

СТАТИСТИЧЕСКАЯ НЕЗАВИСИМОСТЬ В ТЕОРИИ ВЕРОЯТНОСТЕЙ, АНАЛИЗЕ И ТЕОРИИ ЧИСЕЛ

В книге излагаются в очень доступной и увлекательной форме применения некоторых идей теории вероятностей в других областях математики. Основная часть книги посвящена понятию статистической независимости. Автору удалось показать, как это понятие возникает в разных видах в различных математических дисциплинах.

Книга будет полезной и интересной для студентов, она представит несомненный интерес также для специалистов—математиков, физиков и инженеров, занимающихся приложениями теории вероятностей.

ОГЛАВЛЕНИЕ

Глава 1. ОТ ВИЕТА К ПОНЯТИЮ СТАТИСТИЧЕСКОЙ НЕЗАВИСИМОСТИ	13	5. Закон природы или математическая теорема?	73
1. Формула Виета	13	Задачи	81
2. Другой взгляд на формулу Виета	14	Литература	81
3. Случайность или начало чего-либо более глубокого?	17	Глава 4. ПРОСТЫЕ ЧИСЛА «ИГРАЮТ В АЗАРТНУЮ ИГРУ»	82
4. $\left(\frac{1}{2}\right)^2 = \frac{1}{2} \dots \frac{1}{2}$ (n раз)	19	1. Теоретико-числовые функции, плотность, независимость	82
5. Герб или решетка?	21	2. Статистика значений ϕ -функций Эйлера	83
6. Независимость и «независимость»	24	Задачи	94
Задачи	26	3. Другое применение	97
Глава 2. БОРЕЛЬ И ПОСЛЕ НЕГО	29	4. Почти каждое целое m имеет приблизленно $\log \log m$ простых делителей	106
1. «Законы больших чисел»	29	Задачи	110
2. Борель и «нормальные числа»	32	5. Нормальный закон в теории чисел	110
Задачи	36	Задачи	116
3. «Герб или решетка»—более абстрактное изложение	40	Литература	116
4. В чем ценность абстракции?	43	Глава 5. ОТ КИНЕТИЧЕСКОЙ ТЕОРИИ К НЕПРЕРЫВНЫМ ДРОБЯМ	118
5. Пример 1. Сходимость ряда со случайными знаками	45	1. Парадоксы кинетической теории	118
6. Пример 2. Расходимость ряда со случайными знаками	53	2. Предварительные сведения	119
Задачи	57	3. Ответ Больпмана	123
Литература	58	4. Абстрактное изложение	125
Глава 3. НОРМАЛЬНЫЙ ЗАКОН	60	5. Эргодическая теорема и непрерывные дроби	130
1. Муавр	60	Задачи	134
2. Основная идея метода	61	Литература	135
3. Метод Маркова становится строгим	63	ДОБАВЛЕНИЕ Ю. В. Прохоров	136
Задачи	66		
4. Более внимательный взгляд на метод	67		
Задачи	70		

ПРЕДИСЛОВИЕ К РУССКОМУ ИЗДАНИЮ

Издательство иностранной литературы любезно предложило мне написать предисловие к русскому переводу моей маленькой монографии «Статистическая независимость в теории вероятностей, анализе и теории чисел», и я с большим удовольствием выполняю эту просьбу.

Когда переводится твоя работа — это всегда приятно. В особенности приятно, когда переводится на русский язык работа, связанная с теорией вероятностей. Тогда чувствуешь, что твои труды становятся доступными русским читателям и что они продолжают великие традиции Чебышева, Маркова, Ляпунова, Бернштейна и Хинчина, блестяще развиваемые Колмогоровым и многими его сотрудниками и учениками.

В предисловии к английскому изданию я указал цели, с которыми написана эта книга. Я ничего не могу больше добавить, хочу только еще раз подчеркнуть, что книга написана в основном для молодежи, стоящей на пороге огромного и удивительного мира математики.

Многие идеи и способы изложения были задуманы мной, когда я был студентом, и я попытался поделиться с читателями моими собственными волнениями на пороге этого мира. То, что теперь я могу поделиться ими с читателями в Советском Союзе, является для меня источником большой радости и удовлетворения.

Я считаю своим очень приятным долгом поблагодарить профессора Прохорова за перевод, а также за добавление и примечания к моей книге. Любой автор может только мечтать, чтобы его книгу переводил и редактировал такой известный и высококвалифицированный специалист.

М. Кац

Нью-Йорк, июль 1962 г.

