерес.

Намеченные в § 2 первой главы взгляды на пути обоснования применимости аксиоматической теории вероятностей к реальным задачам были развиты мною подробно в [1]. Но и здесь оставались невыясненными причины того, почему мы так часто встречаемся на практике с устойчивостью частот. Новый подход к этому вопросу был мною намечен в [2] и [3] (см. также [4]):

[1] Монография «Математика, ее содержание, методы и значение», изд. АН СССР 1956, глава XI.

[2] А. Н. К о л м о г о р о в, Три подхода к определению понятия «количество информации», Проблемы передачи информации, т. I, вып. 1 (1965).

[3] А. Н. К о л м о г о р о в, К логическим основам теории информации и теории вероятностей, Проблемы передачи информации, т. V, вып. 3 (1969).

[4] А. К. З в о н к и н и Л. А. Л е в и н, Сложность конечных объектов и обоснование теории информации

и случайности с помощью теории алгоритмов, Успехи математических наук, том 25, вып. 6 (1970).

Отмечу специально те вопросы, по которым читателю следует особенно настоятельно рекомендовать сопоставление изложения, данного в этой книжке, с более современным.

1. В § 1 главы V дано определение условной вероятности $P(A | \xi)$, где ξ — случайный элемент некоторого множества X , т. е. отображение Ω в X . С этим отображением можно связать алгебру $\mathcal{F}^\xi \subseteq \mathcal{F}$ всех принадлежащих \mathcal{F} полных прообразов подмножеств множества X . Теперь предпочитают сначала определять условные вероятности по отношению к любой σ -подалгебре $\mathcal{F}' \subseteq \mathcal{F}$ и затем считать, что

$$P(A | \xi) = P(A | \mathcal{F}^\xi).$$

2. Результаты § 4 главы III широко употребляются, но не дают непосредственно приемлемых распределений в имеющих реальный интерес функциональных пространствах (см. об этом на стр. 46).

17 декабря 1973 г.

А. Колмогоров

І. ЭЛЕМЕНТАРНАЯ ТЕОРИЯ ВЕРОЯТНОСТЕЙ

Мы называем *элементарной теорией вероятностей* ту часть теории вероятностей, в которой приходится иметь дело с вероятностями лишь конечного числа событий. Теоремы, которые здесь выводятся, естественно применяются также и к вопросам, связанным с бесконечным числом случайных событий, однако при изучении этих последних применяются также существенно новые принципы. Поэтому единственная аксиома математической теории вероятностей, относящаяся именно к случаю бесконечного числа случайных событий, вводится лишь в начале второй главы (аксиома V).

Теория вероятностей как математическая дисциплина может и должна быть аксиоматизирована совершенно в том же смысле, как геометрия или алгебра. Это означает, что, после того как даны названия изучаемым объектам и их основным отношениям, а также аксиомы, которым эти отношения должны подчиняться, все дальнейшее изложение должно основываться исключительно лишь на этих аксиомах, не опираясь на обычное конкретное значение этих объектов и их отношений.

Соответственно этому в § 1 определяется понятие *поля вероятностей* как системы множеств, удовлетворяющей определенным условиям. Что представляют собой элементы этих множеств, совершенно безразлично для чисто математического развития теории вероятностей (ср. введение основных геометрических понятий в «Основах геометрии» Гильберта или определение групп, колец и тел в абстрактной алгебре).

Всякая аксиоматическая (абстрактная) теория допускает, как известно, бесконечное число конкретных

интерпретаций. Таким образом и математическая теория вероятностей допускает наряду с теми интерпретациями, из которых она возникла, также много других. Так, мы приходим к приложениям математической теории вероятностей к таким областям науки, которые не имеют отношения к понятиям случая и вероятности в собственном смысле этого слова.

Аксиоматизация теории вероятностей может быть проведена различными способами как в отношении выбора аксиом, так и выбора основных понятий и основных соотношений. Если преследовать цель возможной простоты как самой системы аксиом, так и построения из нее дальнейшей теории, то представляется наиболее целесообразным аксиоматизирование понятий случайного события и его вероятности. Существуют также другие системы аксиоматического построения теории вероятностей, а именно такие, в которых понятие вероятности не относится к числу основных понятий, а само выражается через другие понятия¹⁾. При этом стремятся, однако, к другой цели, а именно, по возможности к наиболее тесному смыканию математической теории с эмпирическим возникновением понятия вероятности.

§ 1. Аксиомы²⁾

Пусть Ω — множество элементов ω , которые мы будем называть *элементарными событиями*, а \mathcal{F} — множество подмножеств из Ω . Элементы множества \mathcal{F} будем называть *случайными событиями* (или просто — *событиями*), а Ω — *пространством элементарных событий*.

I. \mathcal{F} является алгеброй множеств³⁾.

¹⁾ Ср., например, R. von Mises [1] и [2] и С. Н. Бернштейн [1].

²⁾ Читатель, желающий сразу же придать конкретный смысл нижеследующим аксиомам, должен начать читать § 2.

³⁾ Система \mathcal{F} подмножеств множества Ω называется *алгеброй*, если $\Omega \in \mathcal{F}$, соединение, пересечение и разность двух множеств системы опять принадлежат этой системе. Мы обозначаем пересечение множеств A и B через $A \cap B$ или AB , их соединение

II. Каждому множеству A из \mathcal{F} поставлено в соответствие неотрицательное действительное число $P(A)$. Это число называется вероятностью события A .

III. $P(\Omega) = 1$.

IV. Если A и B не пересекаются, то

$$P(A + B) = P(A) + P(B).$$

Совокупность объектов (Ω, \mathcal{F}, P) , удовлетворяющую аксиомам I—IV, будем называть *полем вероятностей*.

Наша система аксиом I—IV непротиворечива. Это показывает следующий пример: Ω состоит из единственного элемента ω , \mathcal{F} — из Ω и пустого множества \emptyset , при этом положено $P(\Omega) = 1$, $P(\emptyset) = 0$.

Наша система аксиом, однако, не является *полной*: в разных вопросах теории вероятностей рассматриваются различные поля вероятностей.

Простейшие поля вероятностей строятся следующим образом. Берутся произвольное конечное множество $\Omega = \{\omega_1, \dots, \omega_k\}$ и произвольное множество $\{p_1, \dots, p_k\}$ неотрицательных чисел с суммой $p_1 + \dots + p_k = 1$. За \mathcal{F} принимается совокупность всех подмножеств A из Ω и для $A = \{\omega_{i_1}, \dots, \omega_{i_\lambda}\}$ полагается $P(A) = p_{i_1} + \dots + p_{i_\lambda}$. В этом случае говорят, что p_1, \dots, p_k суть вероятности элементарных событий $\omega_1, \dots, \omega_k$ или просто *элементарные вероятности*. Так получаются все возможные *конечные* поля вероятностей, в которых \mathcal{F} состоит из совокупности всех подмножеств из Ω (при этом поле вероятностей называют конечным, если множество Ω конечно). По поводу дальнейших примеров см. гл. II, § 3.

— через $A \cup B$, разность — через $A \setminus B$. Дополнительное множество $\Omega \setminus A$ к множеству A обозначаем \bar{A} . Через \emptyset обозначается пустое множество. Если множества A и B не пересекаются ($AB = \emptyset$), то их соединение $A \cup B$ будет обозначаться также через $A + B$ и называться суммой. Множества из \mathcal{F} будем в дальнейшем обозначать большими латинскими буквами. Ср. А. Н. Колмогоров, С. В. Фомин, Элементы теории функций и функционального анализа, Изд-во «Наука», М., 1968.

§ 2. Отношение к данным опыта ¹⁾

Применение теории вероятностей к действительному миру опыта происходит по следующей схеме.

1. Предполагают данным некоторый комплекс \mathcal{S} условий, допускающий неограниченное число повторений.

2. Изучают определенный круг событий, которые могут наступать в результате осуществления условий \mathcal{S} . В отдельных случаях эти события могут наступать или не наступать в разных комбинациях. В множество Ω включаются все возможные варианты появлений или неоявлений рассматриваемых событий.

3. Если после реализации условий \mathcal{S} осуществившийся на практике вариант окажется принадлежащим к (определенному какими-либо условиями) множеству A , то говорят, что наступило событие A .

Пр и м е р. Комплекс условий \mathcal{S} заключается в том, что бросают два раза монету. Круг событий, о котором шла речь в п. 2, состоит в том, что при каждом бросании могут появиться решетка или герб. Отсюда следует, что всего возможно четыре различных варианта (элементарных события), именно: решетка — решетка, решетка — герб, герб — решетка, герб — герб. В качестве события A рассматривается «повторение». Это событие состоит из суммы первого и четвертого элементарных событий. Таким образом, можно каждое событие рассматривать как множество элементарных событий.

4. При известных условиях, в которые мы здесь не будем глубже вдаваться, можно предположить, что не-

¹⁾ Читатель, который интересуется лишь чисто математическим развитием теории, может этот параграф не читать — дальнейшее изложение основывается на аксиомах § 1 и не использует рассуждений этого параграфа. В нем мы ограничимся лишь простым указанием на эмпирическое возникновение аксиом теории вероятностей и сознательно оставляем в стороне глубокие философские изыскания о понятии вероятности в мире опыта. В изложении необходимых предпосылок для приложимости теории вероятностей к миру действительных событий автор в значительной мере следует выводам Мизеса, в частности ср. R. von Mises [1], стр. 21—27, параграф «Das Verhältnis der Theorie zur Erfahrungswelt».

которым событиям A , которые могут наступить или же не наступить после осуществления условий \mathfrak{E} , поставлены в соответствие определенные действительные числа $P(A)$, обладающие следующими свойствами:

А. Можно практически быть уверенным, что если комплекс условий \mathfrak{E} будет повторен большое число n раз и если при этом через m обозначено число случаев, при которых событие A наступило, то отношение $\frac{m}{n}$ будет мало отличаться от $P(A)$.

В. Если $P(A)$ очень мало, то можно практически быть уверенным, что при однократной реализации условий \mathfrak{E} событие A не будет иметь места.

Эмпирическая дедукция аксиом. Обычно можно предполагать, что система \mathcal{F} рассматриваемых событий A, B, C, \dots , которым приписаны определенные вероятности, образует алгебру множеств, содержащую в качестве элемента множество Ω (аксиома I, а также первая часть аксиомы II — существование вероятностей). Далее ясно, что $0 \leq \frac{m}{n} \leq 1$, так что вторая часть аксиомы II оказывается вполне естественной. Для события Ω всегда $m = n$, благодаря чему естественно положить $P(\Omega) = 1$ (аксиома III). Если, наконец, A и B несовместны между собой (т. е. множества A и B не пересекаются), то $m = m_1 + m_2$, где m, m_1, m_2 обозначают соответственно число опытов, в которых происходят события $A + B, A, B$. Отсюда следует:

$$\frac{m}{n} = \frac{m_1}{n} + \frac{m_2}{n}.$$

Следовательно, является уместным положить

$$P(A + B) = P(A) + P(B) \quad (\text{аксиома IV}).$$

Примечание I. Из практической достоверности двух утверждений следует практическая достоверность утверждения об их одновременной правильности, хотя степень достоверности при этом несколько понижается. Если, однако, число утверждений очень велико,

то из практической достоверности каждого отдельного из этих утверждений вообще нельзя вывести никаких заключений относительно одновременной правильности всех этих утверждений. Поэтому из принципа А никоим образом не следует, что при очень большом числе серий по n испытаний в каждой серии отношение $\frac{m}{n}$ будет мало отличаться от $P(A)$.

П р и м е ч а н и е II. Невозможному событию (пустому множеству \emptyset) соответствует в силу наших аксиом вероятность $P(\emptyset) = 0$ ¹⁾, в то время как, наоборот, из $P(A) = 0$ не следует еще невозможность события A ; согласно принципу В из обращения вероятности в нуль следует только, что при однократной реализации условий \mathcal{S} событие A практически невозможно. Это, однако, не означает, что при достаточно длинном ряде испытаний событие A также не наступит. Согласно принципу А можно лишь утверждать, что при $P(A) = 0$ и большом n отношение $\frac{m}{n}$ будет мало.

§ 3. Терминологические замечания

Мы определили объекты нашего дальнейшего изучения — случайные события — как множества. Многие теоретико-множественные понятия обозначаются, однако, в теории вероятностей другими именами. Мы приведем здесь краткий указатель таких понятий.

<i>В теории множеств</i>	<i>Для случайных событий</i>
1. A и B не пересекаются, т. е. $AB = \emptyset$.	1. События A и B несовместны.
2. $AB \dots N = \emptyset$.	2. События A, B, \dots, N несовместны.
3. $AB \dots N = X$.	3. Событие X заключается в одновременной реализации всех событий A, B, \dots, N .

¹⁾ Ср. § 4, формула (3).

$$4. A \cup B \cup \dots \cup N = X.$$

5. Дополнительное множество \bar{A} .

$$6. A = \emptyset.$$

$$7. A = \Omega.$$

8. Система \mathfrak{A} множеств A_1, A_2, \dots, A_n образует *разложение* множества Ω , если $A_1 + A_2 + \dots + A_n = \Omega$. (Это уже предполагает, что множества A_i попарно не пересекаются.)

9. B является подмножеством $A : B \subseteq A$.

4. Событие X заключается в наступлении по крайней мере одного из событий A, B, \dots, N .

5. Противоположное событие \bar{A} , состоящее в ненаступлении события A .

6. A невозможно.

7. A должно необходимо наступить.

8. *Испытание* \mathfrak{A} заключается в том, что устанавливают, какое из событий A_1, A_2, \dots, A_n происходит; A_1, A_2, \dots, A_n называются при этом возможными исходами испытания \mathfrak{A} .

9. Из осуществления события B с необходимостью следует осуществление A .

§ 4. Непосредственные следствия из аксиом, условные вероятности, теорема Байеса

Из $A + \bar{A} = \Omega$ и аксиом III и IV следует, что

$$P(A) + P(\bar{A}) = 1, \quad (1)$$

$$P(\bar{A}) = 1 - P(A). \quad (2)$$

Так как $\bar{\Omega} = \emptyset$, то, в частности,

$$P(\emptyset) = 0. \quad (3)$$

Если A, B, \dots, N несовместны, то из аксиомы IV следует формула

$$P(A + B + \dots + N) = P(A) + P(B) + \dots + P(N) \quad (4)$$

(теорема сложения).

Если $P(A) > 0$, то частное

$$P(B|A) = \frac{P(AB)}{P(A)} \quad (5)$$

называют *условной вероятностью* события B при условии A .

Из (5) непосредственно следует, что

$$P(AB) = P(B|A) \cdot P(A). \quad (6)$$

Заключение по индукции дает общую формулу

$$P(A_1 A_2 \dots A_n) = P(A_1) P(A_2 | A_1) P(A_3 | A_1 A_2) \dots P(A_n | A_1 \dots A_{n-1}) \quad (7)$$

(теорема умножения).

Легко доказываются также следующие формулы:

$$P(B|A) \geq 0, \quad (8)$$

$$P(\Omega|A) = 1, \quad (9)$$

$$P(B + C|A) = P(B|A) + P(C|A). \quad (10)$$

Сравнивая формулы (8) и (10) с аксиомами II—IV, получаем, что система множеств \mathcal{F} вместе с функцией множеств $P(B|A)$ (при закреплённом множестве A) образует поле вероятностей. Следовательно, все доказанные для вероятностей $P(B)$ общие теоремы справедливы для условных вероятностей $P(B|A)$ (при фиксированном событии A). Легко также заметить, что

$$P(A|A) = 1. \quad (11)$$

Из (6) и из аналогичной формулы

$$P(AB) = P(A|B) P(B)$$

получаем важную формулу

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)}, \quad (12)$$

содержащую, собственно, теорему Байеса.

Т е о р е м а (о полной вероятности). Пусть $A_1 + A_2 + \dots + A_n = \Omega$ и B — произвольное событие. Тогда

$$P(B) = P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + \dots + P(A_n)P(B|A_n). \quad (13)$$

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

$$B = A_1B + \dots + A_nB,$$

то, согласно (4),

$$P(B) = P(A_1B) + \dots + P(A_nB);$$

согласно (6) при этом имеет место равенство

$$P(A_iB) = P(A_i)P(B|A_i).$$

Т е о р е м а (Байеса). Пусть $A_1 + A_2 + \dots + A_n = \Omega$, а B — произвольное событие, тогда

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{P(A_1)P(B|A_1) + \dots + P(A_n)P(B|A_n)}. \quad (14)$$

Д о к а з а т е л ь с т в о. Согласно формуле (12)

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{P(B)}.$$

Для получения формулы (14) остается только заметить вероятность $P(B)$ ее выражением (13) по теореме о полной вероятности.

(События A_1, A_2, \dots, A_n часто называют «гипотезами» и говорят, что формула (14) дает вероятность $P(A_i|B)$ гипотезы A_i после наступления события B ; $P(A_i)$ означает при этом априорную вероятность A_i .)

§ 5. Независимость

Понятие *независимости* двух или нескольких опытов занимает в известном смысле центральное место в теории вероятностей. В самом деле, мы уже видели, что теорию вероятностей с математической

точки зрения можно рассматривать как специальное применение общей теории аддитивных функций множеств. Естественно задать вопрос, каким же образом теория вероятностей развилась в большую, обладающую своими собственными методами самостоятельную науку?

Для ответа на этот вопрос следует указать на ту специализацию, которую получают в теории вероятностей общие проблемы, касающиеся аддитивных функций множеств.

То обстоятельство, что наша аддитивная функция множеств $P = P(\cdot)$ неотрицательна и удовлетворяет условию $P(\Omega) = 1$, не обуславливает еще собой возникновения новых глубоких проблем. Случайные величины (ср. третью главу) с математической точки зрения представляют собой не что иное, как измеримые (относительно \mathcal{F}) функции, а их математические ожидания являются абстрактными интегралами Лебега. Эта аналогия была впервые полностью разъяснена в работах Фреше¹⁾. Введение упомянутых понятий не может, следовательно, еще доставить никакого базиса для развития большой оригинальной теории.

Исторически независимость испытаний и случайных величин явилась тем математическим понятием, которое придало теории вероятностей своеобразный отпечаток. Классические работы Лапласа, Пуассона, Чебышева, Маркова, Ляпунова, Мизеса и Бернштейна действительно посвящены в основном изучению рядов независимых случайных величин. Если в новейших исследованиях (Марков, Бернштейн и др.) часто отказываются от предположения полной независимости, то оказываются принужденными для получения достаточно содержательных результатов ввести аналогичные ослабленные предположения (ср. в этой главе § 6, о цепях Маркова). Мы приходим, следовательно, к тому, чтобы в понятии независимости видеть по крайней мере первый зародыш своеобразной проблематики теории вероятностей — обстоятельство, которое в этой

¹⁾ Ср. Fréchet [1] и [2].

книге будет мало выделяться, так как в ней мы занимаемся главным образом только логической подготовкой к собственно теоретико-вероятностным исследованиям.

Соответственно этому одной из важнейших задач философии естественных наук, после разъяснения пресловутого вопроса о сущности самого понятия вероятности, являются выяснение и уточнение тех предположений, при которых можно какие-либо данные действительные явления рассматривать как независимые. Этот вопрос выходит, однако, за пределы данной книги.

Перейдем к определению независимости.

Пусть даны n испытаний $\mathfrak{A}^{(1)}, \mathfrak{A}^{(2)}, \dots, \mathfrak{A}^{(n)}$, т. е. n разложений

$$\Omega = A_1^{(i)} + A_2^{(i)} + \dots + A_{r_i}^{(i)} \quad (i = 1, 2, \dots, n)$$

основного множества Ω на сумму (непересекающихся) событий. Тогда можно задать $r = r_1 r_2 \dots r_n$ вероятностей

$$p_{k_1 k_2 \dots k_n} = P(A_{k_1}^{(1)} A_{k_2}^{(2)} \dots A_{k_n}^{(n)}) \geq 0$$

вообще произвольно при единственном условии¹⁾

$$\sum_{(k_1, k_2, \dots, k_n)} p_{k_1 k_2 \dots k_n} = 1. \quad (1)$$

О п р е д е л е н и е 1. Испытания $\mathfrak{A}^{(1)}, \mathfrak{A}^{(2)}, \dots, \mathfrak{A}^{(n)}$ будем называть *независимыми*, если для любых k_1, k_2, \dots, k_n имеет место равенство

$$P(A_{k_1}^{(1)} A_{k_2}^{(2)} \dots A_{k_n}^{(n)}) = P(A_{k_1}^{(1)}) P(A_{k_2}^{(2)}) \dots P(A_{k_n}^{(n)}). \quad (2)$$

¹⁾ Поле вероятностей с произвольными вероятностями, удовлетворяющими только упомянутым условиям, можно построить следующим образом: множество Ω составляется из r элементов $\omega_{k_1 k_2 \dots k_n}$, соответствующие элементарные вероятности пусть будут $p_{k_1 k_2 \dots k_n}$ и $A_k^{(i)}$ определяется как множество всех $\omega_{k_1 k_2 \dots k_n}$, для которых $k_i = k$.

Среди r уравнений (2) имеется только $r - (r_1 + \dots + r_n) + (n - 1)$ независимых ¹⁾.

Т е о р е м а 1. Если испытания $\mathfrak{A}^{(1)}, \mathfrak{A}^{(2)}, \dots, \mathfrak{A}^{(n)}$ независимы, то из них любые m ($m < n$) испытаний $\mathfrak{A}^{(i_1)}, \mathfrak{A}^{(i_2)}, \dots, \mathfrak{A}^{(i_m)}$ также независимы ²⁾.

В случае независимости имеют, следовательно, место равенства

$$P(A_{k_1}^{(i_1)} A_{k_2}^{(i_2)} \dots A_{k_m}^{(i_m)}) = P(A_{k_1}^{(i_1)}) P(A_{k_2}^{(i_2)}) \dots P(A_{k_m}^{(i_m)}) \quad (3)$$

(предполагается, что i_l различны).

О п р е д е л е н и е II. События A_1, A_2, \dots, A_n называются *независимыми*, если разложения (испытания)

$$\Omega = A_k + \bar{A}_k \quad (k = 1, 2, \dots, n)$$

являются независимыми.

¹⁾ В самом деле, в случае независимости можно выбрать произвольно только $r_1 + r_2 + \dots + r_n$ вероятностей $p_k^{(i)} = P(A_k^{(i)})$, притом так, чтобы соблюдались n условий

$$\sum_k p_k^{(i)} = 1.$$

Следовательно, в общем случае имеем $r - 1$ степеней свободы, а в случае независимости только $r_1 + \dots + r_n - n$.

²⁾ Для доказательства достаточно установить, что из независимости n разложений следует независимость первых $n - 1$ из них. Примем, что уравнения (2) удовлетворены. Тогда

$$\begin{aligned} P(A_{k_1}^{(1)} A_{k_2}^{(2)} \dots A_{k_{n-1}}^{(n-1)}) &= \sum_{k_n} P(A_{k_1}^{(1)} A_{k_2}^{(2)} \dots A_{k_n}^{(n)}) = \\ &= P(A_{k_1}^{(1)}) P(A_{k_2}^{(2)}) \dots P(A_{k_{n-1}}^{(n-1)}) \sum_{k_n} P(A_{k_n}^{(n)}) = \\ &= P(A_{k_1}^{(1)}) P(A_{k_2}^{(2)}) \dots P(A_{k_{n-1}}^{(n-1)}). \end{aligned}$$

В этом случае $r_1 = r_2 = \dots = r_n = 2$, $r = 2^n$, следовательно, из 2^n уравнений (2) имеется независимых только $2^n - n - 1$. Для независимости событий A_1, A_2, \dots, A_n необходимы и достаточны следующие условия ¹⁾:

$$P(A_{i_1} A_{i_2} \dots A_{i_m}) = P(A_{i_1}) P(A_{i_2}) \dots P(A_{i_m}), \quad (4)$$

$$m = 1, 2, \dots, n; 1 \leq i_1 < i_2 < \dots < i_m \leq n.$$

Все эти уравнения независимы между собой.

В случае $n = 2$ получаем из (4) только одно ($2^2 - 2 - 1 = 1$) условие для независимости двух событий A_1 и A_2 :

$$P(A_1 A_2) = P(A_1) P(A_2). \quad (5)$$

Система уравнений (2) состоит в этом случае, кроме (5), еще из трех уравнений:

$$P(A_1 \bar{A}_2) = P(A_1) P(\bar{A}_2),$$

$$P(\bar{A}_1 A_2) = P(\bar{A}_1) P(A_2),$$

$$P(\bar{A}_1 \bar{A}_2) = P(\bar{A}_1) P(\bar{A}_2),$$

которые, очевидно, следуют из (5) ²⁾.

Следует при этом еще заметить, что из *попарной* независимости событий A_1, A_2, \dots, A_n , т. е. из соотношений

$$P(A_i A_j) = P(A_i) P(A_j) \quad (i \neq j)$$

в случае $n > 2$ отнюдь не следует независимость этих

¹⁾ Ср. С. Н. Б е р н ш т е й н [1], стр. 47—57. Впрочем, читатель может это сам проверить без труда (заключение по индукции).

²⁾ $P(A_1 \bar{A}_2) = P(A_1) - P(A_1 A_2) = P(A_1) - P(A_1) P(A_2) =$
 $= P(A_1) [1 - P(A_2)] = P(A_1) P(\bar{A}_2)$

событий ¹⁾ (для нее необходимо выполнение всех равенств (4)).

При введении понятия независимости мы не пользовались условными вероятностями. Нашей целью было при этом возможно яснее изложить чисто математически сущность этого понятия. Его приложения опираются, однако, главным образом на свойства некоторых условных вероятностей. Если мы предположим, что все вероятности положительны, то из уравнений (3) следует, что ²⁾

$$P(A_{k_m}^{(i_m)} | A_{k_1}^{(i_1)} A_{k_2}^{(i_2)} \dots A_{k_{m-1}}^{(i_{m-1})}) = P(A_{k_m}^{(i_m)}). \quad (6)$$

Обратно, из формул (6) следуют по теореме умножения (формула (7), § 4) формулы (2). Мы имеем, следовательно, следующую теорему.