ПРЕДИСЛОВИЕ АВТОРА

Во время сессии Американского математического общества, происходившей летом 1955 года, мне была предоставлена возможность прочитать небольшой цикл лекций на чтениях в честь Хедрика (Hedrick Lectures). Я был весьма обрадован, когда несколько позже профессор Т. Радо от имени комитета по изданию серии «Carus Monographs» любезно попросил меня изложить мои лекции в форме монографии.

Через некоторое время я удостоился чести быть приглашенным Хаверфордским колледжем прочитать ряд лекций в этом колледже. Это приглашение дало мне благоприятную возможность испытать задуманную монографию на «живой» аудитории, и настоящая книга является лишь незначительно измененным текстом моих лекций, читанных в Хаверфордском колледже в течение весеннего семестра 1958 года.

Как и в первоначальных лекциях, так и в этом расширенном варианте моей основной целью было показать, что: (а) крайне простые наблюдения часто являются отправной точкой обширных и плодотворных исследова-

ний и (б) многие, на вид не связанные, выводы в действительности оказываются вариациями одной и той же простой темы.

За исключением последней главы, где я имел дело с эффектным применением эргодической теоремы к непрерывным дробям, книга посвящена понятию статистической независимости.

Это понятие возникло в теории вероятностей, и долгое время им пользовались, не понимая четко его сути, что вызвало подозрение в некорректности этого математического понятия.

Теперь мы знаем, как определять статистическую независимость в более общих и отвлеченных терминах. Однако современное стремление к общности и отвлеченности приводит не только к тому, что от внимания ускользает простота первоначальной идеи, но и к тому, что становится неясной возможность применения вероятностных идей вне сферы теории вероятностей.

На последующих страницах я попытался спасти статистическую независимость от этой опасности, показав как в своей простейшей форме она возникает в различных контекстах в нескольких математических дисциплинах.

Что касается степени подготовленности читателей книги, то я предполагаю знакомство с теорией меры и интеграла Лебега, элементарной теорией интегралов Фурье и начальными основами теории чисел. Так как я не хотел предполагать большего, а также не хотел загромождать рассказ слишком многими техническими

деталями, я опустил доказательства некоторых утверждений.

Я прошу простить мне эти пропуски и надеюсь, что читатель достаточно заинтересуется предметом, чтобы восполнить имеющиеся пробелы. Для этого я прилагаю небольшую библиографию, которая не претендует на полноту.

В книгу я включил также некоторое количество задач. Эти задачи большей частью довольно сложны, и читатель не должен чувствовать себя обескураженным, если он не сумеет решить их без значительного усилия.

Я хочу поблагодарить профессоров Хаверфордского колледжа К. Окли и Р. Уиснера за превосходное сотрудничество и за превращение моего путешествия от Итаки до Хаверфорда в истинное удовольствие.

Я был счастлив иметь в числе своих слушателей профессора Пенсильванского университета Г. Радемахера и профессора Джона Окстоби из Брин-Марского колледжа. Их критика, советы и постоянная поддержка поистине неоценимы, и мой долг им велик.

Мои коллеги по Корнелльскому университету, профессор Х. Уидом и М. Шрейбер прочли рукопись и предложили большое количество изменений и усовершенствований. Поблагодарить их за помощь я считаю удовольствием.

Я приношу благодарность также студентам Хаверфордского и Брин-Марского колледжей, которые выступили в качестве «подопытных морских свинок», и особенно Дж. Райлу, составившему библиографию и читавшему корректуру рукописи.

Наконец, не в меньшей степени, я хочу выразить благодарность м-сс Аксельсон из Хаверфордского колледжа и м-с Мартин из Математического отделения Корнелльского университета за решение часто неразрешимой задачи перепечатки рукописи по моим почти неразборчивым записям.

Марк Кац

Итака, Нью-Йорк,
сентябрь 1959 г.

Моему учителю

профессору Гуго Штейнгаузу

ОТ ВИЕТА К ПОНЯТИЮ СТАТИСТИЧЕСКОЙ НЕЗАВИСИМОСТИ

1. Формула Виета. Мы начнем с простой тригонометрии. Запишем

$$\begin{aligned}
 \sin x &= 2 \sin \frac{x}{2} \cos \frac{x}{2} = 2^2 \sin \frac{x}{4} \cos \frac{x}{4} \cos \frac{x}{2} = \\
 &= 2^3 \sin \frac{x}{8} \cos \frac{x}{8} \cos \frac{x}{4} \cos \frac{x}{2} = \\
 &\quad \cdot \\
 &\quad \cdot \\
 &\quad \cdot \\
 &= 2^n \sin \frac{x}{2^n} \prod_{k=1}^n \cos \frac{x}{2^k}. \tag{1.1}
 \end{aligned}$$

Из элементарного анализа мы знаем, что при $x \neq 0$

$$1 = \lim_{n \rightarrow \infty} \frac{\sin \frac{x}{2^n}}{\frac{x}{2^n}} = \frac{1}{x} \lim_{n \rightarrow \infty} 2^n \sin \frac{x}{2^n}$$

и, следовательно,

$$\lim_{n \rightarrow \infty} 2^n \sin \frac{x}{2^n} = x. \tag{1.2}$$

Сопоставляя (1.2) с (1.1), получаем

$$\frac{\sin x}{x} = \prod_{k=1}^{\infty} \cos \frac{x}{2^k}. \quad (1.3)$$

Особенно интересен один частный случай соотношения (1.3). Полагая $x = \pi/2$, находим

$$\frac{2}{\pi} = \prod_{n=1}^{\infty} \cos \frac{\pi}{2^{n+1}} = \frac{\sqrt{2}}{2} \frac{\sqrt{2+\sqrt{2}}}{2} \frac{\sqrt{2+\sqrt{2+\sqrt{2}}}}{2} \dots, \quad (1.4)$$

что является классической формулой, принадлежащей Виету.

2. Другой взгляд на формулу Виета. До сих пор все было легким и хорошо знакомым.

Рассмотрим теперь (1.3) с другой точки зрения.

Известно, что любое действительное число t , $0 \leq t \leq 1$, может быть *однозначно* представлено в виде

$$t = \frac{\varepsilon_1}{2} + \frac{\varepsilon_2}{2^2} + \dots, \quad (2.1)$$

где каждая из величин ε есть или 0, или 1.

Это известное *двоичное разложение* t , и чтобы обеспечить единственность, условимся записывать обрывающиеся разложения в форме, в которой все двоичные цифры, начиная с некоторого места, равны 0. Так, например, запишем

$$\frac{3}{4} = \frac{1}{2} + \frac{1}{2^2} + \frac{0}{2^3} + \frac{0}{2^4} + \dots,$$

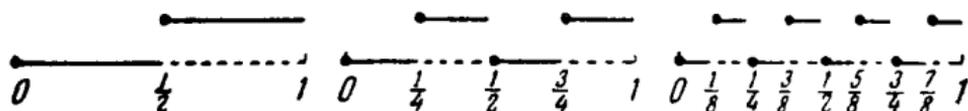
а не

$$\frac{3}{4} = \frac{1}{2} + \frac{0}{2^2} + \frac{1}{2^3} + \frac{1}{2^4} + \dots$$

Двоичные цифры ε_i являются, конечно, функциями от t , поэтому представление (2.1) более точно должно выглядеть как

$$t = \frac{\varepsilon_1(t)}{2} + \frac{\varepsilon_2(t)}{2^2} + \frac{\varepsilon_3(t)}{2^3} + \dots \quad (2.2)$$

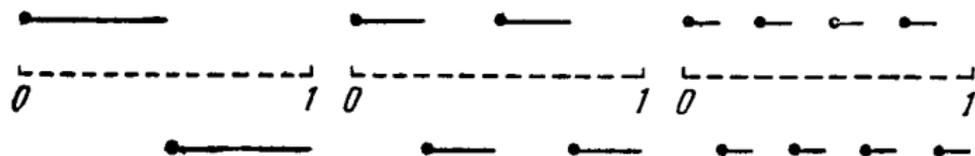
При нашем соглашении относительно записи обрывающихся разложений графики функций $\varepsilon_1(t)$, $\varepsilon_2(t)$, $\varepsilon_3(t)$, ... выглядят следующим образом:



Удобнее ввести функции $r_k(t)$, определяемые равенствами

$$r_k(t) = 1 - 2\varepsilon_k(t), \quad k = 1, 2, 3, \dots, \quad (2.3)$$

и имеющие следующее графическое изображение:



Эти функции, впервые введенные и изученные Г. Радемахером, называются функциями Радемахера. В тер-