Т е о р е м а II. *В случае положительных вероятностей всех $A_{k_i}^{(i)}$ для независимости испытаний $\mathfrak{A}^{(1)}, \mathfrak{A}^{(2)}, \dots, \mathfrak{A}^{(n)}$ необходимо и достаточно следующее условие: условная вероятность исхода $A_{k_i}^{(i)}$ при той гипотезе, что некоторые другие испытания $\mathfrak{A}^{(i_1)}, \mathfrak{A}^{(i_2)}, \dots, \mathfrak{A}^{(i_r)}$ получили определенные исходы $A_{k_1}^{(i_1)}, A_{k_2}^{(i_2)}, \dots, A_{k_r}^{(i_r)}$, равняется абсолютной вероятности $P(A_{k_i}^{(i)})$.*

¹⁾ Это доказывается следующим простым примером (С. Н. Бернштейн): множество Ω состоит из четырех элементов $\omega_1, \omega_2, \omega_3, \omega_4$, соответствующие элементарные вероятности p_1, p_2, p_3, p_4 все полагаются равными $1/4$ и $A = \{\omega_1, \omega_2\}$, $B = \{\omega_1, \omega_3\}$, $C = \{\omega_1, \omega_4\}$.

Легко тогда сосчитать, что

$$P(A) = P(B) = P(C) = \frac{1}{2},$$

$$P(AB) = P(BC) = P(AC) = \frac{1}{4} = \left(\frac{1}{2}\right)^2,$$

$$P(ABC) = \frac{1}{4} \neq \left(\frac{1}{2}\right)^3.$$

²⁾ Для доказательства следует вспомнить определение условной вероятности (формула (5), § 4) и заменить вероятности пересечений на произведения вероятностей по формуле (3).

На основе формул (4) аналогично доказывается следующая теорема.

Т е о р е м а III. *Если все вероятности $P(A_k)$ положительны, то для независимости событий A_1, A_2, \dots, A_n необходимо и достаточно, чтобы*

$$P(A_k | A_{k_1}, A_{k_2}, \dots, A_{k_l}) = P(A_k) \quad (7)$$

для любых попарно различных k, k_1, k_2, \dots, k_l .

В случае $n = 2$ условия (7) сводятся к двум уравнениям:

$$\begin{aligned} P(A_2 | A_1) &= P(A_2), \\ P(A_1 | A_2) &= P(A_1). \end{aligned} \quad (8)$$

Легко усмотреть, что уже одно лишь первое уравнение (8) представляет необходимое и достаточное условие независимости A_1 и A_2 , если только $P(A_1) > 0$.

§ 6. Условные вероятности как случайные величины; цепи Маркова

Пусть \mathfrak{M} является разложением основного множества Ω ,

$$\Omega = A_1 + A_2 + \dots + A_m,$$

а $\xi = \xi(\omega)$ — действительная функция элементарного события ω , которая на каждом множестве A_i принимает значение x_i :

$$\xi(\omega) = \sum_{i=1}^m x_i I_{A_i}(\omega),$$

где $I_{A_i}(\omega)$ — индикатор множества A_i , т. е. $I_{A_i}(\omega) = 1$, если $\omega \in A_i$ и $I_{A_i}(\omega) = 0$, если $\omega \in \bar{A}_i$.

В этом случае говорят, что ξ — случайная величина, принимающая конечное число значений x_1, \dots, x_m , и

$$M_{\xi}^x = \sum_{i=1}^m x_i P(A_i)$$

называют *математическим ожиданием* величины ξ . Теория случайных величин и их математических ожиданий будет развита в третьей и четвертой главах, не ограничиваясь при этом случайными величинами, которые могут принимать только конечное число различных значений.

О п р е д е л е н и е I. Случайную величину, которая на каждом множестве A_i принимает значение $P(B | A_i)$, мы назовем *условной вероятностью события B после данного испытания \mathfrak{A}* и обозначим $P(B | \mathfrak{A})(\omega)$, или просто $P(B | \mathfrak{A})$:

$$P(B | \mathfrak{A}) = \sum_{i=1}^m P(B | A_i) I_{A_i}(\omega).$$

Два испытания $\mathfrak{A}^{(1)}$ и $\mathfrak{A}^{(2)}$ тогда и только тогда независимы, когда

$$P(A_i^{(2)} | \mathfrak{A}^{(1)}) = P(A_i^{(2)}), \quad i = 1, 2, \dots, m_2.$$

Если даны какие-либо испытания $\mathfrak{A}^{(1)}, \mathfrak{A}^{(2)}, \dots, \mathfrak{A}^{(n)}$, то через

$$\mathfrak{A}^{(1)}\mathfrak{A}^{(2)} \dots \mathfrak{A}^{(n)}$$

мы обозначим испытание, соответствующее разложению множества Ω на произведения $A_{i_1}^{(1)} A_{i_2}^{(2)} \dots A_{i_n}^{(n)}$. Испытания $\mathfrak{A}^{(1)}, \mathfrak{A}^{(2)}, \dots, \mathfrak{A}^{(n)}$ тогда и только тогда независимы, когда

$$P(A_i^{(k)} | \mathfrak{A}^{(1)}\mathfrak{A}^{(2)} \dots \mathfrak{A}^{(k-1)}) = P(A_i^{(k)})$$

при любом выборе k и i ¹⁾.

О п р е д е л е н и е II. Последовательность испытаний

$$\mathfrak{A}^{(1)}, \mathfrak{A}^{(2)}, \dots, \mathfrak{A}^{(n)}, \dots$$

¹⁾ Необходимость этих условий следует из теоремы II § 5, а что они также являются достаточными, можно заключить непосредственно из теоремы умножения (формула (7) § 4).

образует *цепь Маркова*, если при любых k и i

$$P(A_i^{(k)} | \mathfrak{A}^{(1)}\mathfrak{A}^{(2)} \dots \mathfrak{A}^{(k-1)}) = P(A_i^{(k)} | \mathfrak{A}^{(k-1)}).$$

Цепи Маркова образуют естественное обобщение последовательностей независимых испытаний. Если положить

$$p_{i_m i_n}(m, n) = P(A_{i_n}^{(n)} | A_{i_m}^{(m)}), \quad m < n,$$

то основная формула теории цепей Маркова будет иметь следующий вид:

$$p_{i_k i_n}(k, n) = \sum_{i_m} p_{i_k i_m}(k, m) p_{i_m i_n}(m, n), \quad (1)$$

$$k < m < n$$

Обозначив матрицу $|| p_{i_m i_n}(m, n) ||$ через $p(m, n)$, можно (1) записать в следующем виде¹⁾:

$$p(k, n) = p(k, m) p(m, n), \quad (2)$$

$$k < m < n.$$

¹⁾ По поводу дальнейшего развития теории цепей Маркова см. R. von Mises [1], § 16 и B. Hostinsky, Méthodes générales du calcul des probabilités, Mém. Sci. Math. 52, Paris, 1931.

образует *цепь Маркова*, если при любых k и i

$$P(A_i^{(k)} | \mathfrak{A}^{(1)} \mathfrak{A}^{(2)} \dots \mathfrak{A}^{(k-1)}) = P(A_i^{(k)} | \mathfrak{A}^{(k-1)}).$$

Цепи Маркова образуют естественное обобщение последовательностей независимых испытаний. Если положить

$$p_{i_m i_n}(m, n) = P(A_{i_n}^{(n)} | A_{i_m}^{(m)}), \quad m < n,$$

то основная формула теории цепей Маркова будет иметь следующий вид:

$$p_{i_k i_n}(k, n) = \sum_{i_m} p_{i_k i_m}(k, m) p_{i_m i_n}(m, n), \quad (1)$$

$$k < m < n$$

Обозначив матрицу $|| p_{i_m i_n}(m, n) ||$ через $p(m, n)$, можно (1) записать в следующем виде¹⁾:

$$p(k, n) = p(k, m) p(m, n) \quad (2)$$

$$k < m < n.$$

¹⁾ По поводу дальнейшего развития теории цепей Маркова см. R. von Mises [1], § 16 и B. Hostinsky, Méthodes générales du calcul des probabilités, Mém. Sci. Math. 52, Paris, 1934.

II. БЕСКОНЕЧНЫЕ ПОЛЯ ВЕРОЯТНОСТЕЙ

§ 1. Аксиома непрерывности

Мы обозначаем, как обычно, через $\bigcap_m A_m$ пересечение множеств A_m (в конечном или бесконечном числе) и через $\bigcup_m A_m$ их соединение. В случае непересекающихся множеств A_m соединение $\bigcup_m A_m$ называем суммой и обозначаем $\sum_m A_m$. Следовательно,

$$\bigcup_m A_m = A_1 \cup A_2 \cup A_3 \cup \dots,$$

$$\sum_m A_m = A_1 + A_2 + A_3 + \dots,$$

$$\bigcap_m A_m = A_1 A_2 A_3 \dots$$

При всех дальнейших рассмотрениях мы предполагаем, что кроме аксиом I—IV (гл. I, § 1) выполняется еще следующая аксиома непрерывности:

V. Для убывающей последовательности

$$A_1 \supseteq A_2 \supseteq \dots \supseteq A_n \supseteq \dots \quad (1)$$

событий из \mathcal{F} такой, что

$$\bigcap_n A_n = \emptyset, \quad (2)$$

имеет место равенство

$$\lim_n \mathbf{P}(A_n) = 0. \quad (3)$$

Во всем дальнейшем изложении мы называем *полем вероятностей* только такое поле вероятностей (Ω, \mathcal{F}, P) в смысле главы первой, которое, кроме того, удовлетворяет аксиоме V. Поля вероятностей в смысле главы первой можно называть *полями вероятностей в расширенном смысле*.

Если система множеств \mathcal{F} конечна, аксиома V следует из аксиом I—IV. В самом деле, в этом случае существует только конечное число различных множеств в последовательности (1). Пусть A_k — наименьшее из них, тогда все множества A_{k+l} совпадают с A_k , и мы, следовательно, получаем

$$A_k = A_{k+l} = \bigcap_n A_n = \emptyset,$$

$$\lim_n P(A_n) = P(\emptyset) = 0.$$

Все примеры с конечными полями вероятностей из первой главы удовлетворяют, следовательно, также аксиоме V. Система аксиом I—V является, таким образом, непротиворечивой и неполной.

Напротив, для бесконечных полей аксиома непрерывности V является независимой от аксиом I—IV. Так как новая аксиома существенна лишь для бесконечных полей вероятностей, то является почти невозможным разъяснить ее эмпирическое значение, например, так, как это было вкратце сделано для аксиом I—IV в § 2 главы первой. При описании какого-либо действительно наблюдаемого случайного процесса можно получать только конечные поля вероятностей. Бесконечные поля вероятностей появляются только как идеализированные схемы действительных случайных явлений. Мы произвольно ограничиваемся при этом такими схемами, которые удовлетворяют аксиоме V. Это ограничение оказывается целесообразным в самых различных исследованиях.

Т е о р е м а I (расширенная теорема сложения). Если $A, A_1, \dots, A_n, \dots$ принадлежат \mathcal{F} и

$$A = \sum_n A_n, \quad (4)$$

то

$$P(A) = \sum_n P(A_n). \quad (5)$$

Доказательство. Положим

$$R_n = \sum_{m>n} A_m.$$

Тогда, очевидно,

$$\bigcap_n R_n = \emptyset,$$

и, следовательно, по аксиоме V

$$\lim_n P(R_n) = 0. \quad (6)$$

С другой стороны, по теореме сложения

$$P(A) = P(A_1) + \dots + P(A_n) + P(R_n). \quad (7)$$

Из (6) и (7) следует непосредственно (5).

Итак, мы доказали, что вероятность $P = P(\cdot)$ является на \mathcal{F} счетно-аддитивной функцией множеств. Обратно, аксиомы IV и V имеют место для всякой определенной на какой-либо алгебре множеств \mathcal{F} счетно-аддитивной функции множеств¹⁾. Можно, следовательно, понятие поля вероятностей определить следующим образом.

Пусть Ω — произвольное множество, \mathcal{F} — алгебра подмножеств Ω , а $P = P(\cdot)$ — неотрицательная счетно-аддитивная функция множеств, определенная на \mathcal{F} , подчиненная условию $P(\Omega) = 1$. Тогда система (Ω, \mathcal{F}, P) образует поле вероятностей.

Т е о р е м а II (о покрытиях). Если $A, A_1, A_2, \dots, A_n, \dots$ принадлежат \mathcal{F} и

$$A \subseteq \bigcup_n A_n, \quad (8)$$

то

$$P(A) \leq \sum_n P(A_n). \quad (9)$$

¹⁾ См., например, книгу А. Н. Колмогорова и С. В. Фомина, цитированную на стр. 11.

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

$$A = A (\cup A_n) = A (A_1 + A_2 \bar{A}_1 + A_3 \bar{A}_1 \bar{A}_2 + \dots),$$

то

$$\begin{aligned} P(A) &= P(AA_1) + P(AA_2 \bar{A}_1) + \dots \leq \\ &\leq P(A_1) + P(A_2) + \dots \end{aligned}$$

§ 2. Борелевские поля вероятностей

Алгебра \mathcal{F} подмножеств множества Ω называется *борелевской алгеброй*, если все счетные суммы $\sum_n A_n$ множеств A_n из \mathcal{F} принадлежат \mathcal{F} . Борелевские алгебры называют также σ -алгебрами. Из формулы

$$\bigcup_n A_n = A_1 + A_2 \bar{A}_1 + A_3 \bar{A}_1 \bar{A}_2 + \dots \quad (1)$$

можно заключить, что σ -алгебра содержит также все множества $\bigcup_n A_n$, составленные из счетного числа множеств A_n . Из формулы

$$\bigcap_n A_n = \Omega \setminus \bigcup_n \bar{A}_n$$

следует то же для пересечений множеств.

Поле вероятностей (Ω, \mathcal{F}, P) называется борелевским полем вероятностей, если соответствующая алгебра \mathcal{F} является борелевской.

При борелевских полях теория вероятностей получает полную свободу действия, не связанную с опасностью прийти к событиям, которые не имеют никакой вероятности. Мы теперь покажем, что можно ограничиться рассмотрением только борелевских полей вероятностей. Это будет следовать из так называемой теоремы о продолжении, к которой мы сейчас перейдем.

Пусть дано поле вероятностей $(\Omega, \mathcal{F}_0, P)$. Как известно ¹⁾, существует наименьшая σ -алгебра $\mathcal{F} = \sigma(\mathcal{F}_0)$, содержащая \mathcal{F}_0 .

Т е о р е м а (о продолжении). *Определенную на (Ω, \mathcal{F}_0) неотрицательную счетно-аддитивную функцию множеств $P = P(\cdot)$ всегда можно продолжить с сохранением обоих свойств (неотрицательности и счетной аддитивности) на все множества из $\mathcal{F} = \sigma(\mathcal{F}_0)$ и притом единственным образом.*

Борелевская алгебра $\mathcal{F} = \sigma(\mathcal{F}_0)$ вместе с продолженной функцией множеств $P = P(\cdot)$ образует некоторое поле вероятностей (Ω, \mathcal{F}, P) . Это поле назовем борелевским расширением поля $(\Omega, \mathcal{F}_0, P)$.

Д о к а з а т е л ь с т в о теоремы о продолжении, которая относится к теории аддитивных функций множеств и которая в основном должна быть известна в различных других трактовках, проводится по следующей схеме.

Пусть A — некоторое произвольное подмножество Ω . Определим $P^*(A)$ как нижнюю границу сумм

$$\sum_n P(A_n)$$

для всех покрытий

$$A \subseteq \bigcup_n A_n$$

множества A конечным или счетным числом множеств A_n из \mathcal{F}_0 . Легко доказать, что $P^*(A)$ является *внешней мерой* в смысле Каратеодори ²⁾. Согласно теореме о покрытиях (§ 1), $P^*(A)$ совпадает с $P(A)$ для всех множеств из \mathcal{F}_0 . Далее доказывается, что все множества из \mathcal{F}_0 измеримы в смысле Каратеодори. Так как все из-

¹⁾ Ф. Хаусдорф, Теория множеств, Гостехиздат, М., 1937; А. Н. Колмогоров, С. Ф. Фомин, Элементы теории функций и функционального анализа, Изд-во «Наука», 1968, стр. 44.

²⁾ С. Carathéodory, Vorlesungen über reelle Funktionen, Teubner, Berlin und Leipzig, 1918, стр. 237—258; А. Н. Колмогоров, С. В. Фомин (цит. выше), гл. V, § 3.

меримые множества образуют σ -алгебру, то, следовательно, все множества из $\sigma(\mathcal{F}_0)$ являются измеримыми. Функция множеств $P^*(A)$ является, следовательно, счетно-аддитивной на $\sigma(\mathcal{F}_0)$, и на $\sigma(\mathcal{F}_0)$ мы можем положить

$$P(A) = P^*(A).$$

Этим доказано существование продолжения. Единственность продолжения следует сразу же из минимальных свойств алгебры $\sigma(\mathcal{F}_0)$.

З а м е ч а н и е. Если множества (события) A из \mathcal{F}_0 могут иметь смысл в качестве действительных и наблюдавшихся (хотя бы приближенно) событий, то отсюда еще не следует, что множества из расширенной алгебры $\sigma(\mathcal{F}_0)$ допускают такое же разумное истолкование в качестве действительно наблюдавшихся событий. Может случиться, что поле вероятностей $(\Omega, \mathcal{F}_0, P)$ рассматривается в качестве (хотя бы идеализированного) образа реальных случайных событий, в то время как расширенное поле вероятностей $(\Omega, \sigma(\mathcal{F}_0), P)$ остается чисто математическим построением.

Множества из $\sigma(\mathcal{F}_0)$ мы рассматриваем только как «идеальные события», которым ничего не соответствует во внешнем мире. Если, однако, рассуждение, которое использует вероятности таких идеальных событий, приводит к определению вероятностей действительного события из \mathcal{F}_0 , то это определение, очевидно, автоматически будет непротиворечивым и с эмпирической точки зрения.

§ 3. Примеры бесконечных полей вероятностей

I. Еще в первой главе, § 1, мы строили различные конечные поля вероятностей. Пусть теперь

$$\Omega = \{\omega_1, \omega_2, \dots\}$$

— счетное множество, а \mathcal{F} совпадает с совокупностью всех подмножеств множества Ω .

Все возможные поля вероятностей с таким множеством \mathcal{F} получаются следующим образом: берется последовательность неотрицательных чисел $\{p_n\}$ при условии

$$p_1 + p_2 + \dots = 1$$

и для каждого множества A полагается

$$P(A) = \sum_n' p_n,$$

причем суммирование \sum_n' распространяется на все те индексы n , для которых ω_n принадлежит A . Эти поля вероятностей, очевидно, являются борелевскими.

II. Теперь предположим, что Ω представляет собой действительную числовую прямую R . Сначала пусть \mathcal{F}_0 образовано из всевозможных конечных сумм полуоткрытых интервалов $[a, b) = \{\omega: a \leq \omega < b\}$ (при этом мы рассматриваем наряду с собственными интервалами с конечными a и b также и несобственные $[-\infty, b)$, $[a, +\infty)$ и $[-\infty, +\infty)$). Легко убедиться, что \mathcal{F}_0 является алгеброй. По теореме о продолжении можно, однако, каждое поле вероятностей $(\Omega, \mathcal{F}_0, P)$ расширить в подобное поле $(\Omega, \sigma(\mathcal{F}_0), P)$. Система множеств $\mathcal{F} = \sigma(\mathcal{F}_0)$ в нашем случае является ни чем иным, как системой всех борелевских множеств числовой прямой.

III. Пусть $\Omega = R$ — действительная числовая прямая, а \mathcal{F} состоит из всех борелевских множеств этой прямой. Для построения вероятностного пространства с данной борелевской алгеброй \mathcal{F} достаточно определить на множествах $A \in \mathcal{F}$ любую неотрицательную счетно-аддитивную функцию множеств $P(A)$, удовлетворяющую условию $P(\Omega) = 1$. Такая функция, как известно ¹⁾, однозначно определяется своими значениями

$$P\{[-\infty, x)\} = F(x) \quad (1)$$

для специальных интервалов $[-\infty, x)$.

¹⁾ См., А. Лебег, Интегрирование и отыскание примитивных функций, ГТТИ, 1934, стр. 127—132.

Функцию $F = F(x)$ называют *функцией распределения* ω . Далее доказывается (третья глава, § 2), что $F(x)$ не убывает, непрерывна слева и имеет следующие предельные значения:

$$\begin{aligned}\lim_{x \rightarrow -\infty} F(x) &= F(-\infty) = 0, \\ \lim_{x \rightarrow +\infty} F(x) &= F(+\infty) = 1.\end{aligned}\quad (2)$$

Обратно, если дана функция $F = F(x)$, удовлетворяющая этим условиям, то она всегда определяет неотрицательную счетно-аддитивную функцию множеств $P(A)$ такую, что $P(\Omega) = 1$ ¹⁾.

IV. Пусть теперь за основное множество Ω принимается n -мерное евклидово координатное пространство R^n , т. е. множество всех упорядоченных комплексов $\omega = \{x_1, x_2, \dots, x_n\}$ из n действительных чисел. Система \mathcal{F} пусть состоит при этом из всех борелевских множеств²⁾ пространства R^n . На основании рассуждений, аналогичных приведенным в примере II, мы можем отказаться от рассмотрения более узких систем множеств, например системы всех n -мерных интервалов.

За вероятностную функцию $P(A)$ здесь, как всегда, можно принять любую неотрицательную и счетно-аддитивную функцию множеств, определенную на \mathcal{F} и удовлетворяющую условию $P(\Omega) = 1$. Такая функция множеств однозначно определяется, если даны ее значения

$$P(\Lambda_{a_1, a_2, \dots, a_n}) = F(a_1, a_2, \dots, a_n) \quad (3)$$

для специальных множеств $\Lambda_{a_1, a_2, \dots, a_n}$, где $\Lambda_{a_1, a_2, \dots, a_n}$ означает множество всех ω , для которых

$$x_i < a_i, \quad i = 1, 2, \dots, n.$$

1) См., например, цитированную выше книгу А. Н. Колмогорова и С. Ф. Фомина, стр. 262.

2) Определение борелевских множеств в R^n см. Ф. Хаусдорф, Теория множеств, Гостехиздат, М., 1937.

Нетрудно подсчитать, что для множества

$$\Lambda_{a_1, \dots, a_n}^{b_1, \dots, b_n} = \{\omega: a_1 \leq x_1 < b_1, \dots, a_n \leq x_n < b_n\}$$

вероятность

$$\begin{aligned} P(\Lambda_{a_1, \dots, a_n}^{b_1, \dots, b_n}) &= \\ &= F(b_1, \dots, b_n) - \sum_{i=1}^n F(b_1, \dots, b_{i-1}, a_i, b_{i+1}, \dots, b_n) + \\ &+ \sum_{i < j} F(b_1, \dots, b_{i-1}, a_i, b_{i+1}, \dots, b_{j-1}, a_j, b_{j+1}, \dots, b_n) - \dots \\ &\dots + (-1)^n F(a_1, \dots, a_n). \quad (4) \end{aligned}$$

За функцию $F(a_1, \dots, a_n)$ можно при этом выбрать любую непрерывную слева, неубывающую по всем переменным функцию, для которой выражение в правой части (4) неотрицательно при любых $a_i \leq b_i$, $i = 1, \dots, n$, и которая удовлетворяет также следующим условиям:

$$\begin{aligned} \lim_{a_i \rightarrow -\infty} F(a_1, \dots, a_n) &= F(a_1, \dots, a_{i-1}, -\infty, a_{i+1}, \dots, a_n) = 0, \\ i &= 1, 2, \dots, n, \end{aligned} \quad (5)$$

$$\lim_{a_1 \rightarrow +\infty, \dots, a_n \rightarrow +\infty} F(a_1, \dots, a_n) = F(+\infty, \dots, +\infty) = 1.$$

Функцию $F(a_1, \dots, a_n)$ называют *функцией распределения величин* x_1, \dots, x_n .

Рассмотрение полей вероятностей вышеопределенного типа достаточно для всех классических проблем теории вероятностей ¹⁾. В частности, вероятностная функция в R^n может быть определена следующим образом. Берется любая определенная в R^n неотрицательная

¹⁾ Ср., например, R. von M i s e s [1], стр. 13—19. Здесь требуется существование вероятностей для «всех практически возможных» множеств n -мерного пространства.

функция $f = f(x_1, \dots, x_n)$ такая, что

$$\int_{-\infty}^{\infty} \dots \int f(x_1, \dots, x_n) dx_1 \dots dx_n = 1,$$

и полагается

$$P(A) = \int_A \dots \int f(x_1, \dots, x_n) dx_1 \dots dx_n. \quad (6)$$

Функция $f(x_1, \dots, x_n)$ является в этом случае *плотностью вероятности* в точке (x_1, x_2, \dots, x_n) (ср. гл. III, § 2).

Другой тип вероятностных функций в R^n получается следующим образом. Пусть $\{\omega_i\}$ — последовательность точек из R^n и $\{p_i\}$ — последовательность неотрицательных действительных чисел такая, что $\sum_i p_i = 1$.

Тогда, так же как и в примере I, полагаем

$$P(A) = \sum_i' p_i,$$

причем суммирование \sum_i' распространяется на все те индексы i , для которых ω_i принадлежит A . Оба упомянутых здесь типа вероятностных функций в R^n не исчерпывают всех возможностей, хотя ими обычно довольствуются в приложениях теории вероятностей.

Можно, однако, себе представить, кроме этой классической области, также и другие, интересные для приложений задачи, в которых элементарные события определяются с помощью бесконечного числа координат. Соответствующие поля вероятностей мы исследуем ближе после введения некоторых необходимых для этого вспомогательных понятий (ср. гл. III, § 3).

III. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

§ 1. Вероятностные функции

Пусть дано отображение множества Ω в множество X , состоящее из каких-либо элементов, т. е. определенная на Ω однозначная функция $\xi = \xi(\omega)$, значения которой принадлежат множеству X . Каждому подмножеству A из X мы ставим в соответствие в качестве его прообраза в Ω множество $\xi^{-1}(A)$ всех элементов из Ω , которые отображаются в один из элементов A . Пусть далее \mathcal{F}_ξ — система всех подмножеств A из X , прообразы которых принадлежат к алгебре множеств \mathcal{F} . Система \mathcal{F}_ξ тогда также является алгеброй. Если при этом \mathcal{F} — борелевская алгебра, то тоже имеет место и для \mathcal{F}_ξ .