минах функций $r_k(t)$ мы можем переписать представление (2.2) в виде

$$1 - 2t = \sum_{k=1}^{\infty} \frac{r_k(t)}{2^k}. \quad (2.4)$$

Заметим теперь, что

$$\int_0^1 e^{ix(1-2t)} dt = \frac{\sin x}{x}$$

и

$$\int_0^1 \exp\left(ix \frac{r_k(t)}{2^k}\right) dt = \cos \frac{x}{2^k}.$$

При этом формула (1.3) превратится в

$$\begin{aligned} \frac{\sin x}{x} &= \int_0^1 e^{ix(1-2t)} dt = \int_0^1 \exp\left(ix \sum_{k=1}^{\infty} \frac{r_k(t)}{2^k}\right) dt = \\ &= \prod_{k=1}^{\infty} \cos \frac{x}{2^k} = \prod_{k=1}^{\infty} \int_0^1 \exp\left(ix \frac{r_k(t)}{2^k}\right) dt \end{aligned}$$

и, в частности, будет иметь место равенство

$$\int_0^1 \prod_{k=1}^{\infty} \exp\left(ix \frac{r_k(t)}{2^k}\right) dt = \prod_{k=1}^{\infty} \int_0^1 \exp\left(ix \frac{r_k(t)}{2^k}\right) dt. \quad (2.5)$$

Интеграл от произведения равен произведению интегралов!

3. Случайность или начало чего-либо более глубокого? Можем ли мы рассматривать равенство (2.5) как случайное совпадение? Конечно нет, до тех пор, пока мы не исследуем вопрос более тщательно.

Взглянем на функцию

$$\sum_{k=1}^n c_k r_k(t).$$

Это ступенчатая функция, которая постоянна на интервалах

$$\left(\frac{s}{2^n}, \frac{s+1}{2^n} \right), \quad s = 0, 1, \dots, 2^n - 1,$$

и значениями которой являются числа

$$\pm c_1 \pm c_2 \pm \dots \pm c_n.$$

Каждая последовательность (длины n), состоящая из $+1$ и -1 , соответствует одному и только одному интервалу $(s/2^n, (s+1)/2^n)$. Таким образом,

$$\int_0^1 \exp \left[i \sum_1^n c_k r_k(t) \right] dt = \frac{1}{2^n} \sum \exp \left(i \sum_1^n \pm c_k \right),$$

где внешняя сумма справа берется по *всем* возможным последовательностям (длины n) из $+1$ и -1 .

Теперь

$$\frac{1}{2^n} \sum \exp \left(i \sum_1^n \pm c_k \right) = \prod_{k=1}^n \left(\frac{e^{ic_k} + e^{-ic_k}}{2} \right) = \prod_{k=1}^n \cos c_k$$

и, следовательно,

$$\int_0^1 \exp \left[i \sum_1^n c_k r_k(t) \right] dt = \prod_{k=1}^n \cos c_k = \prod_{k=1}^n \int_0^1 e^{i c_k r_k(t)} dt. \quad (3.1)$$

Полагая

$$c_k = \frac{x}{2^k},$$

получаем

$$\int_0^1 \exp \left(i x \sum_1^n \frac{r_k(t)}{2^k} \right) dt = \prod_{k=1}^n \cos \frac{x}{2^k},$$

и так как

$$\lim_{n \rightarrow \infty} \sum_1^n \frac{r_k(t)}{2^k} = 1 - 2t,$$

причем сходимость в левой части *равномерна* на $(0, 1)$, то мы имеем

$$\begin{aligned} \frac{\sin x}{x} &= \int_0^1 e^{i x (1-2t)} dt = \lim_{n \rightarrow \infty} \int_0^1 \exp \left(i x \sum_1^n \frac{r_k(t)}{2^k} \right) dt = \\ &= \lim_{n \rightarrow \infty} \prod_{k=1}^n \cos \frac{x}{2^k} = \prod_{k=1}^{\infty} \cos \frac{x}{2^k}. \end{aligned}$$

Таким образом, мы установили другое доказательство формулы (1.3). Лучше ли оно, чем доказательство, данное в п. 1?

Данное доказательство более сложно, но в то же время более поучительно, так как оно как-то связывает формулу Виета с двоичными цифрами.

Какое же свойство двоичных цифр приводит к успеху?

4. $\left(\frac{1}{2}\right)^n = \frac{1}{2} \cdot \dots \cdot \frac{1}{2}$ (n раз). Рассмотрим множество тех t , для которых

$$r_1(t) = +1, \quad r_2(t) = -1, \quad r_3(t) = -1.$$

Одного взгляда на графики r_1 , r_2 и r_3 достаточно, чтобы установить, что это множество (за исключением, быть может, конечных точек) — просто интервал

$$\left(\frac{3}{8}, \frac{4}{8}\right).$$

Очевидно, что длина (или мера) этого интервала равна $\frac{1}{8}$ и

$$\frac{1}{8} = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2}.$$

Это тривиальное наблюдение может быть записано в следующей форме:

$$\begin{aligned} \mu \{r_1(t) = +1, r_2(t) = -1, r_3(t) = -1\} = \\ = \mu \{r_1(t) = +1\} \mu \{r_2(t) = -1\} \mu \{r_3(t) = -1\}, \end{aligned}$$

где μ обозначает меру (длину) множества, определяемого выражением внутри скобок.