Мы полагаем теперь

$$P_\xi(A) = P\{\xi^{-1}(A)\}. \quad (1)$$

Эта определенная на \mathcal{F}_ξ функция множеств P_ξ удовлетворяет относительно алгебры \mathcal{F}_ξ всем нашим аксиомам I—V и, следовательно, является вероятностной функцией на \mathcal{F}_ξ . Прежде чем перейти к доказательству всех только что указанных фактов, мы сформулируем уже теперь следующее

О п р е д е л е н и е. Пусть дана однозначная функция $\xi = \xi(\omega)$ случайного события ω . Тогда функция P_ξ , определенная формулой (1), называется *вероятностной функцией* ξ .

Примечание I. При исследовании поля вероятностей (Ω, \mathcal{F}, P) функцию P называют вероятностной функцией или просто вероятностью, а P_ξ — вероятно-

стной функцией ξ . В случае $\xi(\omega) = \omega$ $P_\xi(A)$ совпадает с $P(A)$.

Примечание II. Событие $\xi^{-1}(A)$ состоит в том, что $\xi(\omega)$ принадлежит множеству A . Следовательно, $P_\xi(A)$ есть вероятность того, что $\xi(\omega) \in A$.

Нам осталось доказать вышеупомянутые свойства \mathcal{F}_ξ и P_ξ . Они следуют, однако, из одного-единственного факта, а именно следующего.

Лемма. Сумма, пересечение и разность каких-либо прообразных множеств $\xi^{-1}(A)$ являются прообразами соответствующих сумм, пересечений и разностей множеств A .

Доказательство этой леммы предоставляет-ся читателю.

Пусть теперь A и B — два множества из \mathcal{F}_ξ , их прообразы A' и B' принадлежат тогда \mathcal{F} . Так как \mathcal{F} — алгебра, то множества $A'B'$, $A' + B'$ и $A' \setminus B'$ также принадлежат \mathcal{F} . Но эти множества являются прообразами множеств AB , $A + B$ и $A \setminus B$, следовательно, последние множества принадлежат к \mathcal{F}_ξ . Итак, мы доказали, что \mathcal{F}_ξ — алгебра. Так же доказывается, что если \mathcal{F} является борелевской алгеброй, то то же справедливо и для \mathcal{F}_ξ .

Далее ясно, что

$$P_\xi(X) = P\{\xi^{-1}(X)\} = P(\Omega) = 1.$$

Что P_ξ всегда неотрицательна, понятно само собой. Следовательно, остается доказать, что P_ξ счетно-аддитивна (ср. конец § 1 гл. II).

Итак, пусть множества A_n , а следовательно, и их прообразы $\xi^{-1}(A_n)$ не пересекаются. Тогда

$$\begin{aligned} P_\xi\left(\sum_n A_n\right) &= P\left\{\xi^{-1}\left(\sum_n A_n\right)\right\} = \\ &= P\left\{\sum_n \xi^{-1}(A_n)\right\} = \sum_n P\{\xi^{-1}(A_n)\} = \sum_n P_\xi(A_n), \end{aligned}$$

чем доказана счетная аддитивность P_ξ .

В заключение заметим еще следующее. Пусть $\xi_1 = \xi_1(\omega)$ — функция, отображающая Ω в X_1 , а $\xi_2 = \xi_2(x_1)$ — другая функция, отображающая X_1 в X_2 . Тогда сложная функция $\xi(\omega) = \xi_2[\xi_1(\omega)]$ отображает множество Ω в X_2 . Мы рассмотрим теперь вероятностные функции $P_{\xi_1}(A_1)$ и $P_{\xi_2}(A_2)$ для функций $\xi_1(\omega)$ и $\xi(\omega) = \xi_2[\xi_1(\omega)]$. Эти две вероятностные функции связаны, как легко подсчитать, следующим соотношением:

$$P_{\xi}(A_2) = P_{\xi_1}\{\xi_2^{-1}(A_2)\}. \quad (2)$$

§ 2. Определение случайных величин, функции распределения

О п р е д е л е н и е I. Однозначную действительную функцию $\xi = \xi(\omega)$, определенную на основном множестве Ω , называют *случайной величиной*, если при каждом выборе действительного числа x множество $\{\xi < x\}$ всех тех ω , для которых справедливо неравенство $\xi(\omega) < x$, принадлежит к системе множеств \mathcal{F} .

Эта функция $\xi(\omega)$ отображает основное множество Ω на множество R всех действительных чисел. Наше определение случайной величины можно теперь сформулировать так: действительная функция $\xi = \xi(\omega)$ является случайной величиной тогда и только тогда, когда \mathcal{F}_{ξ} содержит каждый интервал вида $(-\infty, a)$.

Так как \mathcal{F}_{ξ} — алгебра, то она содержит наряду с интервалами $(-\infty, a)$ также всевозможные конечные суммы полуоткрытых интервалов $[a, b)$. Если наше поле вероятностей борелевское, то \mathcal{F} и \mathcal{F}_{ξ} являются борелевскими алгебрами; следовательно, в этом случае \mathcal{F}_{ξ} содержит все борелевские множества R . Вероятностная функция P_{ξ} случайной величины ξ определена для всех множеств A алгебры \mathcal{F}_{ξ} . В частности, в важнейшем случае борелевского поля вероятностей P_{ξ} определена для всех борелевских множеств R .

О п р е д е л е н и е II. Функция

$$F_{\xi}(x) = P_{\xi}(-\infty, x) = P\{\xi(\omega) < x\},$$

где $-\infty$ и $+\infty$ допускаются в качестве значений x , называется *функцией распределения* случайной величины ξ .

Из определения непосредственно следует, что

$$F_{\xi}(-\infty) = 0, \quad F_{\xi}(+\infty) = 1. \quad (1)$$

Вероятность выполнения неравенств $a \leq \xi < b$, очевидно, задается формулой

$$P(a \leq \xi(\omega) < b) = F_{\xi}(b) - F_{\xi}(a). \quad (2)$$

Отсюда следует, что для $a < b$

$$F_{\xi}(a) \leq F_{\xi}(b).$$

Это означает, что $F_{\xi}(x)$ — *неубывающая* функция.

Пусть далее $a_1 < a_2 < \dots < a_n < \dots \rightarrow b$, тогда

$$\bigcap_n \{\omega: \xi(\omega) \in [a_n, b)\} = \emptyset.$$

Следовательно, согласно аксиоме непрерывности

$$F_{\xi}(b) - F_{\xi}(a_n) = P\{\omega: \xi(\omega) \in [a_n, b)\}$$

стремится к нулю при $n \rightarrow \infty$. Отсюда видно, что функция $F_{\xi}(x)$ *непрерывна слева*.

Аналогично можно доказать, что

$$\lim_{x \rightarrow -\infty} F_{\xi}(x) = F_{\xi}(-\infty) = 0, \quad (3)$$

$$\lim_{x \rightarrow +\infty} F_{\xi}(x) = F_{\xi}(+\infty) = 1. \quad (4)$$

Если поле вероятностей (Ω, \mathcal{F}, P) — борелевское, то значения вероятностной функции $P_{\xi}(A)$ для всех борелевских множеств A из R однозначно определяются знанием функции распределения $F_{\xi}(x)$ (ср. гл. II, § 3, III). Так как главным образом интересуются только этими значениями $P_{\xi}(A)$, то функции распределения играют во всем дальнейшем изложении существенную роль.

Если функция распределения $F_{\xi}(x)$ дифференцируема, то ее производную по x

$$f_{\xi}(x) = \frac{dF_{\xi}(x)}{dx}$$

называют *плотностью вероятности* ξ в точке x .

Если для каждого x

$$F_{\xi}(x) = \int_{-\infty}^x f_{\xi}(y) dy,$$

то для каждого борелевского множества A вероятностную функцию ξ можно выразить через $f_{\xi}(x)$ следующим образом:

$$P_{\xi}(A) = \int_A f_{\xi}(x) dx. \quad (5)$$

В этом случае говорят, что *распределение* ξ *абсолютно непрерывно*. В общем случае по аналогии пишут:

$$P_{\xi}(A) = \int_A dF_{\xi}(x). \quad (6)$$

Все введенные понятия допускают обобщение на случай условных вероятностей. Функция множеств

$$P_{\xi}(A | B) = P \{ \xi \in A | B \}$$

является *условной вероятностной функцией* ξ при гипотезе B (предполагается, что $P(B) > 0$). Неубывающая функция

$$F_{\xi}(x | B) = P \{ \xi < x | B \}$$

есть соответствующая функция распределения, и, наконец (в случае дифференцируемости $F_{\xi}(x | B)$),

$$f_{\xi}(x | B) = \frac{dF_{\xi}(x | B)}{dx}$$

— *условная плотность вероятности* ξ в точке x при гипотезе B .

§ 3. Многомерные функции распределения

Пусть теперь даны n случайных величин $\xi_1, \xi_2, \dots, \xi_n$. Точка $\xi = (\xi_1, \dots, \xi_n)$ n -мерного пространства R^n является функцией элементарного события ω . Следовательно, по общим правилам § 1 получаем алгебру множеств $\mathcal{F}_\xi = \mathcal{F}_{\xi_1, \dots, \xi_n}$, состоящую из подмножеств пространства R^n , и определенную на $\mathcal{F}_\xi = \mathcal{F}_{\xi_1, \dots, \xi_n}$ вероятностную функцию $P_\xi(A) = P_{\xi_1, \dots, \xi_n}(A)$. Эту вероятностную функцию называют *n -мерной вероятностной функцией случайных величин $\xi = (\xi_1, \dots, \xi_n)$* .

Алгебра $\mathcal{F}_{\xi_1, \dots, \xi_n}$ содержит (как это прямо следует из определения случайной величины) при каждом выборе i и a_i , $i = 1, 2, \dots, n$, множество всех точек $x = (x_1, \dots, x_n) \in R^n$, для которых $x_i < a_i$. Следовательно, $\mathcal{F}_{\xi_1, \dots, \xi_n}$ содержит также и пересечение указанных множеств, т. е. множество $\Lambda_{a_1, a_2, \dots, a_n}$ всех точек $x = (x_1, \dots, x_n) \in R^n$, для которых выполняются все неравенства $x_i < a_i$, $i = 1, 2, \dots, n$ ¹⁾.

Если назвать n -мерным полуоткрытым интервалом

$$[a_1, a_2, \dots, a_n; b_1, b_2, \dots, b_n)$$

множество всех точек R^n , для которых выполняются все неравенства $a_i \leq x_i < b_i$, то видно сразу, что каждый такой интервал также принадлежит алгебре $\mathcal{F}_{\xi_1, \dots, \xi_n}$, поскольку

$$\begin{aligned} & [a_1, a_2, \dots, a_n; b_1, b_2, \dots, b_n) = \\ & = \Lambda_{b_1, b_2, \dots, b_n} - \Lambda_{a_1, b_2, \dots, b_n} - \Lambda_{a_1, a_2, b_3, \dots, b_n} - \Lambda_{b_1, \dots, b_{n-1}, a_n}. \end{aligned}$$

Борелевское расширение системы всех n -мерных полуоткрытых интервалов состоит из всех борелевских множеств R^n . Отсюда следует, что алгебра $\mathcal{F}_{\xi_1, \dots, \xi_n}$ в случае борелевского поля вероятностей содержит все борелевские множества пространства R^n .

¹⁾ x_i могут принимать также бесконечные значения $\pm \infty$.

Т е о р е м а. В случае борелевского поля вероятностей и борелевской функции $f(x_1, \dots, x_n)$ функция $\eta(\omega) = f(\xi_1(\omega), \dots, \xi_n(\omega))$ конечного числа случайных величин $\xi_1 = \xi_1(\omega), \dots, \xi_n = \xi_n(\omega)$ тоже является случайной величиной.

Для доказательства достаточно заметить, что множество всех точек (x_1, x_2, \dots, x_n) в R^n , для которых $f(x_1, x_2, \dots, x_n) < a$, является борелевским. В частности, все конечные суммы и произведения случайных величин тоже являются случайными величинами.

О п р е д е л е н и е. Функцию

$$F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) = P_{\xi_1, \dots, \xi_n}(\Lambda_{x_1, \dots, x_n})$$

называем n -мерной функцией распределения случайных величин ξ_1, \dots, ξ_n .

Как и в одномерном случае доказывается, что n -мерная функция распределения $F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n)$ не убывает по всем переменным и непрерывна слева. Аналогично равенствам (3) и (4) из § 2

$$\begin{aligned} \lim_{x_i \rightarrow -\infty} F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) &= \\ &= F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_{i-1}, \infty, x_i, \dots, x_n) = 0, \quad (1) \end{aligned}$$

$$\begin{aligned} \lim_{x_1 \rightarrow +\infty, \dots, x_n \rightarrow +\infty} F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) &= \\ &= F_{\xi_1, \dots, \xi_n}(+\infty, \dots, +\infty) = 1. \quad (2) \end{aligned}$$

Функция распределения $F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n)$ непосредственно дает нам значения P_{ξ_1, \dots, ξ_n} только для специальных множеств $\Lambda_{x_1, \dots, x_n}$. Если, однако, наше поле вероятностей — борелевское, то P_{ξ_1, \dots, ξ_n} однозначно определяется для борелевских множеств через функцию распределения $F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n)$ ¹⁾.

¹⁾ Ср. гл. II, § 3, IV.

Если существует производная

$$f_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) = \frac{\partial^n}{\partial x_1 \dots \partial x_n} F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n),$$

то эту производную $f_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n)$ называют *n-мерной плотностью вероятности* случайных величин ξ_1, \dots, ξ_n в точке (x_1, \dots, x_n) .

Если для каждой точки (x_1, \dots, x_n)

$$\begin{aligned} F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) = \\ = \int_{-\infty}^{x_1} \dots \int_{-\infty}^n f_{\xi_1, \dots, \xi_n}(y_1, \dots, y_n) dy_1 \dots dy_n, \end{aligned}$$

то распределение $\xi = (\xi_1, \dots, \xi_n)$ называют *абсолютно непрерывным*. Для каждого борелевского множества $A \subseteq R^n$ имеет место равенство

$$P_{\xi_1, \dots, \xi_n}(A) = \int_A \dots \int f_{\xi_1, \dots, \xi_n}(y_1, \dots, y_n) dy_1 \dots dy_n. \quad (3)$$

В заключение этого параграфа сделаем еще одно замечание о соотношениях между различными вероятностными функциями и функциями распределения. Пусть дана подстановка

$$S = \begin{pmatrix} 1, & 2, & \dots, & n \\ i_1, & i_2, & \dots, & i_n \end{pmatrix},$$

и пусть ρ_S означает преобразование

$$x'_k = x_{i_k}, \quad k = 1, 2, \dots, n,$$

пространства R^n в самого себя. Тогда ясно, что

$$P_{\xi_{i_1}, \dots, \xi_{i_n}}(A) = P_{\xi_1, \dots, \xi_n}(\rho_S^{-1}(A)). \quad (4)$$

Пусть теперь $x' = \pi_k(x)$ — «проекция» пространства R^n в пространство R^k , $k < n$, при которой точка

$x = (x_1, \dots, x_n)$ отображается в точку $x' = (x_1, \dots, x_k)$. Тогда вследствие формулы (2) § 1

$$P_{\xi_1, \dots, \xi_k}(A) = P_{\xi_1, \dots, \xi_n} \{ \pi_k^{-1}(A) \}. \quad (5)$$

Для соответствующих функций распределения из (4) и (5) следуют равенства

$$F_{\xi_{i_1}, \dots, \xi_{i_n}}(x_{i_1}, \dots, x_{i_n}) = F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n),$$

$$F_{\xi_1, \dots, \xi_k}(x_1, \dots, x_k) = F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_k, +\infty, \dots, +\infty).$$

§ 4. Вероятности в бесконечномерных пространствах

В § 3 главы второй мы видели, как строятся различные применяющиеся в теории вероятностей поля вероятностей. Можно, однако, представить себе интересные также и для приложений проблемы, в которых элементарные события определяются с помощью бесконечного числа координат. Итак, пусть выбрано множество \mathcal{N} индексов v любой мощности \aleph . Совокупность всех систем

$$\omega = \{x_v\}$$

действительных чисел x_v , где v пробегает все множество \mathcal{N} , назовем пространством $R^{\mathcal{N}}$ (для определения элемента ω пространства $R^{\mathcal{N}}$ следует каждому элементу v множества \mathcal{N} поставить в соответствие действительное число x_v или, что то же самое, задать определенную на \mathcal{N} однозначную действительную функцию x_v элемента v ¹⁾).

Если множество \mathcal{N} состоит из n первых натуральных чисел $1, 2, \dots, n$, то $R^{\mathcal{N}}$ есть обычное n -мерное пространство R^n . Если выбрать в качестве множества \mathcal{N} множество всех действительных чисел R^1 , то соответствующее пространство $R^{\mathcal{N}} = R^{R^1}$ состоит из всех

¹⁾ Ср. Ф. Х а у с д о р ф, Теория множеств, Гостехиздат, М., 1937.

действительных функций

$$\omega = \{x_t\}$$

действительного переменного t , $-\infty < t < \infty$.

Множество R^{1^c} при произвольном множестве \mathcal{N} мы примем сейчас за основное множество Ω . Пусть $\omega = \{x_v\}$ — элемент Ω . Через $\pi_{v_1, \dots, v_n}(\omega)$ будем обозначать точку $(x_{v_1}, x_{v_2}, \dots, x_{v_n})$ n -мерного пространства R^1 . Подмножество A из Ω назовем *цилиндрическим* множеством, если оно представимо в форме

$$A = \pi_{v_1, \dots, v_n}^{-1}(A'),$$

где A' есть подмножество R^n . Класс всех цилиндрических множеств совпадает, следовательно, с классом всех таких множеств, которые могут быть определены соотношениями вида

$$f(x_{v_1}, x_{v_2}, \dots, x_{v_n}) = 0. \quad (1)$$

Для того чтобы произвольное цилиндрическое множество $\pi_{v_1, v_2, \dots, v_n}^{-1}(A')$ определить таким соотношением, достаточно принять за f функцию, которая на A' равна нулю, а вне A' равна единице.

Цилиндрическое множество является борелевским цилиндрическим множеством, если соответствующее множество A' — борелевское. Все борелевские цилиндрические множества¹⁾ пространства $R^{\mathcal{N}}$ образуют алгебру, которая в дальнейшем будет обозначаться $\mathcal{F}^{\mathcal{N}}$.

¹⁾ Из вышеприведенных рассуждений следует, что борелевские цилиндрические множества — это те, которые могут быть определены борелевскими соотношениями (1). Пусть теперь A и B — два цилиндрических множества, определенных соотношениями

$$f(x_{v_1}, \dots, x_{v_n}) = 0, \quad g(x_{\lambda_1}, \dots, x_{\lambda_m}) = 0.$$

Тогда множества $A + B$, AB и $A \setminus B$ можно определить соответственно следующими соотношениями:

$$f \cdot g = 0, \quad f^2 + g^2 = 0, \quad f^2 + h(g) = 0,$$

Борелевское расширение алгебры $\mathcal{F}^{\mathcal{N}}$ мы обозначим, как всегда, через $\sigma(\mathcal{F}^{\mathcal{N}})$. Множества из системы $\sigma(\mathcal{F}^{\mathcal{N}})$ мы называем *борелевскими множествами пространства $R^{\mathcal{N}}$* .

Далее будет дан метод для построения и оперирования с вероятностными функциями на $\mathcal{F}^{\mathcal{N}}$ и, следовательно, по теореме о продолжении также и на $\sigma(\mathcal{F}^{\mathcal{N}})$. В случае счетности множества \mathcal{N} получаемые таким образом поля вероятностей достаточны для всех целей. Мы овладаем, следовательно, всеми вопросами, касающимися счетной последовательности случайных величин. Если же \mathcal{N} несчетно, то тогда многие простые и интересные подмножества $R^{\mathcal{N}}$ остаются вне системы $\sigma(\mathcal{F}^{\mathcal{N}})$. Например, множество всех элементов ω , для которых x_v при любом выборе индекса v остается меньше какой-либо постоянной величины c , т. е. множество $\{\omega: x_v < c, v \in \mathcal{N}\}$ не принадлежит к системе $\sigma(\mathcal{F}^{\mathcal{N}})$ в случае несчетности множества \mathcal{N} .

Следовательно, всегда стоит добиваться, если это возможно, приведения всякой проблемы к такой форме, при которой пространство всех элементарных событий ω имеет только счетное множество координат.

Пусть на $\mathcal{F}^{\mathcal{N}}$ определена вероятностная функция P . Каждую координату x_v элементарного события ω можно рассматривать как случайную величину. Следовательно, всякая конечная группа $(x_{v_1}, x_{v_2}, \dots, x_{v_n})$ этих координат имеет n -мерную вероятностную функцию $P_{v_1, v_2, \dots, v_n}(A)$ и соответствующую функцию распределения $F_{v_1, v_2, \dots, v_n}(a_1, a_2, \dots, a_n)$. Ясно, что для всякого борелевского цилиндрического множества

$$A = \pi_{v_1, v_2, \dots, v_n}^{-1}(A')$$

где $h(x) = 0$ для $x \neq 0$ и $h(0) = 1$. Если f и g — борелевские функции, то таковыми же являются $f \cdot g$, $f^2 + g^2$ и $f^2 + h(g)$. Следовательно, $A + B$, AB и $A \setminus B$ являются борелевскими цилиндрическими множествами. Этим доказано, что система множеств $\mathcal{F}^{\mathcal{N}}$ является алгеброй,

имеет место равенство

$$P(A) = P_{\nu_1, \nu_2, \dots, \nu_n}(A'),$$

причем A' является борелевским множеством в R^n . Таким образом, вероятностная функция P однозначно определяется на алгебре $\mathcal{F}^{\mathcal{N}}$ всех цилиндрических множеств через значения всех конечномерных вероятностных функций $P_{\nu_1, \nu_2, \dots, \nu_n}$ для всех борелевских множеств соответствующего пространства R^n . Однако для борелевских множеств значения вероятностной функции $P_{\nu_1, \nu_2, \dots, \nu_n}$ однозначно определяются через соответствующие функции распределения. Следовательно, мы доказали следующую теорему.

Т е о р е м а. *Совокупность всех конечномерных функций распределения $F_{\nu_1, \nu_2, \dots, \nu_n}$ однозначно определяет вероятностную функцию P для всех множеств из $\mathcal{F}^{\mathcal{N}}$. Вероятностная функция P (по теореме о продолжении) определяется однозначно на $\sigma(\mathcal{F}^{\mathcal{N}})$ значениями функций распределения $F_{\nu_1, \nu_2, \dots, \nu_n}$.*

Теперь можно поставить вопрос, при каких условиях а priori заданная система функций распределения $F_{\nu_1, \nu_2, \dots, \nu_n}$ определяет вероятность на $\mathcal{F}^{\mathcal{N}}$ (и, следовательно, на $\sigma(\mathcal{F}^{\mathcal{N}})$).

Сперва отметим, что всякая функция распределения $F_{\nu_1, \nu_2, \dots, \nu_n}$ должна удовлетворять условиям, данным в главе второй (§ 3, III), что, конечно, содержится в самом понятии функции распределения. Кроме того, вследствие формул (13) и (14) из § 2 имеют место еще следующие соотношения:

$$F_{\nu_{i_1}, \nu_{i_2}, \dots, \nu_{i_n}}(x_{i_1}, x_{i_2}, \dots, x_{i_n}) = F_{\nu_1, \nu_2, \dots, \nu_n}(x_1, x_2, \dots, x_n), \quad (2)$$

$$\begin{aligned} F_{\nu_1, \nu_2, \dots, \nu_k}(x_1, x_2, \dots, x_k) = \\ = F_{\nu_1, \nu_2, \dots, \nu_n}(x_1, x_2, \dots, x_k, +\infty, \dots, +\infty), \end{aligned} \quad (3)$$

где $k < n$ и $\begin{pmatrix} 1, 2, \dots, n \\ i_1, i_2, \dots, i_n \end{pmatrix}$ — произвольная подста-

повка. Эти необходимые условия, однако, являются также и достаточными, как это явствует из следующей теоремы.

ОСНОВНАЯ ТЕОРЕМА. *Всякая система функций распределения F_{v_1, v_2, \dots, v_n} , удовлетворяющих условиям (2) и (3), определяет вероятностную функцию P на \mathcal{F}^n , которая удовлетворяет всем аксиомам I—V. Эта вероятностная функция P может быть продолжена (по теореме о продолжении) также и на $\mathcal{F}(\mathcal{F}^n)$.*

Доказательство. Итак, пусть даны функции распределения F_{v_1, v_2, \dots, v_n} , удовлетворяющие общим условиям главы второй (§ 3, III) и условиям (2) и (3). Всякая функция распределения F_{v_1, v_2, \dots, v_n} однозначно определяет соответствующую вероятностную функцию P_{v_1, v_2, \dots, v_n} для всех борелевских множеств из R^n (ср. § 3). В дальнейшем мы будем рассматривать только борелевские множества R^n и борелевские цилиндрические множества в Ω .

Для всякого цилиндрического множества

$$A = \pi_{v_1, v_2, \dots, v_n}^{-1}(A')$$

полагаем

$$P(A) = P_{v_1, v_2, \dots, v_n}(A'). \quad (4)$$

Так как одно и то же цилиндрическое множество A может быть определено через различные множества A' , то нужно сначала доказать, что формула (4) дает всегда одинаковое значение для $P(A)$.

Пусть $(x_{v_1}, x_{v_2}, \dots, x_{v_n})$ — конечная система случайных величин x_v . Исходя из вероятностной функции P_{v_1, v_2, \dots, v_n} этих случайных величин, мы можем, согласно правилам § 2, определить вероятностную функцию $P_{v_{i_1}, v_{i_2}, \dots, v_{i_k}}$ каждой подсистемы $(x_{v_{i_1}}, x_{v_{i_2}}, \dots, x_{v_{i_k}})$. Равенства (2) и (3) имеют своим следствием, что эта определенная по § 2 вероятностная функция совпадает с а priori заданной функцией $P_{v_{i_1}, v_{i_2}, \dots, v_{i_k}}$. Мы теперь пред-

положим, что цилиндрическое множество A определяется через

$$A = \pi_{v_{i_1}, v_{i_2}, \dots, v_{i_k}}^{-1}(A')$$

и одновременно через

$$A = \pi_{v_{j_1}, v_{j_2}, \dots, v_{j_m}}^{-1}(A''),$$

причем все случайные величины x_{v_i} и x_{v_j} принадлежат системе $(x_{v_1}, x_{v_2}, \dots, x_{v_n})$, что, очевидно, не является существенным ограничением. Условия

$$(x_{v_{i_1}}, x_{v_{i_2}}, \dots, x_{v_{i_k}}) \in A'$$

и

$$(x_{v_{j_1}}, x_{v_{j_2}}, \dots, x_{v_{j_m}}) \in A''$$

равнозначны. Следовательно,

$$\begin{aligned} P_{v_{i_1}, v_{i_2}, \dots, v_{i_k}}(A') &= P_{v_1, v_2, \dots, v_n} \{(x_{v_{i_1}}, x_{v_{i_2}}, \dots, x_{v_{i_k}}) \in A'\} = \\ &= P_{v_1, v_2, \dots, v_n} \{(x_{v_{j_1}}, x_{v_{j_2}}, \dots, x_{v_{j_m}}) \in A''\} = P_{v_{j_1}, v_{j_2}, \dots, v_{j_m}}(A''), \end{aligned}$$

что доказывает наше утверждение относительно однозначности определения $P(A)$.

Докажем теперь, что поле вероятностей $(\Omega, \mathcal{F}^{\mathcal{N}}, P)$ удовлетворяет всем аксиомам I—V. Аксиома I утверждает только, что $\mathcal{F}^{\mathcal{N}}$ должна быть алгеброй; это обстоятельство (за исключением требования $\Omega \in \mathcal{F}^{\mathcal{N}}$) было уже доказано выше. Далее, при любом $v \in \mathcal{N}$

$$\Omega = \pi_v^{-1}(R^1),$$

$$P(\Omega) = P_v(R^1) = 1,$$

что доказывает применимость аксиом I и II. Наконец, из определения $P(A)$ непосредственно следует, что $P(A)$ неотрицательна (аксиома II).

Несколько сложнее доказывается применимость аксиомы IV. Для этой цели рассмотрим два цилиндрических множества

$$A = \pi_{v_{i_1}, v_{i_2}, \dots, v_{i_k}}^{-1}(A')$$

и

$$B = \pi_{v_{j_1}, v_{j_2}, \dots, v_{j_m}}^{-1}(B').$$

При этом мы предположим, что все величины x_{v_i} и x_{v_j} принадлежат к объемлющей их конечной системе $(x_{v_1}, x_{v_2}, \dots, x_{v_n})$. Если множества A и B не пересекаются, то соотношения

$$(x_{v_{i_1}}, \dots, x_{v_{i_k}}) \in A',$$

$$(x_{v_{j_1}}, \dots, x_{v_{j_m}}) \in B'$$

несовместны. Следовательно,

$$\begin{aligned} P(A + B) &= \\ &= P_{v_1, v_2, \dots, v_n} \{ (x_{v_{i_1}}, x_{v_{i_2}}, \dots, x_{v_{i_k}}) \in A' \text{ или} \\ &\quad (x_{v_{j_1}}, x_{v_{j_2}}, \dots, x_{v_{j_m}}) \in B' \} = \\ &= P_{v_1, v_2, \dots, v_n} \{ (x_{v_{i_1}}, x_{v_{i_2}}, \dots, x_{v_{i_k}}) \in A' \} + \\ &+ P_{v_1, v_2, \dots, v_n} \{ (x_{v_{j_1}}, x_{v_{j_2}}, \dots, x_{v_{j_m}}) \in B' \} = P(A) + P(B). \end{aligned}$$

Остается проверить аксиому V. Пусть

$$A_1 \supseteq A_2 \supseteq \dots \supseteq A_n \supseteq \dots$$

— убывающая последовательность цилиндрических множеств, удовлетворяющая условию

$$\lim_n P(A_n) = L > 0.$$

Мы докажем, что пересечение всех множеств A_n не пусто. Можно без существенного ограничения постановки вопроса предположить, что в определение n

первых цилиндрических множеств A_k входят только n первых координат x_{v_k} последовательности

$$x_{v_1}, x_{v_2}, \dots, x_{v_k}, \dots,$$

т. е. что

$$A_n = \pi_{v_1, v_2, \dots, v_n}^{-1}(B_n).$$

Положим для краткости

$$P_n(B) = P_{v_1, v_2, \dots, v_n}(B).$$

Тогда, очевидно,

$$P_n(B_n) = P(A_n) \geq L > 0.$$

В каждом множестве B_n можно найти замкнутое ограниченное множество U_n такое, что

$$P_n(B_n \setminus U_n) \leq \frac{\varepsilon}{2^n}.$$

Из этого неравенства для множества

$$V_n = \pi_{v_1, v_2, \dots, v_n}^{-1}(U_n)$$

следует неравенство

$$P(A_n \setminus V_n) \leq \frac{\varepsilon}{2^n}. \quad (5)$$

Пусть далее

$$W_n = V_1 V_2 \dots V_n.$$

Из (5) следует, что

$$P(A_n \setminus W_n) \leq \varepsilon,$$

и так как

$$W_n \subseteq V_n \subseteq A_n,$$

то

$$P(W_n) \geq P(A_n) - \varepsilon \geq L - \varepsilon.$$

Если ε достаточно мало, то $P(W_n) > 0$ и множество W_n не пусто. Мы выберем теперь в каждом множестве

W_n точку $\omega^{(n)}$ с координатами $x_{\nu}^{(n)}$. Всякая точка $\omega^{(n+p)}$, $p = 0, 1, 2, \dots$, принадлежит множеству V_n , следовательно,

$$(x_{\nu_1}^{(n+p)}, x_{\nu_2}^{(n+p)}, \dots, x_{\nu_n}^{(n+p)}) = \pi_{\nu_1, \nu_2, \dots, \nu_n}^{-1}(\omega^{(n+p)}) \in U_n.$$

Так как множества U_n ограничены, то из последовательности $\{\omega^{(n)}\}$ можно (диагональный метод) выбрать подпоследовательность

$$\omega^{(n_1)}, \omega^{(n_2)}, \dots, \omega^{(n_i)}, \dots,$$

для которой соответствующие координаты $x_{\nu_k}^{(n_i)}$ стремятся при любом k к определенному пределу x_k . Пусть, наконец, ω — точка множества Ω с координатами

$$\begin{aligned} x_{\nu_k} &= x_k, \\ x_{\nu} &= 0 \quad (\nu \neq \nu_k, k = 1, 2, \dots). \end{aligned}$$

Точка (x_1, x_2, \dots, x_k) как предел последовательности $(x_1^{(n_i)}, x_2^{(n_i)}, \dots, x_k^{(n_i)})$, $i = 1, 2, \dots$, принадлежит множеству U_k . Следовательно, ω принадлежит множеству

$$A_k \subseteq V_k = \pi_{\nu_1, \nu_2, \dots, \nu_k}^{-1}(U_k)$$

при любом k , а следовательно, и к пересечению

$$A = \bigcap_k A_k.$$

Теорема доказана.

§ 5. Эквивалентные случайные величины, разные виды сходимости

С этого параграфа мы рассматриваем исключительно *борелевские* поля вероятностей (Ω, \mathcal{F}, P) . Это не делает, как было уже разъяснено в § 2 главы второй, никакого существенного ограничения для наших исследований.

О п р е д е л е н и е I. Две случайные величины ξ и η называются *эквивалентными*, если вероятность соотношения $\xi \neq \eta$ равна нулю.

Ясно, что две эквивалентные случайные величины имеют одну и ту же вероятностную функцию:

$$P_{\xi}(A) = P_{\eta}(A).$$

Следовательно, функции распределения F_{ξ} и F_{η} также совпадают. Во многих вопросах теории вероятностей можно заменять какую-либо случайную величину любой эквивалентной ей величиной.

Пусть теперь

$$\xi_1, \xi_2, \dots, \xi_n, \dots \quad (1)$$

— последовательность случайных величин. Рассмотрим множество A всех элементарных событий ω , для которых последовательность (1) сходится. Если через $A_{np}^{(m)}$ обозначить множество всех ω , для которых выполняются все неравенства

$$|\xi_{n+k} - \xi_n| < \frac{1}{m}, \quad k = 1, 2, \dots, p,$$

то непосредственно получаем, что

$$A = \bigcap_m \bigcup_n \bigcap_p A_{np}^{(m)}. \quad (2)$$

Согласно § 3 множества $A_{np}^{(m)}$ принадлежат всегда σ -алгебре множеств \mathcal{F} . Соотношение (2) показывает, что множество A также принадлежит \mathcal{F} . Следовательно, всегда имеет вполне определенный смысл говорить о вероятности сходимости последовательности случайных величин.

Пусть теперь вероятность $P(A)$ множества A равна единице. Тогда мы утверждаем, что последовательность (1) *сходится с вероятностью единица* к некоторой случайной величине ξ , которая определяется однозначно с точностью до эквивалентности.

Для построения этой случайной величины полагаем

$$\xi = \lim_n \xi_n$$

на A и $\xi = 0$ вне A . Нам надо доказать, что ξ — случайная величина, т. е. что множество $A(x)$ элементов ω , для которых $\xi < x$, принадлежит алгебре \mathcal{F} . Но

$$A(x) = A \bigcup_n \bigcap_p \{\omega: \xi_{n+p} < x\}$$

в случае $x \leq 0$ и

$$A(x) = A \bigcup_n \bigcap_p \{\omega: \xi_{n+p} < x\} + \bar{A}$$

в противоположном случае, откуда непосредственно следует наше утверждение.

О п р е д е л е н и е II. Если вероятность сходимости последовательности ξ_1, ξ_2, \dots к ξ равна единице, то мы говорим, что эта последовательность *почти наверное* сходится к ξ .

Однако для теории вероятностей, пожалуй, еще важнее другой вид сходимости.

О п р е д е л е н и е III. Последовательность случайных величин ξ_1, ξ_2, \dots сходится *по вероятности* к случайной величине ξ , если для любого $\varepsilon > 0$ вероятность

$$P \{ |\xi_n - \xi| > \varepsilon \} \rightarrow 0,$$

при $n \rightarrow \infty$ ¹⁾.

I. Если последовательность (1) сходится по вероятности одновременно к ξ и к ξ' , то ξ и ξ' эквивалентны.

В самом деле,

$$\begin{aligned} P \left\{ |\xi - \xi'| > \frac{1}{m} \right\} &\leq P \left\{ |\xi - \xi_n| > \frac{1}{2m} \right\} + \\ &+ P \left\{ |\xi' - \xi_n| > \frac{1}{2m} \right\}. \end{aligned}$$

¹⁾ Это понятие восходит в основном еще к Бернулли, однако в полной общности было введено Е. Е. Слуцким (ср. Slutsky [1]).

Так как последние вероятности при достаточно большом n сколько угодно малы, то отсюда следует, что

$$P \left\{ |\xi - \xi'| > \frac{1}{m} \right\} = 0,$$

и, значит,

$$P \{ \xi \neq \xi' \} \leq \sum_m P \left\{ |\xi - \xi'| > \frac{1}{m} \right\} = 0.$$

II. Если последовательность (1) почти наверное сходится к ξ , то она сходится к ξ также и по вероятности.

Пусть A — множество сходимости последовательности (1). Тогда

$$\begin{aligned} 1 = P(A) &\leq \lim_n P \{ |\xi_{n+p} - \xi| < \varepsilon, p = 0, 1, 2, \dots \} \leq \\ &\leq \lim_n P \{ |\xi_n - \xi| < \varepsilon \}, \end{aligned}$$

откуда следует сходимость по вероятности.

III. Для сходимости по вероятности последовательности (1) необходимо и достаточно следующее условие: для любого $\varepsilon > 0$ существует такое n , что для каждого $p > 0$ имеет место неравенство

$$P \{ |\xi_{n+p} - \xi_n| > \varepsilon \} < \varepsilon.$$

Пусть теперь $F(x)$, $F_1(x)$, $F_2(x)$, ... — функция распределения случайных величин ξ , ξ_1 , ξ_2 , ... Если последовательность ξ_1 , ξ_2 , ... сходится по вероятности к ξ , то функция распределения $F(x)$ однозначно определяется знанием функций $F_n(x)$, так как имеет место

Т е о р е м а. Если последовательность ξ_1 , ξ_2 , ... сходится по вероятности к ξ , то последовательность соответствующих функций распределения $F_n(x)$ сходится к функции распределения $F(x)$ случайной величины ξ в каждой точке непрерывности $F(x)$.

Д о к а з а т е л ь с т в о. Тот факт, что $F(x)$ действительно определяется через $\{F_n(x)\}$, следует из того, что $F(x)$ как непрерывная слева монотонная функция однозначно определяется ее значениями

в точках непрерывности ¹⁾. Для доказательства утверждения о сходимости мы предположим, что x является точкой непрерывности $F(x)$. Пусть $x' < x$. Тогда в случае $\xi < x'$, $\xi_n \geq x$ необходимо $|\xi_n - \xi| > x - x'$. Следовательно,

$$\lim_n P\{\xi < x', \xi_n \geq x\} = 0,$$

$$\begin{aligned} F(x') = P\{\xi < x'\} &\leq P\{\xi_n < x\} + P\{\xi < x', \xi_n \geq x\} = \\ &= F_n(x) + P\{\xi < x', \xi_n \geq x\}, \end{aligned}$$

$$F(x') \leq \liminf_n F_n(x) + \lim_n P\{\xi < x', \xi_n \geq x\},$$

$$F(x') \leq \liminf_n F_n(x). \quad (3)$$

Аналогично доказывается, что если $x'' > x$, то

$$F(x'') \geq \limsup_n F_n(x). \quad (4)$$

Так как $F(x')$ и $F(x'')$ при $x' \uparrow x$ и $x'' \downarrow x$ стремятся к $F(x)$, то из (3) и (4) следует, что

$$\lim_n F_n(x) = F(x).$$

Теорема доказана.

¹⁾ В самом деле, она имеет самое большое лишь счетное множество точек разрыва (ср. Л е б е г, Интегрирование и отыскание примитивных функций, 1934, стр. 70). Поэтому точки непрерывности лежат всюду плотно, и значение функции $F(x)$ в точке разрыва определяется как предел ее значений в лежащих слева точках непрерывности.

IV. МАТЕМАТИЧЕСКИЕ ОЖИДАНИЯ

§ 1. Абстрактные интегралы Лебега

Пусть ξ — случайная величина и A — множество из \mathcal{F}^1). образуем для положительного λ сумму

$$S_\lambda = \sum_{k=-\infty}^{+\infty} k\lambda P[\{k\lambda \leq \xi < (k+1)\lambda\} \cap A]. \quad (1)$$

Если этот ряд абсолютно сходится при любом λ , то при $\lambda \rightarrow 0$ S_λ стремится к определенному пределу, который по определению есть интеграл

$$\int_A \xi(\omega) P(d\omega). \quad (2)$$

В этой абстрактной форме понятие интеграла было введено Фреше ²⁾. В частности, оно оказывается необходимым для теории вероятностей. (В следующих параграфах читатель увидит, что обычное определение условного математического ожидания величины ξ при гипотезе A совпадает с точностью до постоянного множителя с определением интеграла (2).)

Мы дадим здесь краткий перечень важнейших свойств интегралов формы (2). Читатель найдет их доказательство в каждом учебнике по теории функций действительного переменного, хотя они большей

¹⁾ Как было упомянуто в § 5 главы третьей, мы рассматриваем в этой и во всех последующих главах только *борелевские* поля вероятностей.

²⁾ Fréchet, Sur l'intégrale d'une fonctionnel étendue à un ensemble abstrait, Bull. Soc. Math. France, т. 43 (1915), стр. 248.

частью проведены для случая, когда P является лебеговой мерой множеств в R^n . Перенесение этих доказательств на общий случай не представляет никакой новой математической задачи; они остаются большей частью дословно теми же ¹⁾.

I. Если случайная величина ξ интегрируема на A , то она интегрируема на каждом подмножестве A' из A , принадлежащем \mathcal{F} .

II. Если ξ интегрируема на A и A распадается на не более чем счетное число непересекающихся множеств A_n из \mathcal{F} , то

$$\int_A \xi(\omega) P(d\omega) = \sum_n \int_{A_n} \xi(\omega) P(d\omega).$$

III. Вместе с ξ интегрируем всегда и $|\xi|$. При этом

$$\left| \int_A \xi(\omega) P(d\omega) \right| \leq \int_A |\xi(\omega)| P(d\omega).$$

IV. Если для каждого ω выполняются неравенства $0 \leq \eta(\omega) \leq \xi(\omega)$, то вместе с $\xi(\omega)$ интегрируема также и величина $\eta(\omega)$ ²⁾, и при этом

$$\int_A \eta(\omega) P(d\omega) \leq \int_A \xi(\omega) P(d\omega).$$

V. Если $m \leq \xi(\omega) \leq M$, причем и m и M — две постоянные, то

$$mP(A) \leq \int_A \xi(\omega) P(d\omega) \leq M \cdot P(A).$$

VI. Если $\xi(\omega)$ и $\eta(\omega)$ интегрируемы, а K и L — две действительные постоянные, то случайная величина $K\xi(\omega) + L\eta(\omega)$ также интегрируема, и при

¹⁾ См. цитированную на стр. 11 книгу А. Н. Колмогорова и С. В. Фомина.

²⁾ При этом предполагается, что $\eta(\omega)$ — случайная величина, т. е. в терминологии общей теории интегрирования она измерима по отношению к \mathcal{F} .

этом

$$\int_A [K\xi(\omega) + L\eta(\omega)] P(d\omega) = K \int_A \xi(\omega) P(d\omega) + L \int_A \eta(\omega) P(d\omega).$$

VII. Если ряд

$$\sum_n \int_A |\xi_n(\omega)| P(d\omega)$$

сходится, то ряд

$$\sum_n \xi_n(\omega) = \xi(\omega)$$

сходится в каждой точке множества A с точностью до некоторого множества B такого, что $P(B) = 0$. Если вне множества $A \setminus B$ положить $\xi(\omega) = 0$, то

$$\int_A \xi(\omega) P(d\omega) = \sum_n \int_A \xi_n(\omega) P(d\omega).$$

VIII. Если $\xi(\omega)$ и $\eta(\omega)$ эквивалентны, $P\{\xi(\omega) \neq \eta(\omega)\} = 0$, то для каждого множества A из \mathcal{F}

$$\int_A \xi(\omega) P(d\omega) = \int_A \eta(\omega) P(d\omega). \quad (3)$$

IX. Если (3) имеет место для каждого множества A из \mathcal{F} , то $\xi(\omega)$ и $\eta(\omega)$ эквивалентны.

Из вышеупомянутого определения интеграла получается еще следующее свойство, которого нет в обычной теории Лебега.

X. Пусть P_1 и P_2 — две вероятностные функции, определенные на одной и той же σ -алгебре \mathcal{F} , $P = P_1 + P_2$ и $\xi = \xi(\omega)$ интегрируема на множестве A относительно P_1 и P_2 . Тогда

$$\int_A \xi(\omega) P(d\omega) = \int_A \xi(\omega) P_1(d\omega) + \int_A \xi(\omega) P_2(d\omega).$$

XI. Всякая ограниченная случайная величина интегрируема.

§ 2. Абсолютные и условные математические ожидания

Пусть $\xi = \xi(\omega)$ — интегрируемая случайная величина. Интеграл

$$M\xi = \int_{\Omega} \xi(\omega) P(d\omega)$$

называют в теории вероятностей *математическим ожиданием* величины ξ .

Из свойств III, IV, V, VI, VII, VIII, IX следует, что

$$\text{I. } |M\xi| \leq M|\xi|;$$

$$\text{II. } M\eta \leq M\xi, \text{ если } 0 \leq \eta(\omega) \leq \xi(\omega);$$

$$\text{III. } \inf_{\omega} \xi(\omega) \leq M\xi \leq \sup_{\omega} \xi(\omega);$$

$$\text{IV. } M(K\xi + L\eta) = KM\xi + LM\eta;$$

$$\text{V. } M\left(\sum_n \xi_n\right) = \sum_n M\xi_n, \text{ если ряд } \sum_n M|\xi_n| \text{ сходится;}$$

$$\text{VI. Если } \xi \text{ и } \eta \text{ эквивалентны, то}$$

$$M\xi = M\eta;$$

VII. Всякая ограниченная случайная величина имеет математическое ожидание.

По определению интеграла

$$\begin{aligned} M\xi &= \lim_{\lambda \rightarrow 0} \sum_{k=-\infty}^{+\infty} k\lambda P\{k\lambda \leq \xi(\omega) < (k+1)\lambda\} = \\ &= \lim_{\lambda \rightarrow 0} \sum_{k=-\infty}^{+\infty} k\lambda [F_{\xi}((k+1)\lambda) - F_{\xi}(k\lambda)]. \end{aligned}$$

Вторая строка есть не что иное, как обычное определение интеграла Стильеса $\int_{-\infty}^{\infty} xF_{\xi}(dx)$. Поэтому

$$M\xi = \int_{-\infty}^{\infty} xF_{\xi}(dx). \quad (1)$$

Формула (1) может, следовательно, также служить определением математического ожидания $M\xi$.

Пусть теперь $\xi = \xi(\omega)$ — функция элементарного события ω , а η — случайная величина, определенная как однозначная функция от ξ : $\eta = \eta(\xi)$. Тогда

$$P\{k\lambda \leq \eta < (k+1)\lambda\} = P_{\xi}\{x: k\lambda \leq \eta(x) < (k+1)\lambda\},$$

где P_{ξ} — вероятностная функция ξ . Отсюда следует по определению интеграла, что

$$\int_{\Omega} \eta(\xi(\omega)) P(d\omega) = \int_X \eta(x) P_{\xi}(dx)$$

и, следовательно,

$$M\eta = \int_X \eta(x) P_{\xi}(dx), \quad (2)$$

где X означает множество всех возможных значений ξ .

В частности, когда $\xi = \xi(\omega)$ является случайной величиной, то

$$\begin{aligned} M\eta &= \int_{\Omega} \eta(\xi(\omega)) P(d\omega) = \int_R \eta(x) P_{\xi}(dx) = \\ &= \int_{-\infty}^{\infty} \eta(x) F_{\xi}(dx). \end{aligned} \quad (3)$$

Последний интеграл в формуле (3) является в случае непрерывности функции $\eta = \eta(x)$ обыкновенным интегралом Стильеса. Однако отметим при этом, что интеграл

$$\int_{-\infty}^{\infty} \eta(x) F_{\xi}(dx)$$

может существовать также и в случае отсутствия математического ожидания $M\eta$. Для существования $M\eta$

необходима и достаточна конечность интеграла ¹⁾

$$\int_{-\infty}^{\infty} |\eta(x)| F_{\xi}(dx).$$

Если $\xi = (\xi_1, \xi_2, \dots, \xi_n)$ — случайная точка пространства R^n , то вследствие (2)

$$M\eta = \int \dots \int_{R^n} \eta(x_1, x_2, \dots, x_n) P_{\xi_1, \xi_2, \dots, \xi_n}(dx_1, \dots, dx_n). \quad (4)$$

Мы уже видели, что если $P(B) > 0$, то условная вероятность $P(\cdot | B)$ обладает всеми свойствами вероятностной функции (при закрепленном множестве B). Соответствующий интеграл

$$M(\xi | B) = \int_{\Omega} \xi(\omega) P(d\omega | B) \quad (5)$$

мы называем *условным математическим ожиданием* случайной величины $\xi = \xi(\omega)$ относительно события B . Так как

$$P(\bar{B} | B) = 0, \\ \int_{\bar{B}} \xi(\omega) P(d\omega | B) = 0,$$

то из (5) следует равенство

$$\begin{aligned} M(\xi | B) &= \int_{\Omega} \xi(\omega) P(d\omega | B) = \\ &= \int_B \xi(\omega) P(d\omega | B) + \int_{\bar{B}} \xi(\omega) P(d\omega | B) = \\ &= \int_B \xi(\omega) P(d\omega | B). \end{aligned}$$

Вспомнив, что в случае $A \subseteq B$ и $P(B) > 0$

$$P(A | B) = \frac{P(AB)}{P(B)} = \frac{P(A)}{P(B)},$$

¹⁾ Ср. V. G l i v e n k o, Sur les valeurs probables de fonctions, Rend. Accad. Lincei, т. 8 (1928), стр. 480—483.

мы получаем

$$M(\xi | B) = \frac{1}{P(B)} \int_B \xi(\omega) P(d\omega), \quad (6)$$

$$P(B) \cdot M(\xi | B) = \int_B \xi(\omega) P(d\omega). \quad (7)$$

Из (6) и из равенства

$$\int_{A+B} \xi(\omega) P(d\omega) = \int_A \xi(\omega) P(d\omega) + \int_B \xi(\omega) P(d\omega)$$

следует, наконец, что

$$M(\xi | A + B) = \frac{P(A)M(\xi | A) + P(B)M(\xi | B)}{P(A + B)}. \quad (8)$$

В частности, если $0 < P(A) < 1$, то имеет место формула

$$M\xi = P(A)M(\xi | A) + P(\bar{A})M(\xi | \bar{A}). \quad (9)$$

§ 3. Неравенство Чебышева

Пусть $f = f(x)$ — неотрицательная функция действительного аргумента x , которая при $x \geq a$ остается не меньше $b > 0$. Тогда для любой случайной величины $\xi = \xi(\omega)$

$$P\{\xi(\omega) \geq a\} \leq \frac{Mf(\xi)}{b} \quad (1)$$

(предполагается, что математическое ожидание $Mf(\xi)$ существует).

В самом деле,

$$\begin{aligned} Mf(\xi) &= \int_{\Omega} f(\xi(\omega)) P(d\omega) \geq \int_{\{\omega: \xi(\omega) \geq a\}} f(\xi(\omega)) P(d\omega) \geq \\ &\geq bP\{\xi(\omega) \geq a\}, \end{aligned}$$

откуда непосредственно следует (1).