Читатель без труда сможет распространить это наблюдение на случай произвольного числа функций r . Он получит следующий результат: если $\delta_1, \dots, \delta_n$ — последовательность, состоящая из $+1$ и -1 , то

$$\begin{aligned} & \mu \{r_1(t) = \delta_1, \dots, r_n(t) = \delta_n\} = \\ & = \mu \{r_1(t) = \delta_1\} \mu \{r_2(t) = \delta_2\} \dots \mu \{r_n(t) = \delta_n\}. \end{aligned}$$

Может показаться, что мы всего лишь усложненным способом записали равенство

$$\left(\frac{1}{2}\right)^n = \frac{1}{2} \times \frac{1}{2} \times \dots \times \frac{1}{2} \quad (n \text{ раз}),$$

но в действительности это значительно больше. Полученный результат выражает глубокое свойство функций $r_k(t)$ (и, следовательно, двоичных цифр) и служит отправной точкой обширных и плодотворных исследований. Именно это свойство лежит в основе доказательства п. 3. Формула (3.1) может быть теперь доказана следующим образом:

$$\begin{aligned} & \int_0^1 \exp \left[i \sum_1^n c_k r_k(t) \right] dt = \\ & = \sum_{\delta_1, \dots, \delta_n} \exp \left(i \sum_1^n c_k \delta_k \right) \mu \{r_1(t) = \delta_1, \dots, r_n(t) = \delta_n\} = \\ & = \sum_{\delta_1, \dots, \delta_n} \prod_1^n e^{i c_k \delta_k} \prod_1^n \mu \{r_k(t) = \delta_k\} = \end{aligned}$$

$$\begin{aligned}
 &= \sum_{\delta_1, \dots, \delta_n} \prod_{k=1}^n e^{ic_k \delta_k} \mu \{r_k(t) = \delta_k\} = \\
 &= \prod_{k=1}^n \sum_{\delta_k} e^{ic_k \delta_k} \mu \{r_k(t) = \delta_k\} = \prod_{k=1}^n \int_0^1 e^{ic_k r_k(t)} dt.
 \end{aligned}$$

5. Герб или решетка? Элементарное исследование опыта с бросанием монеты начинается с двух предположений:

а) монета симметрична;

б) последовательные бросания монеты *независимы*.

Первое предположение означает, что в каждом отдельном бросании исходы H (герб) и T (решетка) равновероятны¹⁾, т. е. каждому из них приписывается «вероятность» $\frac{1}{2}$. Второе предположение позволяет при-

менить «правило умножения вероятностей». Это правило, грубо говоря, таково: если события A_1, \dots, A_n *независимы*, то вероятность их совместного появления равна произведению вероятностей появления отдельных событий. Другими словами,

$$\begin{aligned}
 &\text{Вер. } \{A_1 \text{ и } A_2 \text{ и } A_3 \dots \text{ и } A_n\} = \\
 &= \text{Вер. } \{A_1\} \text{ Вер. } \{A_2\} \dots \text{ Вер. } \{A_n\}. \quad (5.1)
 \end{aligned}$$

¹⁾ H и T от английских «head» и «tail». — Прим. перев.

Примененное к независимым бросаниям симметричной монеты, это правило позволяет сделать заключение о том, что вероятность появления последовательности (длины n), состоящей из альтернативных исходов H и T (например, $HNHTT\dots T$), равна

$$\frac{1}{2} \times \frac{1}{2} \times \dots \times \frac{1}{2} = \frac{1}{2^n}.$$

Это напоминает выводы п. 4, и мы можем использовать функции $r_k(t)$ в качестве *модели* для опыта с бросанием монеты. Для достижения этой цели составим следующий словарь терминов:

символ H	+ 1
символ T	- 1
k -е бросание ($k = 1, 2, \dots$)	$r_k(t)$ ($k = 1, 2, \dots$)
событие	некоторое множество значений t
вероятность события	мера соответствующего множества значений t

Чтобы понять, как применять этот словарь, рассмотрим следующую задачу. Найти вероятность того, что при n независимых бросаниях симметричной монеты герб выпадает точно l раз. Используя словарь, мы переведем задачу, и она будет читаться в новой форме так:

Найти меру множества точек t , таких, что из n чисел $r_1(t), r_2(t), \dots, r_n(t)$ ровно l равны $+1$. Мы можем решить эту задачу (не прибегая к обычной комбинаторике) способом, с которым мы еще много раз встретимся при различных обстоятельствах впоследствии:

Прежде всего заметим следующее: условие, состоящее в том, что среди $r_1(t), \dots, r_n(t)$ ровно l равны 1 , эквивалентно условию

$$r_1(t) + r_2(t) + \dots + r_n(t) = 2l - n. \quad (5.2)$$

Далее отметим, что для целого m

$$\frac{1}{2\pi} \int_0^{2\pi} e^{imx} dx = \begin{cases} 1, & m = 0, \\ 0, & m \neq 0, \end{cases} \quad (5.3)$$

и, следовательно, функция

$$\varphi(t) = \frac{1}{2\pi} \int_0^{2\pi} e^{ix[r_1(t) + \dots + r_n(t) - (2l - n)]} dx \quad (5.4)$$

равна 1 , если удовлетворяется (5.2), и равна 0 в противном случае.

Таким образом,

$$\begin{aligned} \mu \{r_1(t) + \dots + r_n(t) = 2l - n\} &= \int_0^1 \varphi(t) dt = \\ &= \int_0^1 \frac{1}{2\pi} \int_0^{2\pi} e^{ix[r_1(t) + \dots + r_n(t) - (2l - n)]} dx dt = \\ &= \frac{1}{2\pi} \int_0^{2\pi} e^{-i(2l - n)x} \left(\int_0^1 e^{ix[r_1(t) + \dots + r_n(t)]} dt \right) dx \end{aligned}$$

(последний шаг — изменение порядка интегрирования — обыкновенно оправдывается обращением к теореме Фубини. В нашем случае обоснование этого шага тривиально, так как сумма $r_1(t) + \dots + r_n(t)$ представляет собой ступенчатую функцию).

Вспомним теперь равенство (3.1); с его помощью при $c_1 = c_2 = \dots = c_n = x$ получим

$$\begin{aligned} \mu \{r_1(t) + \dots + r_n(t) = 2l - n\} &= \\ &= \frac{1}{2\pi} \int_0^{2\pi} e^{-i(2l-n)x} \cos^n x \, dx. \end{aligned} \quad (5.5)$$

Наконец, мы оставим в качестве упражнения доказательство того, что

$$\mu \{r_1(t) + \dots + r_n(t) = 2l - n\} = \frac{1}{2^n} \binom{n}{l}. \quad (5.6)$$

6. Независимость и «независимость». Понятие независимости, хотя и является центральным по важности в теории вероятностей, не есть чисто математическое понятие. Правило умножения вероятностей независимых событий представляет собой попытку формализовать это понятие и на этой основе построить некоторое исчисление. При этом возникает склонность рассматривать события, которые кажутся не связанными, как независимые друг от друга. Так, физик, рассматривая события, происходящие в двух удаленных друг от друга пробах газа, будет считать их независимыми (как может быть иначе, если одна проба взята, скажем

в Бисмарке, Северная Дакота, а другая—в Валлингтоне, округ Колумбия?) и радостно обратится к правилу умножения вероятностей.

К сожалению, действуя таким образом, он может (непреднамеренно и произвольно) создать впечатление, что сказанное им—*строгое логическое заключение*.

В действительности то, о чем говорилось, является *определением* независимости, объединяемым с убеждением (основанным, несомненно, на наблюдениях и опыте), что это определение применимо в некоторых случаях.

Существует, таким образом, независимость в расплывчатом интуитивном понимании, и «независимость» в том узком, но точно определенном смысле, что применимо правило умножения вероятностей.

Длительное время основным мотивом и движущей силой развития теории вероятностей были неопределенные интуитивные представления.

И в то время, как создавался впечатляющий формализм, математики (за очень немногим исключением) оставались в стороне, так как не очень ясно понимали, к каким объектам формализм применим¹⁾.

¹⁾ Предположим, что книга по дифференциальным уравнениям написана исключительно в терминах масс, сил, ускорений и попадает в руки кого-нибудь, кто никогда не слышал о механике. Богатое чисто математическое содержание такой книги могло бы быть в значительной степени потеряно для этого гипотетического читателя.

Затем в 1909 г. Э. Борель заметил, что двоичные цифры $\varepsilon_k(t)$ [или, что эквивалентно, функции Радемахера $r_k(t)$] являются «независимыми» [см. (4.1)].