Например, для любого положительного c

$$P\{\xi(\omega) \geq a\} \leq \frac{Me^{c\xi}}{e^{ca}}. \quad (2)$$

Пусть теперь $f = f(x)$ — неотрицательная, четная и при положительном x неубывающая функция. Тогда для каждой случайной величины $\xi = \xi(\omega)$ и при любом выборе постоянной $a > 0$ имеет место неравенство

$$P\{|\xi(\omega)| \geq a\} \leq \frac{Mf(\xi)}{f(a)}. \quad (3)$$

В частности,

$$P\{|\xi - M\xi| \geq a\} \leq \frac{Mf(\xi - M\xi)}{f(a)}. \quad (4)$$

Особо важным является случай $f(x) = x^2$. В этом случае из (3) получаем *неравенство Чебышева*

$$P\{|\xi(\omega)| \geq a\} \leq \frac{M\xi^2}{a^2}. \quad (5)$$

Из (4) находим также, что

$$P\{|\xi - M\xi| \geq a\} \leq \frac{M(\xi - M\xi)^2}{a^2} = \frac{D\xi}{a^2}. \quad (6)$$

Величину

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

называют *дисперсией* случайной величины ξ . Легко подсчитать, что

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

Если функция $f(x)$ ограничена,

$$|f(x)| \leq K,$$

то $P\{|\xi(\omega)| \geq a\}$ можно оценить также и снизу. В самом деле,

$$\begin{aligned} Mf(\xi) &= \int_{\Omega} f(\xi(\omega)) P(d\omega) = \\ &= \int_{\{\omega: |\xi(\omega)| < a\}} f(\xi(\omega)) P(d\omega) + \int_{\{\omega: |\xi(\omega)| \geq a\}} f(\xi(\omega)) P(d\omega) \leq \\ &\leq f(a) P\{|\xi(\omega)| < a\} + KP\{|\xi(\omega)| \geq a\} \leq \\ &\leq f(a) + K \cdot P\{|\xi(\omega)| \geq a\} \end{aligned}$$

и, следовательно,

$$P\{|\xi(\omega)| \geq a\} \geq \frac{Mf(\xi) - f(a)}{K}. \quad (7)$$

Если вместо ограниченности функции $f(x)$ потребовать ограниченности самой случайной величины $\xi = \xi(\omega)$,

$$|\xi(\omega)| \leq M,$$

то $f(\xi(\omega)) \leq f(M)$, и вместо (7) мы получаем формулу

$$P\{|\xi(\omega)| \geq a\} \geq \frac{Mf(\xi) - f(a)}{f(M)}. \quad (8)$$

В случае $f(x) = x^2$ из (8) находим, что

$$P\{|\xi(\omega)| \geq a\} \geq \frac{M\xi^2 - a^2}{M^2}. \quad (9)$$

§ 4. Некоторые признаки сходимости

Пусть

$$\xi_1, \xi_2, \dots, \xi_n, \dots \quad (1)$$

— последовательность случайных величин, и $f = f(x)$ — неотрицательная четная и при положительном x монотонно возрастающая функция¹⁾. Тогда справедливы следующие утверждения.

I. Для сходимости последовательности (1) по вероятности достаточно следующее условие: для каждого $\varepsilon > 0$ существует такое n , что для каждого $p > 0$ справедливо неравенство

$$Mf(\xi_{n+p} - \xi_n) < \varepsilon. \quad (2)$$

II. Для сходимости по вероятности последовательности (1) к случайной величине ξ достаточно условие

$$\lim_{n \rightarrow +\infty} Mf(\xi_n - \xi) = 0. \quad (3)$$

¹⁾ Следовательно, $f(x) > 0$, если $x \neq 0$.

III. Если $f(x)$ ограничена, непрерывна и $f(0) = 0$, то условия I и II являются также и необходимыми.

IV. Если $f(x)$ непрерывна, $f(0) = 0$ и все случайные величины

$$\xi, \xi_1, \xi_2, \dots$$

ограничены в своей совокупности, то условия I и II являются также и необходимыми.

Из II и IV следует, в частности,

V. Для сходимости по вероятности последовательности (1) к ξ достаточным является условие

$$\lim_{n \rightarrow \infty} M(\xi_n - \xi)^2 = 0. \quad (4)$$

Если при этом ξ, ξ_1, ξ_2, \dots ограничены в своей совокупности, то это условие является и необходимым.

Доказательство утверждений I—IV см. в работах Slutsky [4] и Fréchet [1]. Впрочем, эти теоремы следуют почти непосредственно из формул (3) и (8) предшествующего параграфа.

§ 5. Дифференцирование и интегрирование математических ожиданий по параметру

Пусть каждому элементарному событию ω поставлена в соответствие определенная вещественная функция $\xi_t(\omega)$ действительного переменного t .

О п р е д е л е н и е. Мы говорим, что $\xi = \{\xi_t(\omega), -\infty < t < \infty\}$ есть *случайная функция*, если при каждом фиксированном t величина $\xi_t(\omega)$ является случайной величиной.

Возникает тогда вопрос, при каких условиях знак математического ожидания является переместительным со знаками интегрирования и дифференцирования. Две следующие теоремы могут, не исчерпывая всей проблемы, дать во многих простых случаях удовлетворительный ответ на этот вопрос.

Т е о р е м а I. Если для любого t математическое ожидание $M\xi_t(\omega)$ конечно, $\xi_t(\omega)$ дифференцируема (по t для всех ω) и производная $\xi'_t(\omega) = \frac{d\xi_t(\omega)}{dt}$ по аб-

солютному значению всегда меньше некоторой определенной постоянной, то

$$\frac{d}{dt} M\xi_t(\omega) = M\xi'_t(\omega).$$

Т е о р е м а II. Если $\xi_t(\omega)$ по абсолютному значению всегда остается меньше некоторой постоянной K и интегрируема по t в смысле Римана, то

$$\int_a^b M\xi_t(\omega) dt = M \left[\int_a^b \xi_t(\omega) dt \right],$$

если только $M\xi_t(\omega)$ интегрируемо в смысле Римана.

Д о к а з а т е л ь с т в о т е о р е м ы I. Прежде всего заметим, что $\xi'_t(\omega)$ как предел случайных величин

$$\frac{\xi_{t+h}(\omega) - \xi_t(\omega)}{h}, \quad h = 1, \frac{1}{2}, \dots, \frac{1}{n}, \dots$$

является случайной величиной. Так как $\xi'_t(\omega)$ ограничена, то существует математическое ожидание $M\xi'_t(\omega)$ (свойство VII математических ожиданий из § 2). Выберем теперь постоянное t и обозначим через A событие

$$\left\{ \omega: \left| \frac{\xi_{t+h}(\omega) - \xi_t(\omega)}{h} - \xi'_t(\omega) \right| > \varepsilon \right\}.$$

Вероятность $P(A)$ стремится к нулю при $h \rightarrow 0$ для любого $\varepsilon > 0$. Так как всегда

$$\left| \frac{\xi_{t+h}(\omega) - \xi_t(\omega)}{h} \right| \leq M, \quad |\xi'_t(\omega)| \leq M$$

и, кроме того, в случае $\omega \in \bar{A}$

$$\left| \frac{\xi_{t+h}(\omega) - \xi_t(\omega)}{h} - \xi'_t(\omega) \right| \leq \varepsilon,$$

то

$$\begin{aligned}
& \left| \frac{M\xi_{t+h}(\omega) - M\xi_t(\omega)}{h} - M\xi'_t(\omega) \right| \leq \\
& \leq M \left| \frac{\xi_{t+h}(\omega) - \xi_t(\omega)}{h} - \xi'_t(\omega) \right| = \\
& = P(A) M \left\{ \left| \frac{\xi_{t+h}(\omega) - \xi_t(\omega)}{h} - \xi'_t(\omega) \right| \middle| A \right\} + \\
& + P(\bar{A}) M \left\{ \left| \frac{\xi_{t+h}(\omega) - \xi_t(\omega)}{h} - \xi'_t(\omega) \right| \middle| \bar{A} \right\} \leq \\
& \leq 2M \cdot P(A) + \varepsilon.
\end{aligned}$$

Поскольку $\varepsilon > 0$ можно выбрать произвольно, а вероятность $P(A)$ сколь угодно мала при достаточно малом h , то

$$\frac{d}{dt} M\xi_t(\omega) = \lim_{h \rightarrow 0} \frac{M\xi_{t+h}(\omega) - M\xi_t(\omega)}{h} = M\xi'_t(\omega),$$

что и требовалось доказать.

Доказательство теоремы II. Пусть

$$S_n = \frac{1}{n} \sum_{k=1}^n \xi_{t+kh}(\omega), \quad h = \frac{b-a}{n}.$$

Так как S_n сходится к $S = \int_a^b \xi_t(\omega) dt$, то можно для любого $\varepsilon > 0$ выбрать такое N , что для всех $n \geq N$

$$P\{|S_n - S| > \varepsilon\} < \varepsilon.$$

Если положить $A = \{|S_n - S| > \varepsilon\}$ и

$$S_n^* = \frac{1}{n} \sum_{k=1}^n M\xi_{t+kh}(\omega) = MS_n,$$

то

$$\begin{aligned}
|S_n^* - MS| &= |M(S_n - S)| \leq M|S_n - S| = \\
&= P(A) M\{|S_n - S| \mid A\} + P(\bar{A}) M\{|S_n - S| \mid \bar{A}\} \leq \\
&\leq 2KP(A) + \varepsilon \leq (2K + 1)\varepsilon.
\end{aligned}$$

Следовательно, S_n^* стремится к MS , откуда следует равенство

$$\int_a^b M\xi_t(\omega) dt = \lim_n S_n^* = MS.$$

Теорема II может быть без всяких новых затруднений обобщена для двойных и тройных интегралов. Мы дадим применение этой теоремы для одной задачи геометрической теории вероятностей.

Пусть $G = G(\omega)$ — квадратируемая область на плоскости, вид которой зависит от случая, т. е. пусть каждому элементарному событию $\omega \in \Omega$ поставлена в соответствие определенная квадратируемая область $G(\omega)$ на плоскости. Через S_G обозначим площадь области G , а через $P(x, y)$ — вероятность того, что точка (x, y) принадлежит области G . Тогда

$$MS_G = \iint P(x, y) dx dy.$$

Для доказательства достаточно заметить, что

$$S_G = \iint f_G(x, y) dx dy,$$

$$P(x, y) = M f_G(x, y),$$

где $f_G(x, y)$ — характеристическая функция области G [$f_G(x, y) = 1$ на G и $f_G(x, y) = 0$ вне G] ¹⁾.

¹⁾ Ср. A. Kolmogoroff und M. Leontowitsch, Zur Berechnung der mittleren Brownschen Fläche, Physik. Zeitschr. d. Sowjetunion, т. 4 (1933).

У. УСЛОВНЫЕ ВЕРОЯТНОСТИ И МАТЕМАТИЧЕСКИЕ ОЖИДАНИЯ

§ 1. Условные вероятности

В главе первой, § 6 мы определили условную вероятность $P(B | \mathfrak{A})$ события B относительно испытания \mathfrak{A} . При этом там было предположено, что \mathfrak{A} допускает лишь *конечное* число различных возможных исходов. Можно, однако, определить $P(B | \mathfrak{A})$ также и для случая испытания \mathfrak{A} с *бесконечным* множеством возможных исходов, т. е. для разложения множества на бесконечное число непересекающихся подмножеств. В частности, такое разложение получается, если рассматривать произвольную функцию $\xi = \xi(\omega)$ от ω и определить в качестве элементов разложения $\mathfrak{A} = \mathfrak{A}_\xi$ множества $\{\omega: \xi(\omega) = \text{const}\}$.

Условная вероятность $P(B | \mathfrak{A}_\xi)$ будет обозначаться также через $P(B | \xi)$. Любое разложение \mathfrak{A} множества Ω можно определить как разложение \mathfrak{A}_ξ , которое «индуцируется» функцией $\xi = \xi(\omega)$ от ω , если каждому ω поставлено в соответствие в качестве $\xi(\omega)$ то множество из разложения \mathfrak{A} , которое содержит ω .

Две функции ξ и η от ω определяют тогда и только тогда одно и то же разложение ($\mathfrak{A}_\xi = \mathfrak{A}_\eta$) множества Ω , если существует такое взаимно однозначное соответствие $y = f(x)$ между их значениями, при которых тождественно $\eta(\omega) = f(\xi(\omega))$. Читатель может легко показать, что определяемые ниже *случайные* величины $P(B | \xi)$ и $P(B | \eta)$ в этом случае совпадают; следовательно, в основном они определяются самим разложением $\mathfrak{A}_\xi = \mathfrak{A}_\eta$.

Для определения $P(B | \xi)$ можно применить следующее равенство:

$$P(B | \xi \in A) = M [P(B | \xi) | \xi \in A]. \quad (1)$$

Легко показать, что в случае *конечности* множества X возможных значений ξ равенство (1) выполняется при любом выборе множества A (причем $P(B | \xi)$ определяется согласно § 6 главы первой). В общем случае (для которого случайная величина $P(B | \xi)$ еще не определена) мы докажем, что *всегда существует* одна и, с точностью до эквивалентных величин, только одна случайная величина $P(B | \xi)$, определяемая как (борелевская) функция от ξ и удовлетворяющая при каждом A из \mathcal{F}_ξ , для которого $P_\xi(A) > 0$, уравнению (1). *Определенную таким образом (с точностью до эквивалентности) функцию $P(B | \xi)$ от ξ мы называем условной вероятностью B относительно ξ (или при данном ξ).* Значение $P(B | \xi)$ при $\xi = x$ мы будем обозначать $P(B | \xi = x)$.

Доказательство существования и единственности $P(B | \xi)$. Умножив (1) на $P\{\xi \in A\} = P_\xi(A)$, получаем слева

$$P\{\xi \in A\} P(B | \xi \in A) = P(B \cap \{\xi \in A\}) = \\ = P(B \cap \xi^{-1}(A)),$$

а справа

$$P\{\xi \in A\} M[P(B | \xi) | \xi \in A] = \\ = \int_{\{\omega: \xi \in A\}} P(B | \xi) P(d\omega) = \int_A P(B | \xi = x) P_\xi(dx).$$

Следовательно,

$$P(B \cap \xi^{-1}(A)) = \int_A P(B | \xi = x) P_\xi(dx). \quad (2)$$

Обратно, из (2) следует формула (1).

В случае $P_\xi(A) = 0$, при котором (1) не имеет смысла, равенство (2) тривиально. Требование (2), следовательно, эквивалентно (1).

По свойству IX интеграла (глава четвертая, § 1) случайные величины $\eta = \eta(\omega)$ с точностью до эквивалентности однозначно определяются по значениям интегралов

$$\int_B \eta(\omega) P(d\omega), \quad B \in \mathcal{F}.$$

Так как $P(B \mid \xi = x)$ — случайная величина, определенная на поле вероятностей $(X, \mathcal{F}_\xi, P_\xi)$, то отсюда следует, что формула (2) однозначно (с точностью до эквивалентности) определяет эту величину.

Нам остается доказать существование $P(B \mid \xi)$. Для этой цели применим следующую теорему Радона — Никодима¹⁾.

Т е о р е м а. Пусть \mathcal{U} — борелевская алгебра множеств в Y , P — определенная на (Y, \mathcal{U}) неотрицательная счетно-аддитивная функция множеств, а \tilde{P} — вторая определенная также на (Y, \mathcal{U}) , счетно-аддитивная функция множеств, причем из $P(A) = 0$ следует равенство $\tilde{P}(A) = 0$. Тогда существует измеримая (по отношению к \mathcal{U}) функция $\varphi = \varphi(y)$ (в теоретико-вероятностной терминологии — случайная величина), удовлетворяющая для каждого множества A из \mathcal{U} равенству

$$\tilde{P}(A) = \int_A \varphi(y) P(dy).$$

Для применения этой теоремы к нашему случаю остается доказать: 1° что $\tilde{P}(A) = P(B \cap \xi^{-1}(A))$ счетно-аддитивна на (X, \mathcal{F}_ξ) , 2° что из $\tilde{P}(A) \neq 0$ следует неравенство $P_\xi(A) > 0$.

Утверждение 2° следует из того, что

$$0 \leq P(B \cap \xi^{-1}(A)) \leq P(\xi^{-1}(A)) = P_\xi(A).$$

¹⁾ O. N i k o d y m, Sur une généralisation des integrales de M. J. Radon, Fund. Math., т. 15 (1930), 168 (théoreme III). См. также цит. выше книгу А. Н. К о л м о г о р о в а и С. В. Ф о м и н а.

Для доказательства утверждения 1° положим

$$A = \sum_n A_n.$$

Тогда

$$\xi^{-1}(A) = \sum_n \xi^{-1}(A_n)$$

и

$$B \cap \xi^{-1}(A) = \sum_n B \cap \xi^{-1}(A_n).$$

Так как P счетно-аддитивна, то

$$P(B \cap \xi^{-1}(A)) = \sum_n P(B \cap \xi^{-1}(A_n)),$$

что и требовалось доказать.

Из равенства (1) следует, в частности (если положить $A = X$), важная формула:

$$P(B) = M[P(B | \xi)]. \quad (3)$$

Теперь мы перейдем к доказательству двух следующих фундаментальных свойств условных вероятностей.

Т е о р е м а I. *Почти наверное*

$$0 \leq P(B | \xi) \leq 1. \quad (4)$$

Т е о р е м а II. *Если события B, B_1, \dots принадлежат \mathcal{F} и*

$$B = \sum_n B_n,$$

то почти наверное

$$P(B | \xi) = \sum_n P(B_n | \xi). \quad (5)$$

Эти два свойства $P(B | \xi)$ соответствуют двум характеристическим свойствам вероятностной функции

$P(B)$: всегда $0 \leq P(B) \leq 1$ и $P(B)$ счетно-аддитивна. Они позволяют перенести на условные вероятности $P(\cdot | \xi)$ многие дальнейшие существенные свойства абсолютных вероятностей $P(\cdot)$. Однако при этом не следует забывать, что $P(B | \xi)$ при фиксированном множестве B является величиной, определенной лишь с точностью до эквивалентности.

Доказательство теоремы I. Если мы предположим — в противоположность доказываемому утверждению, — что на множестве M с $P_{\xi}(M) > 0$ выполняется неравенство $P(B | \xi) \geq 1 + \varepsilon$, то по формуле (1)

$$P(B | \xi \in M) = M[P(B | \xi) | \xi \in M] \geq 1 + \varepsilon,$$

что, очевидно, невозможно. Так же доказывается, что почти наверное $P(B | \xi) \geq 0$.

Доказательство теоремы II. Из сходимости ряда

$$\sum_n M[P(B_n | \xi)] = \sum_n M[P(B_n | \xi)] = \sum_n P(B_n) = P(B)$$

следует по свойству V математических ожиданий (гл. IV, § 2), что ряд

$$\sum_n P(B_n | \xi)$$

сходится почти наверное. Так как ряд

$$\sum_n M[P(B_n | \xi) | \xi \in A] = \sum_n P(B_n | \xi \in A) = P(B | \xi \in A)$$

сходится при каждом выборе множества A такого, что $P_{\xi}(A) > 0$, то из упомянутого свойства V математических ожиданий следует, что для каждого множества A указанного вида имеет место соотношение

$$\begin{aligned} M\left[\sum_n P(B_n | \xi) | \xi \in A\right] &= \sum_n M[P(B_n | \xi) | \xi \in A] = \\ &= P(B | \xi \in A) = M[P(B | \xi) | \xi \in A], \end{aligned}$$

из которого непосредственно следует равенство (5).

В заключение этого параграфа укажем на два частных случая. Пусть $\xi(\omega) \equiv C$ — постоянная, тогда почти наверное $P(A | C) = P(A)$. Если положить, напротив, $\xi(\omega) = \omega$, то получим сразу, что $P(A | \omega)$ почти наверное равна единице на A и нулю на \bar{A} . Следовательно, $P(A | \omega)$ является *характеристической функцией* множества A .

§ 2. Объяснение одного парадокса Бореля

Пусть за основное множество Ω выбрано множество всех точек сферической поверхности. За \mathcal{F} мы примем совокупность всех борелевских множеств сферических поверхностей. Наконец, пусть $P(A)$ пропорциональна мере множества A . Выберем теперь две диаметрально противоположные точки в качестве полюсов. Тогда каждый меридианный круг однозначно определяется соответствующей географической долготой ψ ($0 \leq \psi < \pi$). Так как ψ изменяется только от 0 до π , т. е. мы рассматриваем полные меридианные круги (а не полуокружности), то и широта θ должна изменяться от $-\pi$ до π (а не от $-\frac{\pi}{2}$ до $\frac{\pi}{2}$). Борелем поставлена следующая задача: определить «условное распределение вероятностей» для широты θ , $-\pi \leq \theta < \pi$, при заданной долготе ψ . Легко подсчитать, что

$$P(\theta_1 \leq \theta < \theta_2 | \psi) = \frac{1}{4} \int_{\theta_1}^{\theta_2} |\cos \theta| d\theta$$

и, значит, распределение вероятностей для θ при заданном ψ *неравномерно*.

Если же теперь предположить, что условное распределение вероятностей для θ «при гипотезе, что ω лежит на меридианном круге», должно быть равномерным, то получается противоречие.

Это обстоятельство показывает, что понятие условной вероятности относительно изолированно заданной гипотезы, вероятность которой равна нулю, является недопустимым: только тогда мы получим на меридиан-

ном круге распределение вероятностей для θ , если будем рассматривать этот меридианный круг в качестве элемента разложения всей сферической поверхности на меридианные круги с заданными полюсами.

§ 3. Условные вероятности относительно случайной величины

Если $\xi = \xi(\omega)$ — случайная величина, то условную вероятность $P(B | \xi)$ можно определить также элементарным путём. В случае $P(B) = 0$ полагаем $P(B | \xi) = 0$. Пусть теперь $P(B) > 0$. Тогда формуле (2) из § 1 можно придать следующий вид:

$$P(B) P(\xi \in A | B) = \int_A P(B | \xi = x) P_\xi(dx) \quad (1)$$

или

$$P(B) P_\xi(A | B) = \int_A P(B | \xi = x) P_\xi(dx).$$

Отсюда получается непосредственно, что

$$P(B) F_\xi(a | B) = \int_{-\infty}^a P(B | \xi = x) F_\xi(dx). \quad (2)$$

Согласно одной теореме Лебега ¹⁾ из (2) следует, что

$$P(B | \xi = x) = P(B) \cdot \lim_{h \rightarrow 0} \frac{F_\xi(x+h | B) - F_\xi(x | B)}{F_\xi(x+h) - F_\xi(x)} \quad (3)$$

с точностью до множества H точек x такого, что $P_\xi(H) = 0$. Если рассматривать теперь формулу (3) как определение $P(B | \xi = x)$, причем в случае несуществования предела в правой части (3) положить $P(B | \xi = x) = 0$, то эта новая величина удовлетворяет всем требованиям § 1.

¹⁾ Ле б е г, Интегрирование и отыскание примитивных функций, 1934, стр. 245—246.

Если, кроме того, существуют плотности вероятностей $f_{\xi}(x)$ и $f_{\xi}(x|B)$ и если $f_{\xi}(x) > 0$, то формула (3) превращается в следующую:

$$P(B|\xi = x) = P(B) \frac{f_{\xi}(x|B)}{f_{\xi}(x)}. \quad (4)$$

Из формулы (3) следует также, что существование предела (3) и плотности вероятности $f_{\xi}(x)$ имеет своим следствием существование $f_{\xi}(x|B)$. При этом

$$P(B) f_{\xi}(x|B) \leq f_{\xi}(x). \quad (5)$$

Если $P(B) > 0$, то из (4) следует равенство:

$$f_{\xi}(x|B) = \frac{P(B|\xi = x) f_{\xi}(x)}{P(B)}. \quad (6)$$

В случае $f_{\xi}(x) = 0$, согласно (5), $f_{\xi}(x|B) = 0$ и, следовательно, (6) также верно. Если при этом распределение ξ абсолютно непрерывно, то

$$\begin{aligned} P(B) &= M[P(B|\xi)] = \int_{\Omega} P(B|\xi) P(d\omega) = \\ &= \int_{-\infty}^{\infty} P(B|\xi = x) F_{\xi}(dx) = \int_{-\infty}^{\infty} P(B|\xi = x) f_{\xi}(x) dx. \end{aligned} \quad (7)$$

Из (6) и (7) следует, что

$$f_{\xi}(x|B) = \frac{P(B|\xi = x) f_{\xi}(x)}{\int_{-\infty}^{\infty} P(B|\xi = x) f_{\xi}(x) dx}. \quad (8)$$

Это равенство дает нам так называемую *теорему Байеса для абсолютно непрерывных распределений*. Предположения, при которых эта теорема справедлива, следующие: $P(B|\xi = x)$ измерима в смысле Бореля и определена по формуле (3), распределение ξ абсолютно непрерывно (в точке x существует плотность вероятности $f_{\xi}(x)$).

§ 4. Условные математические ожидания

Пусть $\xi = \xi(\omega)$ — произвольная функция от ω , а $\eta = \eta(\omega)$ — случайная величина. Случайная величина $M(\eta | \xi)$, представляемая как функция от ξ и удовлетворяющая для любого множества A из \mathcal{F}_ξ с $P_\xi(A) > 0$ условию

$$M(\eta | \xi \in A) = M[M(\eta | \xi) | \xi \in A] \quad (1)$$

называется (в случае, если она существует) *условным математическим ожиданием случайной величины η при известном значении ξ* .

Умножив (1) на $P_\xi(A)$, получаем

$$\int_{\{\xi \in A\}} \eta(\omega) P(d\omega) = \int_{\{\xi \in A\}} M(\eta | \xi) P(d\omega). \quad (2)$$

Обратно, из (2) следует формула (1). В случае $P_\xi(A) = 0$, при котором (1) не имеет смысла, (2) тривиально. Так же как и в случае условных вероятностей (ср. § 1), доказывается, что $M(\eta | \xi)$ определяется через (2) однозначно (с точностью до эквивалентности).

Значение $M(\eta | \xi)$ при $\xi(\omega) = x$ мы обозначаем через $M(\eta | \xi = x)$. Заметим еще, что $M(\eta | \xi)$, так же как и $P(B | \xi)$, зависит только от разложения \mathcal{H}_ξ и может быть обозначено через $M(\eta | \mathcal{H}_\xi)$.

При определении $M(\eta | \xi)$ уже предположено существование $M\eta$ (если положить $A = X$, тогда $M(\eta | \xi \in A) = M\eta$). Мы теперь докажем, что существование $M\eta$ и достаточно для существования $M(\eta | \xi)$. Для этого достаточно показать, что по теореме Радона — Никодима (ср. § 1) функция множеств

$$Q(A) = \int_{\{\xi \in A\}} \eta(\omega) P(d\omega)$$

счетно-аддитивна на \mathcal{F}_ξ и абсолютно непрерывна относительно $P_\xi(A)$. Первое обстоятельство доказывается дословно так же, как и в случае условных вероятностей (ср. § 1). Второе требование — абсолютной непрерывности — заключается в том, что из $Q(A) \neq 0$

должно следовать неравенство $P_{\xi}(A) > 0$. Если мы предположим, что $P_{\xi}(A) = P(\xi \in A) = 0$, то ясно, что

$$Q(A) = \int_{\{\xi \in A\}} \eta(\omega) P(d\omega) = 0.$$

Таким образом, наше второе требование также выполнено.

Если в равенстве (1) положить $A = X$, то получится формула

$$M\eta = M[M(\eta | \xi)]. \quad (3)$$

Далее доказывается, что почти наверное

$$M[a\eta + b\xi | \xi] = a M(\eta | \xi) + b M(\xi | \xi), \quad (4)$$

где a и b — две произвольные постоянные, $M|\eta| < \infty$, $M|\xi| < \infty$. (Проведение доказательств предоставляется читателю.)

Если ξ и ζ — две функции элементарного события ω , то пара (ξ, ζ) может также рассматриваться как функция ω . Имеет место следующее важное равенство:

$$M[M(\eta | \xi, \zeta) | \xi] = M(\eta | \xi). \quad (5)$$

В самом деле, $M(\eta | \xi)$ определяется через соотношение

$$M(\eta | \xi \in A) = M[M(\eta | \xi) | \xi \in A].$$

Следовательно, нужно доказать, что величина $M[M(\eta | \xi, \zeta) | \xi]$ удовлетворяет равенству

$$M(\eta | \xi \in A) = M\{M[M(\eta | \xi, \zeta) | \xi] | \xi \in A\}. \quad (6)$$

Из определения $M(\eta | \xi, \zeta)$ следует, что

$$M(\eta | \xi \in A) = M[M(\eta | \xi, \zeta) | \xi \in A]. \quad (7)$$

Из определения $M[M(\eta | \xi, \zeta) | \xi]$ следует далее, что

$$\begin{aligned} M[M(\eta | \xi, \zeta) | \xi \in A] &= \\ &= M\{M[M(\eta | \xi, \zeta) | \xi] | \xi \in A\}. \end{aligned} \quad (8)$$

Равенство (6) является следствием равенств (7) и (8), чем наше утверждение доказано.

Если положить $\eta(\omega)$ равным единице на B и нулю вне B , то

$$M(\eta | \xi) = P(B | \xi),$$

$$M(\eta | \xi, \zeta) = P(B | \xi, \zeta).$$

В этом случае из формулы (5) получается формула

$$M(P(B | \xi, \zeta) | \xi) = P(B | \xi). \quad (9)$$

Условные математические ожидания можно определить также и непосредственно через соответствующие условные вероятности.

Для этой цели рассматриваются следующие суммы:

$$S_\lambda(\xi) = \sum_{k=-\infty}^{\infty} k\lambda P\{k\lambda \leq \eta < (k+1)\lambda | \xi\} \left(= \sum_{k=-\infty}^{\infty} R_k \right). \quad (10)$$

Если $M\eta$ существует, то ряд (10) сходится почти наверное. В самом деле, по формуле (3), § 1

$$M | R_k | = k\lambda P\{k\lambda \leq \eta < (k+1)\lambda\},$$

а сходимость ряда

$$\sum_{k=-\infty}^{\infty} |k\lambda| P\{k\lambda \leq \eta < (k+1)\lambda\} = \sum_{k=-\infty}^{\infty} M | R_k |$$

является необходимым условием существования $M\eta$ (ср. главу четвертую, § 1). Из этой сходимости следует, что ряд (10) сходится почти наверное (ср. главу четвертую, § 2, V). Далее доказываем, так же как и в теории интеграла Лебега, что из сходимости (10) при каком-либо λ следует сходимость при всяком λ и что в случае сходимости ряда (10) $S_\lambda(\xi)$ стремится к определенному пределу при $\lambda \rightarrow 0^1$). Можно тогда в порядке определе-

¹⁾ При этом мы рассматриваем только счетную последовательность значений λ : в этом случае все вероятности $P\{k\lambda \leq \eta < (k+1)\lambda | \xi\}$ определены почти наверное для всех этих значений.

ния положить

$$M(\eta | \xi) = \lim_{\lambda \rightarrow 0} S_{\lambda}(\xi). \quad (11)$$

Для того чтобы доказать, что определенное соотношением (11) условное ожидание $M(\eta | \xi)$ удовлетворяет поставленным прежде требованиям, нужно только убедиться в том, что определенная, согласно (11), величина $M(\eta | \xi)$ удовлетворяет равенству (1). Это доказательство протекает следующим образом:

$$\begin{aligned} M[M(\eta | \xi) | \xi \in A] &= \lim_{\lambda \rightarrow 0} M[S_{\lambda}(\xi) | \xi \in A] = \\ &= \lim_{\lambda \rightarrow 0} \sum_{k=-\infty}^{\infty} k\lambda P[k\lambda \leq \eta < (k+1)\lambda | \xi \in A] = M[\eta | \xi \in A]. \end{aligned}$$

Перестановка знаков математического ожидания и предела допустима при этой выкладке, так как $S_{\lambda}(\xi)$ сходится равномерно к $M(\eta | \xi)$ при $\lambda \rightarrow 0$ (простое следствие свойства у математических ожиданий из § 2). Перестановка знаков математического ожидания и суммирования также оправдана, так как ряд

$$\begin{aligned} \sum_{k=-\infty}^{\infty} M\{|k\lambda| P(k\lambda \leq \eta < (k+1)\lambda | \xi) | \xi \in A\} &= \\ &= \sum_{k=-\infty}^{\infty} |k\lambda| P\{k\lambda \leq \eta < (k+1)\lambda | \xi \in A\} \end{aligned}$$

сходящийся (непосредственное применение свойства V математических ожиданий).

Вместо (11) можно написать

$$M(\eta | \xi) = \int_{\Omega} \eta(\omega) P(d\omega | \xi). \quad (12)$$

Не следует, однако, при этом забывать, что (12) не является интегралом в смысле § 1 главы четвертой, так что (12) — только символическое обозначение.

Если ξ — случайная величина, то функцию от ξ и y

$$F_{\eta}(y | \xi) = P(\eta < y | \xi)$$

мы называем условной функцией распределения η при известном ξ . Функция $F_\eta(y | \xi)$ определена почти наверное при всяком y . Если $y_1 \leq y_2$, то почти наверное

$$F_\eta(y_1 | \xi) \leq F_\eta(y_2 | \xi).$$

Из (11) и (10) следует, что почти наверное

$$M(\eta | \xi) = \lim_{\lambda \rightarrow 0} \sum_{k=-\infty}^{\infty} k\lambda [F_\eta((k+1)\lambda | \xi) - F_\eta(k\lambda | \xi)]. \quad (13)$$

Это обстоятельство можно символически выразить формулой

$$M(\eta | \xi) = \int_{-\infty}^{\infty} y F_\eta(dy | \xi). \quad (14)$$

С помощью нового определения математического ожидания (10), (11) легко доказать, что для действительной функции $f(\xi)$ с $M | f(\xi) \eta | < \infty$ справедливо (почти наверное) равенство

$$M [f(\xi) \eta | \xi] = f(\xi) M [\eta | \xi].$$

VI. НЕЗАВИСИМОСТЬ. ЗАКОН БОЛЬШИХ ЧИСЕЛ

§ 1. Независимость

О п р е д е л е н и е I. Две функции $\xi = \xi(\omega)$ и $\eta = \eta(\omega)$ называются *независимыми*, если для любых двух множеств A из \mathcal{F}_ξ и B из \mathcal{F}_η справедливо следующее равенство:

$$P(\xi \in A, \eta \in B) = P(\xi \in A) P(\eta \in B). \quad (1)$$

Если множества X и Y состоят лишь из конечного числа элементов:

$$X = x_1 + \dots + x_n,$$

$$Y = y_1 + \dots + y_m,$$

то наше определение независимости ξ и η совпадает, согласно § 5 главы первой, с определением независимости разложений

$$\Omega = \sum_{k=1}^n \{\omega : \xi(\omega) = x_k\},$$

$$\Omega = \sum_{k=1}^m \{\omega : \eta(\omega) = y_k\}.$$

Для независимости ξ и η *необходимо и достаточно* следующее условие: при любом выборе множества A из \mathcal{F}_ξ почти наверное выполнено равенство

$$P(\xi \in A \mid \eta) = P(\xi \in A). \quad (2)$$

В самом деле, в случае, когда $P_\eta(B) = 0$, оба равенства (1) и (2) удовлетворяются. Следовательно, до-

статочно доказать их эквивалентность в случае $P_\eta(B) > 0$. В этом случае (1) эквивалентно соотношению

$$P(\xi \in A \mid \eta \in B) = P(\xi \in A) \quad (3)$$

и, значит, соотношению

$$M[P(\xi \in A \mid \eta) \mid \eta \in B] = P(\xi \in A). \quad (4)$$

С другой стороны, видно, что из (2) следует равенство (4). Обратно, так как $P(\xi \in A \mid \eta)$ однозначно определяется из (4) с точностью до множеств вероятности нуль, то из (4) почти наверное следует равенство (2).

О п р е д е л е н и е II. Пусть Σ — множество функций $\xi_v(\omega)$, где $v \in N$. Эти функции называются *независимыми в совокупности*, если выполняется следующее условие: пусть Σ' и Σ'' — два непересекающихся подмножества из Σ , A' (соответственно A'') — множество из \mathcal{F} , определяемое соотношением между $\xi_v(\omega)$ из Σ' (соответственно Σ''); тогда

$$P(A' \cap A'') = P(A') P(A'').$$

Совокупность всех $\xi_v(\omega)$ из Σ' (соответственно из Σ''), можно рассматривать как координаты некоторой функции ξ' (соответственно ξ''). Определение II требует лишь независимости ξ' и ξ'' в смысле определения I при каждом выборе непересекающихся множеств Σ' и Σ'' .

Если $\xi_1, \xi_2, \dots, \xi_n$ независимы, то всегда

$$\begin{aligned} P(\xi_1 \in A_1, \xi_2 \in A_2, \dots, \xi_n \in A_n) = \\ = P(\xi_1 \in A_1) P(\xi_2 \in A_2) \dots P(\xi_n \in A_n), \end{aligned} \quad (5)$$

если только множества A_k принадлежат соответствующим \mathcal{F}_{ξ_k} (доказательство посредством *индукции*). Этого равенства, однако, в общем случае никоим образом недостаточно для независимости $\xi_1, \xi_2, \dots, \xi_n$.

Равенство (5) без затруднений обобщается на случай счетного произведения.

Из независимости случайных величин ξ_{v_k} в каждой конечной группе $(\xi_{v_1}, \xi_{v_2}, \dots, \xi_{v_k})$, вообще говоря, еще не следует, что все ξ_v независимы.

Наконец, легко заметить, что независимость функций ξ_v , собственно говоря, является свойством соответствующих разложений \mathfrak{M}_{ξ_v} . Если, далее, ξ'_v — однозначные функции соответствующих ξ_v , то из независимости ξ_v следует независимость ξ'_v .

§ 2. Независимые случайные величины

Если $\xi_1, \xi_2, \dots, \xi_n$ — независимые случайные величины, то из равенства (2) предыдущего параграфа следует, в частности, формула

$$F_{\xi_1, \xi_2, \dots, \xi_n}(x_1, x_2, \dots, x_n) = F_{\xi_1}(x_1) F_{\xi_2}(x_2) \dots F_{\xi_n}(x_n). \quad (1)$$

Т е о р е м а 1. Если алгебра множеств $\mathcal{F}_{\xi_1, \xi_2, \dots, \xi_n}$ состоит только из борелевских множеств пространства R^n , то условие (1) является также достаточным для независимости величин $\xi_1, \xi_2, \dots, \xi_n$.

Д о к а з а т е л ь с т в о. Пусть $\xi' = (\xi_{i_1}, \xi_{i_2}, \dots, \xi_{i_k})$ и $\xi'' = (\xi_{j_1}, \xi_{j_2}, \dots, \xi_{j_m})$ — две непересекающиеся подсистемы величин $\xi_1, \xi_2, \dots, \xi_n$. На основании формулы (1) следует доказать, что для любых двух борелевских множеств A' и A'' соответственно из R^k и R^m выполняется равенство

$$P(\xi' \in A', \xi'' \in A'') = P(\xi' \in A') P(\xi'' \in A''). \quad (2)$$

Для множеств вида

$$A' = \{(x_{i_1}, \dots, x_{i_k}) : x_{i_1} < a_1, \dots, x_{i_k} < a_k\},$$

$$A'' = \{(x_{j_1}, \dots, x_{j_m}) : x_{j_1} < b_1, \dots, x_{j_m} < b_m\}$$

это следует непосредственно из (1).

Далее доказывается, что это свойство сохраняется для сумм и разностей множеств указанного вида, откуда равенство (2) следует и для всех борелевских множеств.

Т е о р е м а II. Пусть $\xi = \{\xi_v\}$ — произвольная (вообще бесконечная) совокупность случайных величин. Если алгебра множеств \mathcal{F}_ξ совпадает с σ -алгеброй $\sigma(\mathcal{F}^{\mathcal{N}})$ (\mathcal{N} есть множество всех v)¹⁾, то совокупность равенств

$$F_{\xi_{v_1}, \dots, \xi_{v_n}}(x_1, \dots, x_n) = F_{\xi_{v_1}}(x_1) \dots F_{\xi_{v_n}}(x_n) \quad (3)$$

необходима и достаточна для независимости величин ξ_v , $v \in \mathcal{N}$.

Д о к а з а т е л ь с т в о. Необходимость этого условия следует непосредственно из формулы (1). Мы теперь докажем, что оно также и достаточное.

Пусть \mathcal{N}' и \mathcal{N}'' — два непересекающихся подмножества из множества \mathcal{N} всех индексов v , A' (соответственно A'') — множество из $\sigma(\mathcal{F}^{\mathcal{N}'})$, определенное соотношением между ξ_v с индексами v из \mathcal{N}' (соответственно из \mathcal{N}''). Следует доказать, что тогда выполнено равенство

$$P(A' \cap A'') = P(A') P(A''). \quad (4)$$

Действительно, если A' и A'' — цилиндрические множества, то мы имеем дело с соотношениями между конечным числом величин ξ_v , и равенство (4) представляет в этом случае простое следствие предшествующих результатов (формула (2)). А так как соотношение (4) сохраняется для сумм и разностей множеств A' (соответственно A''), то (4) доказано для всех множеств из $\sigma(\mathcal{F}^{\mathcal{N}'})$.

Пусть теперь для каждого v из некоторого множества \mathcal{N} задана а priori функция распределения $F_v = F_v(x)$.

Т е о р е м а III. Существует поле вероятностей (Ω, \mathcal{F}, P) и определенные на нем случайные величины ξ_v , $v \in \mathcal{N}$, которые независимы в совокупности, а каждая из величин ξ_v имеет в качестве функции распределения а priori данную функцию $F_v(x)$.

Д о к а з а т е л ь с т в о. Примем за основное множество Ω пространство $R^{\mathcal{N}}$ (множество всех систем $\omega = \{x_v\}$ действительных чисел x_v , $v \in \mathcal{N}$), а в каче-

¹⁾ См. § 4 гл. III.

стве \mathcal{F} возьмем σ -алгебру ¹⁾ $\sigma(\mathcal{F}^{\mathcal{N}^c})$. Положим далее $\xi_v(\omega) = x_v$ и определим функцию распределения F_{v_1, \dots, v_n} равенством

$$F_{v_1, \dots, v_n}(x_1, \dots, x_n) = F_{v_1}(x_1) \dots F_{v_n}(x_n).$$

Тогда, согласно основной теореме (§ 4, гл. III), на (Ω, \mathcal{F}) однозначно определена вероятностная функция P , причем $P\{\xi_v < x\} = F_v(x)$ для любого $v \in \mathcal{N}$.

Заметим еще, что, как было видно выше, из независимости всякой конечной группы величин ξ_v (равенство (3)) следует независимость всех ξ_v на $\sigma(\mathcal{F}^{\mathcal{N}^c})$. В объемлющих полях вероятностей это свойство может утеряться.

В заключение этого параграфа дадим еще некоторые признаки независимости для двух случайных величин.

Если две случайные величины ξ и η независимы и если $M\xi$ и $M\eta$ конечны, то почти наверное

$$\begin{aligned} M(\eta | \xi) &= M\eta, \\ M(\xi | \eta) &= M\xi. \end{aligned} \quad (5)$$

Эти формулы представляют непосредственное следствие второго определения условных математических ожиданий (формулы (10) и (11) § 4 пятой главы). Следовательно, в случае независимости ξ и η величины

$$f^2 = \frac{M[\eta - M(\eta | \xi)]^2}{D\eta}, \quad g^2 = \frac{M[\xi - M(\xi | \eta)]^2}{D\xi}$$

равны нулю (предполагается, что $D\xi > 0$ и $D\eta > 0$). Число f^2 называется *корреляционным отношением η по ξ* , g^2 — *корреляционным отношением ξ по η* (Пирсон).

Из (5) далее следует равенство

$$M \xi \eta = M \xi \cdot M \eta, \quad (6)$$

для доказательства которого надо воспользоваться формулой (15) из § 4 пятой главы:

$$\begin{aligned} M \xi \eta &= M[M(\xi \eta) | \xi] = M[\xi M(\eta | \xi)] = M[\xi \cdot M\eta] = \\ &= M\xi \cdot M\eta. \end{aligned}$$

¹⁾ См. § 4 гл. III.

Следовательно, в случае независимости ξ и η величина

$$\rho = \frac{M\xi\eta - M\xi \cdot M\eta}{\sqrt{D\xi \cdot D\eta}}$$

также равна нулю. Как известно, ρ — коэффициент корреляции между ξ и η .

Если две случайные величины ξ и η удовлетворяют равенству (6), то они называются *некоррелированными*. Для суммы

$$s = \xi_1 + \xi_2 + \dots + \xi_n$$

попарно некоррелированных величин $\xi_1, \xi_2, \dots, \xi_n$ легко сосчитать, что

$$Ds = D\xi_1 + D\xi_2 + \dots + D\xi_n. \quad (7)$$

В частности, равенство (7) справедливо для независимых величин $\xi_1, \xi_2, \dots, \xi_n$.

§ 3. Закон больших чисел

О п р е д е л е н и е. Случайные величины η_n последовательности

$$\eta_1, \eta_2, \dots, \eta_n, \dots$$

называются *устойчивыми*, если существует такая числовая последовательность

$$d_1, d_2, \dots, d_n, \dots,$$

что для любого положительного ε

$$P \{ |\eta_n - d_n| \geq \varepsilon \} \rightarrow 0, \quad n \rightarrow \infty.$$

Если существуют все $M\eta_n$ и если можно положить

$$d_n = M\eta_n,$$

то говорят, что устойчивость *нормальная*.

Если все η_n равномерно ограничены, то из

$$P \{ |\eta_n - d_n| \geq \varepsilon \} \rightarrow 0, \quad n \rightarrow \infty, \quad (1)$$

следует соотношение

$$|M \eta_n - d_n| \rightarrow 0, \quad n \rightarrow \infty,$$

и, следовательно,

$$P \{ |\eta_n - M\eta_n| \geq \varepsilon \} \rightarrow 0, \quad n \rightarrow \infty. \quad (2)$$

Таким образом, *устойчивость ограниченной последовательности необходимо нормальна*.

Пусть

$$\sigma_n^2 = D \eta_n (= M (\eta_n - M\eta_n)^2).$$

По неравенству Чебышева

$$P \{ |\eta_n - M\eta_n| \geq \varepsilon \} \leq \frac{\sigma_n^2}{\varepsilon^2}.$$

Следовательно, *условие Маркова*:

$$\sigma_n^2 \rightarrow 0, \quad n \rightarrow \infty, \quad (3)$$

достаточно для нормальной устойчивости.

Если $\eta_n - M\eta_n$ равномерно ограничены,

$$|\eta_n - M\eta_n| \leq C,$$

то по неравенству (9) из § 3 четвертой главы

$$P \{ |\eta_n - M\eta_n| \geq \varepsilon \} \geq \frac{\sigma_n^2 - \varepsilon^2}{C^2}.$$

Следовательно, в этом случае условие Маркова (3) является также и *необходимым* для нормальной устойчивости η_n .

Если

$$\eta_n = \frac{\xi_1 + \xi_2 + \dots + \xi_n}{n}$$

и величины ξ_n попарно некоррелированы, то

$$D\eta_n = \frac{1}{n^2} [D\xi_1 + D\xi_2 + \dots + D\xi_n].$$

Следовательно, в этом случае для нормальной устойчивости средних арифметических η_n , т. е. для того,

чтобы для всякого $\varepsilon > 0$

$$\lim_n P \left\{ \left| \frac{\xi_1 + \dots + \xi_n}{n} - \frac{M\xi_1 + \dots + M\xi_n}{n} \right| \geq \varepsilon \right\} = 0,$$

достаточно выполнения следующего условия:

$$\lim_n \frac{1}{n^2} \sum_{i=1}^n D\xi_i = 0 \quad (4)$$

(теорема Чебышева). В частности, условие (4) выполнено, если все величины ξ_n равномерно ограничены.

1. Можно обобщить эту теорему на случай *слабо коррелированных величин* ξ_n . Если предположить, что коэффициент корреляции ρ_{mn} ¹⁾ между ξ_m и ξ_n удовлетворяет неравенству

$$\rho_{mn} \leq c (|m - n|)$$

и что $c(k) \geq 0$, то для нормальной устойчивости средних арифметических, т. е. для того, чтобы для всякого $\varepsilon > 0$

$$\lim_n P \left\{ \left| \frac{\xi_1 + \dots + \xi_n}{n} - \frac{M\xi_1 + \dots + M\xi_n}{n} \right| \geq \varepsilon \right\} = 0,$$

достаточно выполнения условия²⁾

$$\lim_n \frac{C_n}{n^2} \cdot \sum_{i=1}^n D\xi_i = 0, \quad (5)$$

где $C_n = \sum_{k=0}^{n-1} c(k)$.

2. В случае *независимых* слагаемых ξ_n можно дать также необходимое и достаточное условие для устойчивости средних арифметических η_n .

¹⁾ Ясно, что всегда $\rho_{nn} = 1$.

²⁾ Ср. A. Khintchine, Sur la loi forte des grandes nombres, C. R. de l'Acad. Sci., Paris, т. 186 (1928), 285.

Для каждого ξ_n существует константа m_n (медиана ξ_n), удовлетворяющая следующим условиям:

$$P(\xi_n \leq m_n) \leq \frac{1}{2},$$

$$P(\xi_n \geq m_n) \leq \frac{1}{2}.$$

Положим

$$\xi_{nk} = \begin{cases} \xi_k, & \text{если } |\xi_k - m_k| \leq n, \\ 0, & \text{если } |\xi_k - m_k| > n, \end{cases}$$

$$\eta_n^* = \frac{\xi_{n1} + \dots + \xi_{nn}}{n}.$$

Т е о р е м а ¹⁾. Пусть ξ_1, ξ_2, \dots — последовательность взаимно независимых случайных величин. Тогда условия

$$\sum_{k=1}^n P(|\xi_k - m_k| > n) = \sum_{k=1}^n P(\xi_{nk} \neq \xi_k) \rightarrow 0, \quad n \rightarrow \infty, \quad (6)$$

$$\frac{1}{n^2} \sum_{k=1}^n D\xi_{nk} \rightarrow 0, \quad n \rightarrow \infty \quad (7)$$

необходимы и достаточны для устойчивости величин η_n , $n = 1, 2, \dots$.

При этом постоянные d_n , $n = 1, 2, \dots$, можно принять равными $M\eta_n^*$, так что в случае

$$M\eta_n^* - M\eta_n \rightarrow 0, \quad n \rightarrow \infty$$

(и только в этом случае) устойчивость нормальная.

¹⁾ Ср. A. K o l m o g o r o v, Über die Summen durch den Zufall bestimmter unabhängiger Grössen, Math. Ann., т. 99 (1928), стр. 309—319 (исправления и замечания к этой работе, т. 102 (1929), стр. 484—488), теорема VIII и добавление к ней на стр. 318.

Д о к а з а т е л ь с т в о. Достаточность условий (6), (7) устанавливается просто. В самом деле, поскольку

$$P(\eta_n \neq \eta_n^*) \leq \sum_{k=1}^n P(\xi_{nk} \neq \xi_k) \rightarrow 0, \quad n \rightarrow \infty,$$

а согласно неравенству Чебышева

$$P\{|\eta_n^* - M\eta_n^*| \geq \varepsilon\} \leq \frac{1}{n^2 \varepsilon^2} \sum_{k=1}^n D\xi_{nk} \rightarrow 0, \quad n \rightarrow \infty,$$

то

$$P\{|\eta_n - M\eta_n^*| \geq \varepsilon\} \rightarrow 0, \quad n \rightarrow \infty.$$

Для доказательства необходимости нам понадобится ряд вспомогательных предложений.

Л е м м а 1. Пусть A_1, A_2, \dots, A_n — независимые события, $P(A_i) > 0$, $i = 1, 2, \dots, n$ и для некоторого $u \geq 0$ $P(\bigcup_{i=1}^n A_i) \geq u$. Если, кроме того, событие U таково, что для каждого $i = 1, 2, \dots, n$

$$P(U | A_i) \geq u,$$

то тогда

$$P(U) \geq \frac{1}{9} u^2. \quad (8)$$

Д о к а з а т е л ь с т в о. Если существует такое u , что $P(A_i) \geq \frac{1}{3} u$, то

$$P(U) \geq P(U | A_i) P(A_i) \geq \frac{1}{3} u^2.$$

Пусть теперь для всех $i = 1, 2, \dots, n$

$$P(A_i) < \frac{1}{3} u.$$

Тогда найдется такое k , что

$$\frac{1}{3} u \leq P(A_1 \cup \dots \cup A_k) \leq \frac{2}{3} u,$$

и, значит, для всех $i \leq k$

$$\begin{aligned} P(A_2 \cup \dots \cup A_{i-1} | A_i) &= P(A_1 \cup \dots \cup A_{i-1}) \leq \\ &\leq P(A_1 \cup \dots \cup A_k) \leq \frac{2}{3} u, \end{aligned}$$

$$P(U \cap \overline{(A_1 \cup \dots \cup A_{i-1})} | A_i) \geq \frac{1}{3} u,$$

$$P(U \cap \overline{(A_1 \cup \dots \cup A_{i-1})} \cap A_i) \geq \frac{1}{3} u P(A_i).$$

Отсюда

$$\begin{aligned} P(U) &\geq \sum_{i=1}^k P(U \cap \overline{(A_1 \cup \dots \cup A_{i-1})} \cap A_i) \geq \\ &\geq \frac{1}{3} u \sum_{i=1}^k P(A_i) \geq \frac{1}{3} u P(A_1 \cup \dots \cup A_k) \geq \frac{1}{9} u^2. \end{aligned}$$

Л е м м а 2. Пусть $\xi_1, \xi_2, \dots, \xi_n$ — независимые, ограниченные, $|\xi_i| \leq c$, $i = 1, \dots, n$, случайные величины с нулевыми средними. Тогда для всякого $\alpha > 0$ и целого m

$$P\{\max_{k \leq n} |\xi_1 + \dots + \xi_k| \geq m(\alpha D + c)\} \leq \frac{1}{\alpha^{2m}}, \quad (9)$$

где

$$D^2 = \sum_{i=1}^n D\xi_i.$$

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

$$R = \alpha D + c, \quad s_0 = 0, \quad s_k = \xi_1 + \dots + \xi_k,$$

$$B_{ik} = \{\omega : |s_j| < iR, \quad j < k; \quad |s_k| \geq iR\},$$

$$B_i = \sum_{k=1}^n B_{ik}.$$

Замечая, что на множестве B_{ik}

$$|s_k| \leq iR + c,$$

получаем

$$\begin{aligned} P(B_{i+1} | B_{ik}) &= P\left\{ \max_{1 \leq k \leq n} |s_k| \geq (i+1)R | B_{ik} \right\} \leq \\ &\leq P\left\{ \max_{k+1 \leq p \leq n} \left| \sum_{j=k+1}^p \xi_j \right| \geq \alpha D | B_{ik} \right\} = \\ &= P\left\{ \max_{k+1 \leq p \leq n} \left| \sum_{j=k+1}^p \xi_j \right| \geq \alpha D \right\}. \end{aligned}$$

Из доказываемого далее неравенства (3) теоремы II § 5 следует, что

$$P\left\{ \max_{k+1 \leq p \leq n} \left| \sum_{j=k+1}^p \xi_j \right| \geq \alpha D \right\} \leq \frac{\sum_{j=k+1}^p D\xi_j}{\alpha^2 D^2} \leq \frac{D^2}{\alpha^2 D^2} = \frac{1}{\alpha^2}.$$

Поэтому $P(B_{i+1} | B_{ik}) \leq \frac{1}{\alpha^2}$ при любом $k = 1, 2, \dots, n$.

Значит,

$$P(B_{i+1} | B_i) \leq \frac{1}{\alpha^2}$$

и

$$\begin{aligned} P\left\{ \max_{1 \leq k \leq n} |s_k| \geq mR \right\} &= P(B_m) = \\ &= P(B_m | B_{m-1}) P(B_{m-1} | B_{m-2}) \dots P(B_1 | B_0) \leq \frac{1}{\alpha^{2m}}. \end{aligned}$$

Л е м м а 3. Пусть $\xi_1, \xi_2, \dots, \xi_n$ — независимые ограниченные случайные величины, причем $|\xi_i - M \xi_i| \leq C, i = 1, 2, \dots, n$. Тогда

$$P\{| \xi_1 + \dots + \xi_n | \geq a\} \geq \frac{1}{1600} \left[1 - \frac{4a^2 + C^2}{\sum_{i=1}^n D\xi_i} \right]. \quad (10)$$

Д о к а з а т е л ь с т в о. Обозначим $s = \xi_1 + \dots + \xi_n, D^2 = \sum_{i=1}^n D\xi_i$. Если $C > D$ или $2a > D$, то правая часть в (10) отрицательна и неравенство очевидно.

Пусть теперь одновременно $C \leq D$, $2a \leq D$. Тогда достаточно показать, что

$$P\left(|s| \geq \frac{D}{2}\right) \geq \frac{1}{1600},$$

поскольку, очевидно,

$$P(|s| \geq a) \geq P\left(|s| \geq \frac{D}{2}\right) \geq \frac{1}{1600} \geq \frac{1}{1600} \left[1 - \frac{4a^2 + C^2}{D^2}\right].$$

$$\begin{aligned} \text{Обозначим } A = \{|s| \geq \frac{D}{2}\}. \text{ Если } |Ms| \geq 2D, \text{ то} \\ D^2 \geq P(\bar{A}) M\{(s - Ms)^2 | \bar{A}\} \geq \left(2D - \frac{D}{2}\right)^2 P(\bar{A}) = \\ = \frac{9}{4} D^2 P(\bar{A}) \end{aligned}$$

и, значит,

$$P(A) \geq \frac{1}{2}.$$

Предположим, теперь, что $|Ms| < 2D$. Обозначая $A_m = \left\{\omega : 3mD \leq |s - Ms| < 3(m+1)D, |s| \geq \frac{D}{2}\right\}$

и применяя лемму 2, находим

$$\begin{aligned} P(A_m) &\leq P\{|s - Ms| \geq 3mD\} = P\{|s - Ms| \geq \\ &\geq m[2D + D]\} \leq P\{|s - Ms| \geq m[2D + C]\} \leq \frac{1}{2^{2m}}. \end{aligned}$$

Отсюда

$$\begin{aligned} M(s^2 | A_m) P(A_m) &= M[(s - Ms)^2 + 2(s - Ms)Ms + \\ &+ (Ms)^2 | A_m] P(A_m) \leq 2s \frac{(m+1)^2}{2^{2m}} \cdot D^2. \end{aligned}$$

На множестве $A' = \sum_{m=0}^5 A_m$

$$|s| \leq |s - Ms| + |Ms| \leq 20 \cdot D.$$

Поэтому

$$M(s^2 | A') P(A') \leq 400 D^2 P(A).$$

Ясно также, что

$$M(s^2 | \bar{A}) P(\bar{A}) \leq \frac{1}{4} D^2.$$

Следовательно,

$$\begin{aligned} D^2 &\leq M s^2 = M[s^2 | \bar{A}] P(\bar{A}) + M[s^2 | A'] P(A') + \\ &\quad + \sum_{m=6}^{\infty} M[s^2 | A_m] P(A_m) \leq \\ &\leq \frac{1}{4} D^2 + 400 D^2 P(A) + \sum_{m=6}^{\infty} \frac{(m+1)^2}{2^{2m}} \cdot 25 \cdot D^2, \end{aligned}$$

и, значит,

$$P(A) \geq \frac{1}{1600}.$$

Доказательство теоремы. Необходимость. Пусть последовательность d_n , $n = 1, 2, \dots$, такова, что для любого $\varepsilon > 0$ $P(|\eta_n - d_n| \geq \varepsilon) \rightarrow 0, n \rightarrow \infty$. Покажем, что тогда

$$\sum_{k=1}^n P\{|\xi_k - m_k| \geq n\varepsilon\} \rightarrow 0, \quad n \rightarrow \infty. \quad (11)$$

Обозначим для данного n

$$U = \left\{ |\eta_n - d_n| \geq \frac{\varepsilon}{2} \right\};$$

$$A_k = \{|\xi_k - m_k| \geq n\varepsilon\},$$

$$\begin{aligned} B_k &= \left\{ \left| \frac{\sum_{i \neq k} \xi_i + m_k}{n} - d_n \right| \geq \frac{\varepsilon}{2} \right\} = \\ &= \left\{ \left| (\eta_n - d_n) + \frac{m_k - \xi_k}{n} \right| \geq \frac{\varepsilon}{2} \right\}. \end{aligned}$$

Поскольку m_k — медиана ξ_k , то

$$P(U | B_k) \geq \frac{1}{2}.$$

Для достаточно больших n

$$P(U) \leq \frac{1}{4},$$

поэтому

$$\frac{1}{4} \geq P(U) \geq P(U | B_k) P(B_k) \geq \frac{1}{2} P(B_k),$$

т. е.

$$P(B_k) \leq \frac{1}{2}.$$

Далее, если событие A_k выполняется, а B_k нет, то выполняется событие U и, значит,

$$P(B_k | A_k) + P(U | A_k) \geq 1.$$

Но

$$P(B_k | A_k) = P(B_k) \leq \frac{1}{2}.$$

Следовательно,

$$P(U | A_k) \geq 1 - P(B_k | A_k) \geq \frac{1}{2} \geq \frac{1}{2} P\left(\bigcup_{k=1}^n A_k\right).$$

Применим лемму 1, взяв

$$u = \frac{1}{2} P\left(\bigcup_{k=1}^n A_k\right).$$

Тогда

$$P(U) \geq \frac{1}{36} [P\left(\bigcup_{k=1}^n A_k\right)]^2. \quad (12)$$

События A_1, \dots, A_n независимы, поэтому

$$P\left(\bigcup_{k=1}^n A_k\right) = 1 - \prod_{k=1}^n (1 - P(A_k)). \quad (13)$$

Поскольку по условию $P(U) \rightarrow 0$, $n \rightarrow \infty$, то из (12) и (13) получаем искомое соотношение (11).

Положим теперь

$$\bar{\xi}_{nk} = \begin{cases} \xi_k, & \text{если } |\xi_k - m_k| \leq n, \\ m_k, & \text{если } |\xi_k - m_k| > n, \end{cases}$$

и

$$\bar{\xi}_{nk}(\varepsilon) = \begin{cases} \xi_k, & \text{если } |\xi_k - m_k| \leq \varepsilon n, \\ m_k, & \text{если } |\xi_k - m_k| > \varepsilon n. \end{cases}$$

Из (11) следует, что если $P\{|\eta_n - d_n| \geq \varepsilon\} \rightarrow 0$, $n \rightarrow \infty$, то и

$$P\left\{\left|\frac{1}{n} \sum_{k=1}^n \bar{\xi}_{nk} - d_n\right| \geq \varepsilon\right\} \rightarrow 0,$$

$$P\left\{\left|\frac{1}{n} \sum_{k=1}^n \bar{\xi}_{nk}(\varepsilon) - d_n\right| \geq \varepsilon\right\} \rightarrow 0.$$

Обозначим $\bar{\xi}_{nk}(\varepsilon) = \bar{\xi}_{nk}(\varepsilon) - M\bar{\xi}_{nk}(\varepsilon)$. Тогда $|\bar{\xi}_{nk}(\varepsilon)| \leq 2n\varepsilon$ и по лемме 3

$$\begin{aligned} P\left\{\left|\frac{1}{n} \sum_{k=1}^n \bar{\xi}_{nk}(\varepsilon) - d_n\right| \geq \varepsilon\right\} &= \\ &= P\left\{\left|\sum_{k=1}^n \bar{\xi}_{nk}(\varepsilon) - nd_n\right| \geq \varepsilon \cdot n\right\} = \\ &= P\left\{\left|\sum_{k=1}^n (\bar{\xi}_{nk}(\varepsilon) - d_n)\right| \geq \varepsilon n\right\} \geq \frac{1}{1600} \left[1 - \frac{8\varepsilon^2 n^3}{\sum_{k=1}^n D\bar{\xi}_{nk}(\varepsilon)}\right], \end{aligned}$$

откуда

$$\limsup_n \frac{1}{n^2} \sum_{k=1}^n D\bar{\xi}_{nk}(\varepsilon) \leq 8\varepsilon^2. \quad (14)$$

Для $\varepsilon \leq 1$

$$|D\bar{\xi}_{nk}(\varepsilon) - D\bar{\xi}_{nk}| \leq 8n^2 P\{|\xi_k - m_k| > n\varepsilon\}.$$

Тогда из (11), (13) и (14) следует, что

$$\limsup_n \frac{1}{n^2} \sum_{k=1}^n D\bar{\xi}_{nk} \leq 8\varepsilon^2,$$

а значит, в силу произвольности $\varepsilon \in (0, 1]$

$$\lim_n \frac{1}{n^2} \sum_{k=1}^n D\xi_{nk} = \lim_n \frac{1}{n^2} \sum_{k=1}^n D\tilde{\xi}_{nk} = 0.$$

3. Дальнейшее обобщение теоремы Чебышева получается, если предположить, что η_n каким-нибудь образом зависят от исходов каких-либо n испытаний

$$\mathfrak{A}_1, \mathfrak{A}_2, \dots, \mathfrak{A}_n,$$

так что после каждого определенного исхода всех этих n испытаний η_n принимает определенное значение. Общая идея всех теорем, известных под названием *закона больших чисел*, состоит в том, что если зависимость величины η_n от каждого отдельного испытания \mathfrak{A}_k , $k = 1, 2, \dots, n$, очень мала при больших n , то величины η_n устойчивы. Если рассматривать

$$\beta_{nk}^2 = M[M(\eta_n | \mathfrak{A}_1, \mathfrak{A}_2, \dots, \mathfrak{A}_k) - M(\eta_n | \mathfrak{A}_1, \mathfrak{A}_2, \dots, \mathfrak{A}_{k-1})]^2$$

как разумную меру зависимости величины η_n от испытания \mathfrak{A}_k , то вышеупомянутая общая идея закона больших чисел может быть конкретизирована следующими рассуждениями ¹⁾.

Пусть

$$\zeta_{nk} = M[\eta_n | \mathfrak{A}_1, \dots, \mathfrak{A}_k] - M[\eta_n | \mathfrak{A}_1, \dots, \mathfrak{A}_{k-1}].$$

Тогда

$$\eta_n - M\eta_n = \zeta_{n1} + \zeta_{n2} + \dots + \zeta_{nn},$$

$$\begin{aligned} M\zeta_{nk} &= MM[\eta_n | \mathfrak{A}_1, \dots, \mathfrak{A}_k] - MM[\eta_n | \mathfrak{A}_1, \dots, \mathfrak{A}_{k-1}] = \\ &= M\eta_n - M\eta_n = 0, \end{aligned}$$

$$D\zeta_{nk} = M\zeta_{nk}^2 = \beta_{nk}^2.$$

Легко, далее, подсчитать, что случайные величины ζ_{nk} , $k = 1, 2, \dots, n$, некоррелированы. В самом деле,

¹⁾ Ср. А. К о л м о г о р о в, Sur la loi des grandes nombres, Rend. Acad. Lincei, т. 9 (1929), 470—474.

пусть $i < k$, тогда ¹⁾

$$\begin{aligned} M[\xi_{ni}\xi_{nk} | \mathfrak{A}_1, \dots, \mathfrak{A}_{k-1}] &= \xi_{ni} M[\xi_{nk} | \mathfrak{A}_1, \dots, \mathfrak{A}_{k-1}] = \\ &= \xi_{ni} M\{M(\eta_n | \mathfrak{A}_1, \dots, \mathfrak{A}_k) - \\ &- M(\eta_n | \mathfrak{A}_1, \dots, \mathfrak{A}_{k-1}) | \mathfrak{A}_1, \dots, \mathfrak{A}_{k-1}\} = \\ &= \xi_{ni} [M(\eta_n | \mathfrak{A}_1, \dots, \mathfrak{A}_{k-1}) - M(\eta_n | \mathfrak{A}_1, \dots, \mathfrak{A}_{k-1})] = 0 \end{aligned}$$

и, следовательно,

$$M\xi_{ni}\xi_{nk} = 0, \quad i < k.$$

Итак,

$$D\eta_n = \sum_{i=1}^n D\xi_{ni} = \sum_{i=1}^n \beta_{ni}^2.$$

Таким образом, условие

$$\sum_{i=1}^n \beta_{ni}^2 \rightarrow 0, \quad n \rightarrow \infty$$

достаточно для нормальной устойчивости величины η_n .

§ 4. Замечания к понятию математического ожидания

Мы определили математическое ожидание величины ξ как

$$M\xi = \int_{\Omega} \xi(\omega) P(d\omega) = \int_{-\infty}^{\infty} xF_{\xi}(dx).$$

При этом интеграл в правой части понимается в смысле

$$M\xi = \int_{-\infty}^{\infty} xF_{\xi}(dx) = \lim_{\substack{a \rightarrow -\infty \\ b \rightarrow +\infty}} \int_a^b xF_{\xi}(dx). \quad (1)$$

Напрашивается мысль о том, чтобы рассматривать выражение

$$\widetilde{M}\xi = \lim_{a \rightarrow +\infty} \int_{-a}^a xF_{\xi}(dx) \quad (2)$$

¹⁾ Применение формулы (15) § 4 гл. V.

в качестве обобщенного математического ожидания. Однако при этом теряются некоторые простые свойства математических ожиданий. Например, формула

$$\tilde{M}[\xi + \eta] = \tilde{M}\xi + \tilde{M}\eta$$

не всегда верна. Далее мы увидим, что при некоторых ограничивающих предположениях определение (2) является вполне естественным и пригодным.

Именно, можно поставить вопрос следующим образом. Пусть

$$\xi_1, \xi_2, \dots, \xi_n, \dots$$

— последовательность независимых величин, имеющих ту же функцию распределения $F_{\xi}(x) = F_{\xi_n}(x)$, $n = 1, 2, \dots$, что и ξ . Пусть далее

$$\eta_n = \frac{\xi_1 + \xi_2 + \dots + \xi_n}{n}.$$

Спрашивается, существует ли постоянная $M^\circ \xi$ такая, что при каждом $\varepsilon > 0$

$$\lim_n P \{ |\eta_n - M^\circ \xi| > \varepsilon \} = 0. \quad (3)$$

Ответ гласит: *если такая постоянная $M^\circ \xi$ существует, то $M^\circ \xi = \tilde{M}\xi$. Необходимое и достаточное условие для существования $M^\circ \xi$ заключается в существовании предела (2) и выполнении соотношения*

$$P(|\xi| > n) = o\left(\frac{1}{n}\right). \quad (4)$$

Этот результат непосредственно вытекает из следующей теоремы.

Т е о р е м а. Пусть ξ_1, ξ_2, \dots — последовательность независимых одинаково распределенных случайных величин с функцией распределения $F(x)$. Для устойчивости средних арифметических

$$\eta_n = \frac{\xi_1 + \xi_2 + \dots + \xi_n}{n}, \quad n = 1, 2, \dots$$

необходимо и достаточно выполнение условия (4). При этом в случае устойчивости можно положить

$$d_n = \int_{-n}^n xF(dx).$$

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

$$\xi_{nk} = \begin{cases} \xi_k, & \text{если } |\xi_k - m| \leq n, \\ 0, & \text{если } |\xi_k - m| > n, \end{cases}$$

где m — медиана случайной величины ξ_1 . Из условия (4) следует, что

$$\sum_{k=1}^n \mathbf{P}(\xi_{nk} \neq \xi_k) = n\mathbf{P}(|\xi_1 - m| > n) \rightarrow 0, \quad n \rightarrow \infty$$

и

$$\overline{\lim}_n \frac{1}{n} \int_{\{x: |x-m| \leq n\}} x^2 F(dx) = \overline{\lim}_n \frac{1}{n} \int_{|x| \leq n} x^2 F(dx). \quad (5)$$

Но

$$\begin{aligned} \frac{1}{n} \int_{|x| \leq n} x^2 F(dx) &= \frac{1}{n} \sum_{k=1}^n \int_{k-1 < |x| \leq k} x^2 F(dx) \leq \\ &\leq \frac{1}{n} \sum_{k=1}^n k^2 \mathbf{P}(k-1 < |\xi_1| \leq k) \leq \\ &\leq \frac{2}{n} \sum_{k=1}^n \left[\sum_{l=1}^k l \mathbf{P}(k-1 < |\xi_1| \leq k) \right] \leq \\ &\leq \frac{2}{n} \sum_{l=1}^n l \mathbf{P}(l-1 < |\xi_1| \leq l) \leq \frac{2}{n} \sum_{l=1}^n l \mathbf{P}(|\xi_1| > l-1) \rightarrow 0, \end{aligned}$$

поскольку

$$l \mathbf{P}\{|\xi_1| > l-1\} \rightarrow 0, \quad l \rightarrow \infty.$$

Поэтому

$$\begin{aligned} \frac{1}{n^2} \sum_{k=1}^n D\xi_{nk} &\leq \frac{1}{n^2} \sum_{k=1}^n M\xi_{nk}^2 = \\ &= \frac{1}{n^2} \int_{|x-m| \leq n} x^2 F(dx) \rightarrow 0, \quad n \rightarrow \infty. \end{aligned}$$

Таким образом, условия (6), (7) теоремы из § 3 выполнены и для всякого $\varepsilon > 0$

$$P(|\eta_n - M\eta_n^*| \geq \varepsilon) \rightarrow 0, \quad n \rightarrow \infty,$$

где

$$M\eta_n^* = M \frac{\xi_{n1} + \dots + \xi_{nn}}{n} = \int_{|x-m| \leq n} xF(dx).$$

Но в силу условия (4)

$$\int_{|x-m| \leq n} xF(dx) - \int_{|x| \leq n} xF(dx) \rightarrow 0, \quad n \rightarrow \infty,$$

поэтому

$$P(|\eta_n - d_n| \geq \varepsilon) \rightarrow 0, \quad n \rightarrow \infty,$$

где

$$d_n = \int_{-n}^n xF(dx).$$

Обратно, если последовательность η_n , $n = 1, 2, \dots$ устойчива, то в силу соотношения (11) из § 3

$$nP\{|\xi_1 - m| \geq n\} \rightarrow 0, \quad n \rightarrow \infty,$$

откуда следует условие (4).

Если существует математическое ожидание в прежнем смысле (формула (1)), то условие (4) всегда выполнено ¹⁾ и $M\xi = M^\circ\xi$.

Таким образом, $M^\circ(\xi)$ является законным обобщением математического ожидания $M(\xi)$. Для него сохраняются свойства I—VII (глава четвертая, § 2), однако в общем случае из существования $M^\circ(\xi)$ не следует существование $M^\circ|\xi|$.

Для того чтобы доказать, что новое понятие математического ожидания является действительно более общим, чем предыдущее, достаточно следующего примера.

¹⁾ Ср. A. K o l m o g o r o f f, Bemerkungen zu meiner Arbeit, Über die Summen zufälliger Grössen, Math. Ann., т. 102 (1930), стр. 484—488, теорема XII.

Возьмем плотность вероятности $f_{\xi}(x)$ равной

$$f_{\xi}(x) = \frac{C}{(|x| + 2)^2 \ln(|x| + 2)},$$

причем постоянная C определяется из условия

$$\int_{-\infty}^{\infty} f_{\xi}(x) dx = 1.$$

Легко далее подсчитать, что в этом случае выполнено условие (4). Формула (2) дает значение

$$M^{\circ} \xi = 0,$$

однако интеграл

$$\int_{-\infty}^{\infty} |x| F_{\xi}(dx) = \int_{-\infty}^{\infty} |x| f_{\xi}(x) dx$$

расходится.

Если Ω есть отрезок $[0,1]$ и P — лебеговская мера, то $M^{\circ} \xi$ есть не что иное, как A -интеграл ¹⁾

$$(A) \int_0^1 \xi(\omega) P(d\omega).$$

§ 5. Усиленный закон больших чисел, сходимость рядов

О п р е д е л е н и е. Случайные величины η_n последовательности

$$\eta_1, \eta_2, \dots, \eta_n, \dots$$

сильно устойчивы, если существует такая числовая последовательность

$$d_1, d_2, \dots, d_n, \dots,$$

¹⁾ См., например, Н. К. Б а р и, Тригонометрические ряды, Физматгиз, 1961.

что случайные величины

$$\eta_n - d_n$$

почти наверное стремятся к нулю при $n \rightarrow \infty$:

$$P \{ \lim_n (\eta_n - d_n) = 0 \} = 1.$$

Из сильной устойчивости следует, очевидно, обычная устойчивость. Если можно выбрать

$$d_n = M\eta_n,$$

то говорят, что сильная устойчивость нормальна.

Т е о р е м а I. Пусть ξ_1, ξ_2, \dots — последовательность независимых случайных величин. Для сильной нормальной устойчивости средних арифметических

$$\eta_n = \frac{\xi_1 + \dots + \xi_n}{n}, \quad n = 1, 2, \dots, \quad (1)$$

достаточно выполнения условия ¹⁾

$$\sum_{n=1}^{\infty} \frac{D\xi_n}{n^2} < \infty. \quad (2)$$

Это условие является наилучшим в том смысле, что для любого ряда неотрицательных постоянных $b_1,$

b_2, \dots таких, что $\sum_{n=1}^{\infty} \frac{b_n}{n^2} = \infty$, можно построить та-

кой ряд из независимых случайных величин ξ_1, ξ_2, \dots , что $D\xi_n = b_n$, а соответствующие средние арифметические $\eta_n, n = 1, 2, \dots$, не будут сильно устойчивы.

Для доказательства нам понадобится неравенство (3) из следующей теоремы.

Т е о р е м а II. Пусть $\xi_1, \xi_2, \dots, \xi_n$ — независимые случайные величины с нулевыми средними. Тогда

¹⁾ А. К о л м о г о р о ф ф, Sur la loi forte des grandes nombres, C. R. Acad. Sci., Paris, v. 191 (1930), 910—911.

для любого $a > 0$

$$P\{\max_{1 \leq k \leq n} |\xi_1 + \dots + \xi_k| \geq a\} \leq \frac{\sum_{k=1}^n D\xi_k}{a^2}. \quad (3)$$

Если к тому же случайные величины ξ_1, \dots, ξ_n ограничены, $|\xi_i| \leq c, i = 1, \dots, n$, то справедлива также оценка снизу

$$P\{\max_{1 \leq k \leq n} |\xi_1 + \dots + \xi_k| \geq a\} \geq 1 - \frac{(c+a)^2}{n \sum_{k=1}^n D\xi_k}. \quad (4)$$

Доказательство теоремы II. Положим $s_0 = 0, s_k = \xi_1 + \dots + \xi_k$,

$$A = \{\max_{1 \leq k \leq n} |s_k| \geq a\}, \quad A_k = \{\max_{j \leq k \leq 1} |s_j| < a, |s_k| \geq a\},$$

где $k = 1, 2, \dots, n$. Тогда $\sum_{k=1}^n A_k = A$, и, обозначая I_B — индикатор множества B , находим

$$\begin{aligned} \sum_{k=1}^n D\xi_k &= M s_n^2 \geq M[s_n^2 I_A] = \sum_{k=1}^n M[s_n^2 I_{A_k}] = \\ &= \sum_{k=1}^n M\{[s_k + (s_n - s_k)]^2 I_{A_k}\} = \sum_{k=1}^n M I_{A_k} s_k^2 + \\ &+ \sum_{k=1}^n M I_{A_k} (s_n - s_k)^2 + 2 \sum_{k=1}^n M[I_{A_k} s_k (s_n - s_k)] \geq \\ &\geq a^2 \sum P(A_k) = a^2 P(A), \end{aligned} \quad (5)$$

где мы использовали то, что

$$M[I_{A_k} s_k (s_n - s_k)] = M[I_{A_k} s_k M(s_n - s_k | \xi_1, \dots, \xi_k)] = 0.$$

Из (5) следует требуемое неравенство (3).

Чтобы установить (4), заметим, что, с одной стороны,

$$\begin{aligned} M[s_n^2 I_A] &= Ms_n^2 - Ms_n^2 I_{\bar{A}} \geqslant \\ &\geqslant \sum_{k=1}^n D\xi_k - a^2 P(\bar{A}) = \sum_{k=1}^n D\xi_k - a^2 + a^2 P(A). \end{aligned} \quad (6)$$

С другой стороны, воспользуемся тем, что на множествах $A_k \mid s_{k-1} \mid \leqslant a$, и, значит, $|s_k| \leqslant a + c$. Тогда

$$\begin{aligned} M[s_n^2 I_A] &= \sum_{k=1}^n M[s_k^2 I_{A_k}] + \sum_{k=1}^n M[I_{A_k} (s_n - s_k)^2] \leqslant \\ &\leqslant (a + c)^2 \sum_{k=1}^n P(A_k) + \sum_{k=1}^n P(A_k) \sum_{j=k+1}^n D\xi_j \leqslant \\ &\leqslant \left[(a + c)^2 + \sum_{j=1}^n D\xi_j \right] P(A). \end{aligned} \quad (7)$$

Из (6) и (7) получаем искомое неравенство

$$\begin{aligned} P(A) &\geqslant \frac{\sum_{k=1}^n D\xi_k - a^2}{(a + c)^2 + \sum_{k=1}^n D\xi_k - a^2} = \\ &= 1 - \frac{(a + c)^2}{(a + c)^2 + \sum_{k=1}^n D\xi_k - a^2} \geqslant 1 - \frac{(a + c)^2}{\sum_{k=1}^n D\xi_k}. \end{aligned}$$

Доказательство теоремы I. Не ограничивая общности, можно считать $M\xi_n = 0$, $n = 1, 2, \dots$. Тогда для доказательства того, что средние арифметические $\eta_n = \frac{s_n}{n}$ сходятся с вероятностью единица к нулю, достаточно доказать, что для любого $\varepsilon > 0$ вероятность

$$P \left\{ \limsup_n \left| \frac{s_n}{n} \right| \geqslant \varepsilon \right\} = 0.$$

В силу неравенства (3)

$$P_m \equiv P \left\{ \max_{2^m \leq n < 2^{m+1}} \left| \frac{s_n}{n} \right| \geq \varepsilon \right\} \leq \\ \leq P \left\{ \max_{1 \leq n < 2^{m+1}} |s_n| \geq 2^m \varepsilon \right\} \leq \left(\frac{1}{2^m \varepsilon} \right)^2 \sum_{n=1}^{2^{m+1}} D\xi_n.$$

Поэтому

$$P \equiv P \left\{ \limsup \left| \frac{s_n}{n} \right| \geq \varepsilon \right\} \leq \sum_{m=0}^{\infty} P_m \leq \\ \leq \frac{1}{\varepsilon^2} \sum_{m=0}^{\infty} \left(\frac{1}{2^m} \right)^2 \sum_{n=1}^{2^{m+1}} D\xi_n \leq \frac{1}{\varepsilon^2} \sum_{i=0}^{\infty} \sum_{m=i}^{\infty} \left(\frac{1}{2^m} \right)^2 \sum_{n=2^i}^{2^{i+1}} D\xi_n \leq \\ \leq \frac{1}{\varepsilon^2} \sum_{i=0}^{\infty} \left(\frac{1}{2^{i-1}} \right)^2 \sum_{n=2^i}^{2^{i+1}} D\xi_n \leq \frac{8}{\varepsilon^2} \sum_{n=1}^{\infty} \frac{D\xi_n}{n^2}.$$

Заметим теперь, что вероятность P не изменится, если заменить нулем любое конечное число первых членов последовательности ξ_1, ξ_2, \dots . Значит, для любого N

$$P \leq \frac{8}{\varepsilon^2} \sum_{n=N}^{\infty} \frac{D\xi_n}{n^2},$$

что в силу сходимости ряда $\sum_{n=1}^{\infty} \frac{D\xi_n}{n^2}$ доказывает, что $P = 0$.

Для доказательства заключительного утверждения теоремы построим последовательность независимых случайных величин ξ_1, ξ_2, \dots следующим образом.

Если $\frac{b_n}{n^2} \leq 1$, то возьмем

$$\xi_n = n, \quad \bar{\xi}_n = -n, \quad \xi_n = 0$$

с вероятностями

$$\frac{b_n}{2n^2}, \quad \frac{b_n}{2n^2}, \quad 1 - \frac{b_n}{n^2}$$

соответственно.

Если же $\frac{b_n}{n^2} \geq 1$, то положим

$$\xi_n = \sqrt{b_n}, \quad \xi_n = -\sqrt{b_n}$$

с вероятностями, равными $\frac{1}{2}$.

Тогда $M\xi_n = 0$, $M\xi_n^2 = b_n$. Если теперь предположить, что на множестве положительной вероятности $\frac{s_n}{n} \rightarrow 0$, $n \rightarrow \infty$, то тогда из равенства

$$\frac{\xi_n}{n} = \frac{s_n}{n} - \left(\frac{n-1}{n} \right) \frac{s_{n-1}}{n-1}$$

будет следовать, что с положительной вероятностью

$\frac{\xi_n}{n} \rightarrow 0$, $n \rightarrow \infty$. Но в силу расходимости ряда $\sum_{n=1}^{\infty} \frac{b_n}{n^3}$

ряд

$$\sum_{n=1}^{\infty} P \left\{ \left| \frac{\xi_n}{n} \right| \geq 1 \right\} = \infty.$$

Отсюда в силу леммы Бореля — Каптелли¹⁾ следует, что $\frac{\xi_n}{n}$ не может с положительной вероятностью сходиться к нулю.

В случае, когда независимые случайные величины ξ_1, ξ_2, \dots имеют одну и ту же функцию распределения $F = F(x)$, можно дать необходимое и достаточное условие для сильной устойчивости средних арифметических.

¹⁾ Здесь и далее используется следующий вариант леммы Бореля — Каптелли: если A_1, A_2, \dots — последовательность событий таких, что их индикаторы I_{A_1}, I_{A_2}, \dots независимы, то

$\sum_{n=1}^{\infty} I_{A_n} < \infty$ с вероятностью единица тогда и только тогда,

когда $\sum_{n=1}^{\infty} P(A_n) < \infty$.

Теорема III. Если ξ_1, ξ_2, \dots — последовательность независимых одинаково распределенных случайных величин, то условие $M |\xi_1| < \infty$ является необходимым и достаточным для сильной нормальной устойчивости средних арифметических $\eta_n = \frac{s_n}{n}$, $n = 1, 2, \dots$

Доказательство. Пусть $M |\xi_1| < \infty$; покажем, что тогда с вероятностью единица

$$\frac{s_n}{n} \rightarrow M \xi_1, \quad n \rightarrow \infty. \quad (8)$$

Наряду с величинами ξ_n рассмотрим «урезанные» величины

$$\xi_n^* = \begin{cases} \xi_n, & \text{если } |\xi_n| \leq n, \\ 0, & \text{если } |\xi_n| > n. \end{cases}$$

Тогда

$$\begin{aligned} \sum_{n=1}^{\infty} P \{ \xi_n \neq \xi_n^* \} &= \sum_{n=1}^{\infty} P \{ |\xi_n| > n \} = \sum_{n=1}^{\infty} P \{ |\xi_1| > n \} = \\ &= \sum_{n=1}^{\infty} \sum_{k \geq n} P \{ k < |\xi_1| \leq k+1 \} = \\ &= \sum_{n=1}^{\infty} n P \{ n < |\xi_1| \leq n+1 \} \leq \\ &\leq \sum_{k=0}^{\infty} \int_{\{k < |x| \leq k+1\}} |x| F(dx) \leq \int_{-\infty}^{\infty} |x| F(dx) = M |\xi_1| < \infty. \end{aligned}$$

Поэтому по лемме Бореля — Кантелли с вероятностью единица произойдет лишь конечное число событий $A_n = \{ \xi_n \neq \xi_n^* \}$, $n = 1, 2, \dots$. Отсюда вытекает, что достаточно доказать, что с вероятностью единица

$$\frac{\xi_1^* + \dots + \xi_n^*}{n} \rightarrow M \xi_1.$$

Поскольку

$$M\xi_n^* \rightarrow M\xi_1,$$

то отсюда нетрудно заключить, что

$$\frac{M\xi_1^* + \dots + M\xi_n^*}{n} \rightarrow M\xi_1,$$

и, значит, надо установить, что с вероятностью единица

$$\frac{\tilde{\xi}_1 + \dots + \tilde{\xi}_n}{n} \rightarrow 0,$$

где

$$\tilde{\xi}_n = \xi_n^* - M\xi_n^*.$$

Для доказательства этого достаточно лишь убедиться в том, что для последовательности $\tilde{\xi}_1, \tilde{\xi}_2, \dots$ выполнено условие $\sum_{n=1}^{\infty} \frac{M\tilde{\xi}_n^2}{n^2} < \infty$ теоремы I.

Имеем

$$\begin{aligned} \sum_{n=1}^{\infty} \frac{M\tilde{\xi}_n^2}{n^2} &\leq \sum_{n=1}^{\infty} \frac{M(\xi_n^*)^2}{n^2} = \sum_{n=1}^{\infty} \frac{1}{n^2} \int_{|x| \leq n} x^2 F(dx) = \\ &= \sum_{n=1}^{\infty} \frac{1}{n^2} \left\{ \sum_{k=1}^{\infty} \int_{\{k-1 < |x| \leq k\}} x^2 F(dx) \right\} \leq \\ &\leq \sum_{n=1}^{\infty} \frac{1}{n^2} \sum_{k=1}^n k \int_{\{k-1 < |x| \leq k\}} |x| F(dx) = \\ &= \sum_{k=1}^{\infty} k \int_{\{k-1 < |x| \leq k\}} |x| F(dx) \sum_{n \geq k} \frac{1}{n^2} \leq \\ &\leq 2 \sum_{k=1}^{\infty} \int_{\{k-1 < |x| \leq k\}} |x| F(dx) = 2M|\xi_1| < \infty. \end{aligned}$$

Итак, условие (2) теоремы I выполнено, и, следовательно, с вероятностью единица $\frac{\xi_1 + \dots + \xi_n}{n} \rightarrow 0, n \rightarrow \infty$.

Пусть теперь с вероятностью единица $\frac{s_n}{n} \rightarrow M\xi_1, n \rightarrow \infty$.

Тогда из равенства

$$\frac{\xi_n}{n} = \frac{s_n}{n} - \left(\frac{n-1}{n} \right) \frac{s_{n-1}}{n-1}$$

следует, что с вероятностью единица $\frac{\xi_n}{n} \rightarrow 0, n \rightarrow \infty$. Поэтому для всякого $\varepsilon > 0$

$$P \left\{ \left| \frac{\xi_n}{n} \right| > \varepsilon \text{ для бесконечно многих } n \right\} = 0,$$

что эквивалентно по лемме Бореля — Кантелли неравенству

$$\sum_{n=1}^{\infty} P \left\{ \left| \frac{\xi_n}{n} \right| > \varepsilon \right\} < \infty.$$

Но в случае одинаково распределенных величин это условие равносильно, как легко проверить, условию $M |\xi_1| < \infty$.

Теорема III доказана.

Пусть теперь опять ξ_1, ξ_2, \dots — последовательность независимых случайных величин. Тогда вероятность сходимости ряда $\sum_{n=1}^{\infty} \xi_n$ равна либо нулю, либо единице (ср. с теоремой Дополнения).

Теорема IV. Для сходимости с вероятностью единица ряда

$$\sum_{n=1}^{\infty} \xi_n$$

достаточно одновременной сходимости двух рядов

$$\sum_{n=1}^{\infty} M\xi_n, \quad \sum_{n=1}^{\infty} D\xi_n.$$

Если к тому же случайные величины ξ_1, ξ_2, \dots равномерно ограничены, $|\xi_n| \leq C, n = 1, 2, \dots$, то это условие является и необходимым.

Доказательство. Из сходимости ряда $\sum_{n=1}^{\infty} D\xi_n$ и неравенства (3) следует, что с вероятностью единица сходится ряд $\sum_{n=1}^{\infty} (\xi_n - M\xi_n)$. Вместе со сходимостью ряда $\sum_{n=1}^{\infty} M\xi_n$ отсюда вытекает требуемое утверждение

о сходимости с вероятностью единица ряда $\sum_{n=1}^{\infty} \xi_n$.

Пусть теперь величины ξ_1, ξ_2, \dots равномерно ограничены и ряд $\sum_{n=1}^{\infty} \xi_n$ сходится с вероятностью единица.

Построим последовательность независимых случайных величин $\tilde{\xi}_1, \tilde{\xi}_2, \dots$, имеющих те же самые распределения, что и ξ_1, ξ_2, \dots . Тогда с вероятностью единица

сходятся также и каждый из рядов $\sum_{n=1}^{\infty} \tilde{\xi}_n, \sum_{n=1}^{\infty} (\xi_n - \tilde{\xi}_n)$.

Поскольку $M(\xi_n - \tilde{\xi}_n) = 0$, то из сходимости с вероятностью единица ряда $\sum_{n=1}^{\infty} (\xi_n - \tilde{\xi}_n)$ и неравенства (4)

следует, что $\sum_{n=1}^{\infty} D(\xi_n - \tilde{\xi}_n) < \infty$. А значит, и $\sum_{n=1}^{\infty} D\xi_n =$

$= \frac{1}{2} \sum_{n=1}^{\infty} D(\xi_n - \tilde{\xi}_n) < \infty$. Тогда в силу первой части теоремы с вероятностью единица сходится ряд

$\sum_{n=1}^{\infty} (\xi_n - M\xi_n)$, что вместе со сходимостью ряда $\sum_{n=1}^{\infty} \xi_n$

обеспечивает сходимость ряда $\sum_{n=1}^{\infty} M\xi_n$.

Теорема IV доказана.

Пусть

$$\xi_n^C = \begin{cases} \xi_n & \text{в случае } |\xi_n| \leq C, \\ 0 & \text{в случае } |\xi_n| > C. \end{cases}$$

Т е о р е м а V. Для сходимости ряда

$$\sum_{n=1}^{\infty} \xi_n,$$

из независимых случайных величин ξ_1, ξ_2, \dots необходима и достаточна одновременная сходимость при некотором $C > 0$ каждого из трех рядов ¹⁾

$$\sum_{n=1}^{\infty} P\{|\xi_n| > C\}, \quad \sum_{n=1}^{\infty} M\xi_n^C, \quad \sum_{n=1}^{\infty} D\xi_n^C.$$

Доказательство. **Достаточность.** По-

скольку $\sum_{n=1}^{\infty} P\{|\xi_n| > C\} < \infty$, то по лемме Бореля —

Кантелли почти наверное $\xi_n = \xi_n^C$ для всех n , за исключением, быть может, конечного числа. Тогда по теореме 4 с вероятностью единица сходится ряд $\sum_{n=1}^{\infty} \xi_n^C$, а значит, и ряд $\sum_{n=1}^{\infty} \xi_n$.

Необходимость. Если ряд $\sum_{n=1}^{\infty} \xi_n$ сходится с вероятностью единица, то $\xi_n \rightarrow 0$ почти наверное, $n \rightarrow \infty$.

¹⁾ Ср. A. Khintchine und A. Kolmogoroff, Über Konvergenz von Reihen, Матем. сб., т. 32 (1925), 668—677.

Поэтому для всякого $C > 0$ может произойти самое большее конечное число событий $\{|\xi_n| > C\}$, а значит,

по лемме Бореля — Кантелли $\sum_{n=1}^{\infty} P \{|\xi_n| > C\} < \infty$. Далее,

из сходимости ряда $\sum_{n=1}^{\infty} \xi_n$ следует, что сходится также

и ряд $\sum_{n=1}^{\infty} \xi_n^C$. Тогда сходимость рядов $\sum_{n=1}^{\infty} M \xi_n^C$ и

$\sum_{n=1}^{\infty} D \xi_n^C$ вытекает из теоремы IV.

ДОПОЛНЕНИЕ

ОДНА ЗАМЕЧАТЕЛЬНАЯ ТЕОРЕМА ТЕОРИИ ВЕРОЯТНОСТЕЙ

Наблюдались уже многие случаи, в которых известные предельные вероятности с необходимостью равны нулю или единице. Например, вероятность сходимости ряда из независимых случайных величин может принимать только эти два значения¹⁾. Мы докажем теперь общую теорему, обнимающую много таких случаев.

Т е о р е м а (закон «0 или 1»). Пусть ξ_1, ξ_2, \dots — какие-либо случайные величины, а $f(x) = f(x_1, x_2, \dots)$ — бэровская функция²⁾ переменных $x = (x_1, x_2, \dots)$ такая, что условная вероятность

$$P\{f(\xi_1, \xi_2, \dots) = 0 \mid \xi_1, \xi_2, \dots, \xi_n\}$$

соотношения

$$f(\xi_1, \xi_2, \dots) = 0$$

при известных n первых величинах $\xi_1, \xi_2, \dots, \xi_n$ остается равной абсолютной вероятности

$$P\{f(\xi_1, \xi_2, \dots) = 0\} \quad (1)$$

¹⁾ Ср. главу шестую, § 5. То же имеет место и для вероятности

$$P\{\eta_n - d_n \rightarrow 0, n \rightarrow \infty\}$$

в усиленном законе больших чисел, по крайней мере когда величины ξ_n взаимно независимы.

²⁾ Под бэровской функцией понимается функция, которая может быть представлена, исходя из полиномов, через последовательные итерированные предельные переходы.

для каждого n . При этих условиях вероятность (1) равна нулю или единице.

В частности, предпосылки этой теоремы выполняются, если случайные величины ξ_1, ξ_2, \dots независимы и значение функции $f(x)$ остается неизменным при изменении лишь конечного числа величин x_n .

Доказательство. Обозначим $\xi = (\xi_1, \xi_2, \dots)$, и пусть

$$A = \{\omega : f(\xi) = 0\}.$$

Наряду с этим событием мы рассмотрим алгебру \mathfrak{M} всех событий, которые могут быть определены через какие-либо соотношения между конечным числом величин ξ_n . Если событие B принадлежит \mathfrak{M} , то по условиям теоремы

$$P(A | B) = P(A). \quad (2)$$

В случае $P(A) = 0$ наша теорема уже доказана. Пусть теперь $P(A) > 0$. Тогда из (2) следует формула

$$P(B | A) = \frac{P(A | B)P(B)}{P(A)} = P(B). \quad (3)$$

Итак, $P(B)$ и $P(B | A)$ — две счетно-аддитивные функции множеств, совпадающие на \mathfrak{M} , следовательно, они должны оставаться равными друг другу на каждом множестве борелевского расширения $\sigma(\mathfrak{M})$ алгебры \mathfrak{M} . Поэтому, в частности,

$$P(A) = P(A | A) = 1.$$

Теорема доказана.

Некоторые другие случаи, в которых можно утверждать об известных вероятностях, что они принимают только значения нуль и единица, были открыты П. Леви¹⁾.

¹⁾ См. по этому поводу P. L é v i, Sur un théorème de A. Khintchine, Bull. des Sci. Math., т. 55 (1931), 145—160, теорема II.

ЛИТЕРАТУРА

- [1] Б е р н ш т е й н С. Н., Опыт аксиоматического обоснования теории вероятностей, Зап. Харьк. матем. об-ва, 1917, стр. 209—274.
- [2] — Теория вероятностей, 2-е изд., Москва, ГТТИ, 1934.
- [1] B o r e l E., Les probabilités dénombrables et leurs applications arithmétiques, Rend. Circ. mat., Palermo, т. 27, 1909, стр. 247—271.
- [2] — Principes et formules classiques, Fasc. 1 du tome I du Traité des Probabilités par E. Borel et divers auteurs, Paris, Cauchier-Villars, 1925.
- [3] — Applications à l'arithmétique et à la théorie des fonctions, Fasc. 1 du tome II du Traité des Probabilités par E. Borel et divers auteurs, Paris, Cauchier-Villars, 1926.
- [1] C a n t e l l i F. P., Una teoria astratta del Calcolo delle probabilità, Giorn. Ist. Ital. Attuari, т. 3, 1932, стр. 257—265.
- [2] — Sulla legge dei grandi numeri, Mem. Acad. Lincei, т. 11, 1916.
- [3] — Sulla probabilità come limite della frequenza, Rend. Accad. Lincei, т. 26, 1917, стр. 39—45.
- [1] C o p e l a n d H., The theory of probability from the point of view of admissible numbers, Ann. Math. Statist., т. 3, 1932, стр. 143—156.
- [1] D ö r g e K., Zu der von Mises gegebenen Begründung der Wahrscheinlichkeitsrechnung, Math. Z., т. 32, 1930, стр. 232—258.
- [1] F r é c h e t M., Sur la convergence en probabilité, Metron, т. 8, 1930, стр. 1—48.
- [2] — Recherches théoriques modernes, sur théorie des probabilités, Fasc. 3 du tome I du Traité des probabilités par E. Borel et divers auteurs, Paris, Gauthier-Villars.
- [1] K o l m o g o r o f f A., Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung, Math. Ann., т. 104, 1931, стр. 415—458.

- [2] К о л м о г о р о в А., Общая теория меры и теория вероятностей. Сборник трудов секции точных наук Коммунистической академии, т. 1, 1929, стр. 8—21.
- [1] L é v y P., Calcul des Probabilités, Paris, Gauthier-Villars, 1925.
- [1] L o m n i c k i A., Nouveaux fondements du calcul des probabilités, Fundam. Math., т. 4, 1923, стр. 34—71.
- [1] v o n M i s e s R., Wahrscheinlichkeitsrechnung, Leipzig u. Wien, Fr. Deuticke, 1931.
- [2] — Grundlagen der Wahrscheinlichkeitsrechnung, Math. Z., т. 5, 1919, стр. 52—99.
- [3] — Wahrscheinlichkeitsrechnung, Statistik und Wahrheit. Wien, Julius Springer, 1928.
- [1] R e i c h e n b a c h H., Axiomatik der Wahrscheinlichkeitsrechnung, Math. Z., т. 34, 1932, стр. 568—619.
- [1] S l u t s k y E., Über stochastische Asymptoten und Grenzwerte, Metron, т. 5, 1925, стр. 3—89.
- [2] С л у ц к и й Е., К вопросу о логических основах теории вероятности, Вестник статистики, т. 12, 1922, стр. 13—21.
- [1] S t e i n h a n s H., Les probabilités dénombrables et leur rapport à la théorie de la mesure, Fundam. Math., т. 4, 1923, стр. 286—310.
- [1] T o r n i e r E., Wahrscheinlichkeitsrechnung und Zahlentheorie, J. reine angew. Math., т. 160, 1929, стр. 177—198.
- [2] — Grundlagen der Wahrscheinlichkeitsrechnung, Acta math, т. 60, 1933, стр. 239—380.

СОДЕРЖАНИЕ

Предисловие к первому изданию	5
Предисловие ко второму изданию	7

I. Элементарная теория вероятностей

§ 1. Аксиомы	10
§ 2. Отношение к данным опыта	12
§ 3. Терминологические замечания	14
§ 4. Непосредственные следствия из аксиом, условные вероятности, теорема Байеса	15
§ 5. Независимость	17
§ 6. Условные вероятности как случайные величины; цепи Маркова	23

II. Бесконечные поля вероятностей

§ 1. Аксиома непрерывности	26
§ 2. Борелевские поля вероятностей	29
§ 3. Примеры бесконечных полей вероятностей	31

III. Случайные величины

§ 1. Вероятностные функции	36
§ 2. Определение случайных величин, функции распределения	38
§ 3. Многомерные функции распределения	41
§ 4. Вероятности в бесконечномерных пространствах	44
§ 5. Эквивалентные случайные величины, разные виды сходимости	52

IV. Математические ожидания

§ 1. Абстрактные интегралы Лебега	57
§ 2. Абсолютные и условные математические ожидания	60
§ 3. Неравенство Чебышева	63
§ 4. Некоторые признаки сходимости	65
§ 5. Дифференцирование и интегрирование математических ожиданий по параметру	66

V. Условные вероятности и математические ожидания

§ 1. Условные вероятности	70
§ 2. Объяснение одного парадокса Бореля	75
§ 3. Условные вероятности относительно случайной величины	76
§ 4. Условные математические ожидания	78

VI. Независимость. Закон больших чисел

§ 1. Независимость	83
§ 2. Независимые случайные величины	85
§ 3. Закон больших чисел	88
§ 4. Замечания к понятию математического ожидания	100
§ 5. Усиленный закон больших чисел, сходимость рядов	104

Дополнение. Одна замечательная теорема теории вероятностей	116
--	-----

Литература	118
----------------------	-----

Андрей Николаевич Колмогоров

**ОСНОВНЫЕ ПОНЯТИЯ
ТЕОРИИ ВЕРОЯТНОСТЕЙ**

**(Серия: «Теория вероятностей
и математическая статистика»)**

М., 1974 г., 120 стр.