Наконец, появились точно определенные объекты, к которым теория вероятностей для независимых событий могла быть применена без опасности обращения с монетами, событиями, бросаниями и экспериментами.

Появление классического мемуара Бореля «Sur les probabilités dénombrables et leurs applications arithmétiques» отметило начало современной теории вероятностей, и в следующей главе мы будем обсуждать некоторые из направлений, по которым эта теория развивается.

ЗАДАЧИ

1. Записать троичное разложение t , $0 \leq t \leq 1$, в виде

$$t = \frac{\eta_1(t)}{3} + \frac{\eta_2(t)}{3^2} + \frac{\eta_3(t)}{3^3} + \dots$$

(каждое η_k может принимать значения 0, 1 и 2) и доказать, что величины η независимы.

2. Доказать, что

$$\frac{\sin x}{x} = \prod_{k=1}^{\infty} \frac{1 + 2 \cos \frac{2x}{3^k}}{3},$$

и обобщить это соотношение.

3. Показать, что если $k_1 < k_2 < \dots < k_s$, то

$$\int_0^1 r_{k_1}(t) r_{k_2}(t) \dots r_{k_s}(t) dt = 0.$$

4. Пусть число $2n$ (четное, положительное) записано в двоичной системе:

$$2n = 2^{n_1} + 2^{n_2} + \dots + 2^{n_k}, \quad 1 \leq n_1 < n_2 < \dots < n_k.$$

Зададим функции $w_n(t)$ (функции Уолша—Качмаша) следующим образом:

$$\begin{aligned} w_0(t) &= 1, \\ w_n(t) &= r_{n_1}(t) \dots r_{n_k}(t), \quad n \geq 1. \end{aligned}$$

Доказать, что

$$(a) \quad \int_0^1 w_m(t) w_n(t) dt = \delta_{m,n};$$

(б) если $f(t)$ — интегрируема и

$$\int_0^1 f(t) w_n(t) dt = 0, \quad n = 0, 1, 2, \dots$$

то $f(t) = 0$ почти всюду;

$$(в) \quad \int_0^1 \int_0^1 \left| \sum_{k=0}^{2^n} w_k(t) w_k(s) \right| dt ds = 1.$$

5. Используя формулу

$$|z| = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{1 - \cos zx}{x^2} dx,$$

доказать сначала, что

$$\int_0^1 \left| \sum_1^n r_k(t) \right| dt = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{1 - \cos^n x}{x^2} dx > \frac{1}{\pi} \int_{-1/\sqrt{n}}^{1/\sqrt{n}} \frac{1 - \cos^n x}{x^2} dx,$$

а затем, что

$$\int_0^1 \left| \sum_1^n r_k(t) \right| dt > A \sqrt{n},$$

где

$$A = \frac{1}{\pi} \int_{-1}^1 \frac{1 - e^{-y^2/2}}{y^2} dy.$$

Замечание: неравенство Шварца вместе с результатом задачи 3 при $s=2$ дает

$$\int_0^1 \left| \sum_1^n r_k(t) \right| dt \leq \sqrt{n}.$$

БОРЕЛЬ И ПОСЛЕ НЕГО

1. «Законы больших чисел». Все вы имеете представление о том, что если играть в безобидную игру, то в конце концов маловероятно, чтобы вы разбогатели. «Закон средних позаботится об этом» — вот что разумно подсказывается в этой и подобных ситуациях. Что же это за «закон средних»? Является ли он чем-то вроде физического закона или это есть чисто математическое утверждение? Большей частью все же верно последнее, хотя этот закон и хорошо согласуется с экспериментальными данными. Давайте забудем об экспериментальной стороне проблемы и сконцентрируем внимание на математических вопросах. Предположим, что я бросаю симметричную монету, выигрывая доллар каждый раз, когда появляется H и проигрывая доллар всякий раз при выпадении T . Что я могу сказать о своем выигрыше после n бросаний? Используя наш словарь п. 4 гл. 1, мы можем представить этот выигрыш так:

$$r_1(t) + r_2(t) + \dots + r_n(t). \quad (1.1)$$

Очевиден интерес игрока к следующему вопросу: каковы его шансы на то, что после n бросаний мо-

неты его выигрыш превзойдет заданное число A_n . Согласно нашему словарю, это эквивалентно задаче определения меры множества точек t , для которых

$$r_1(t) + r_2(t) + \dots + r_n(t) > A_n. \quad (1.2)$$

Если действительно маловероятно то, что я разбогатею в этой игре, то при A_n «достаточно большом» мера множества, определенного (1.2), будет «мала». (По тем же соображениям будет неправдоподобен проигрыш больший, чем A_n .) Мы придадим всем этим выводам необходимую точность, доказав следующую теорему:

При любом $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} \mu \{ |r_1(t) + \dots + r_n(t)| > \varepsilon n \} = 0. \quad (1.3)$$

Очевидный подход может быть основан на формуле (5.6) гл. 1. Действительно, мы имеем

$$\begin{aligned} & \mu \{ |r_1(t) + \dots + r_n(t)| > \varepsilon n \} = \\ & = \sum_{|2l-n| > \varepsilon n} \mu \{ r_1(t) + \dots + r_n(t) = 2l - n \} = \\ & = \sum_{|2l-n| > \varepsilon n} \frac{1}{2^n} \binom{n}{l}, \end{aligned}$$

и все сводится к доказательству того, что при любом $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} \sum_{|2l-n| > \varepsilon n} \frac{1}{2^n} \binom{n}{l} = 0. \quad (1.4)$$

Попытайтесь это сделать! Это будет не трудно, но и не очень легко, если следовать естественному влечению и использовать формулу Стирлинга. Если вы добьетесь успеха, то этим вы повторите первоначальное доказательство Бернулли. Однако существует и лучший и более легкий способ, принадлежащий Чебышеву.

Вы просто записываете

$$\begin{aligned} \int_0^1 (r_1(t) + \dots + r_n(t))^2 dt &\geq \\ &\geq \int_{|r_1(t) + \dots + r_n(t)| > \varepsilon n} (r_1(t) + \dots + r_n(t))^2 dt > \\ &> \varepsilon^2 n^2 \mu \{ |r_1(t) + \dots + r_n(t)| > \varepsilon n \}. \end{aligned} \quad (1.5)$$

Если вы решили задачу 3 в конце гл. 1, то вы сможете получить

$$\int_0^1 (r_1(t) + \dots + r_n(t))^2 dt = n \quad (1.6)$$

и, используя (1.5),

$$\mu \{ |r_1(t) + \dots + r_n(t)| > \varepsilon n \} < \frac{1}{\varepsilon^2 n}, \quad (1.7)$$

что доказывает (1.3) «с большим запасом».

Запомните этот изящный прием Чебышева; мы еще с ним встретимся!

Утверждение (1.3) дает простейший пример того, что в специальных терминах называется «слабым законом больших чисел»¹⁾. Прилагательное «слабый» не имеет намерением умалить значение, но употребляется для различения с другим законом больших чисел, называемым обыкновенно «усиленным законом». Характеристика «усиленный» не имеет целью превозношение этого закона; просто для случая игры в орлянку из него вытекает «слабый закон», и потому он сильнее в логическом смысле.

Оба закона были довольно сильно обобщены, и в своей окончательной форме ни один из них не влечет за собой другого. Это, однако, технические вопросы, которые нас в данном случае не интересуют. Математическое содержание слабого закона больших чисел относительно бедно. В форме (1.4) он превращается в занимательную теорему о биномиальных коэффициентах. Может ли тогда он быть выражением таинственного «закона средних», упоминавшегося выше? Боюсь, что так. Это по существу все, что мы можем надеяться получить от чисто математической теории.

2. Борель и «нормальные числа». Другой закон больших чисел был найден Борелем. Борель доказал, что для почти всех t (т. е. для всех точек t , исклю-

¹⁾ «the weak law of large numbers». В русской терминологии прилагательное «слабый» обычно опускают. — *Прим. перев.*

чая множество лебеговой меры 0)

$$\lim_{n \rightarrow \infty} \frac{r_1(t) + r_2(t) + \dots + r_n(t)}{n} = 0. \quad (2.1)$$

Доказательство несложно и основано на хорошо известной теореме из теории меры и интеграла Лебега. Теорема, о которой идет речь, такова:

Если $\{f_n(t)\}$ — последовательность неотрицательных интегрируемых по Лебегу функций, то из сходимости

$$\sum_{n=1}^{\infty} \int_0^1 f_n(t) dt \quad (2.2)$$

следует сходимость почти всюду ряда

$$\sum_{n=1}^{\infty} f_n(t). \quad (2.3)$$

Положим

$$f_n(t) = \left(\frac{r_1(t) + \dots + r_n(t)}{n} \right)^4 \quad (2.4)$$

и рассмотрим

$$\int_0^1 \left(\frac{r_1(t) + \dots + r_n(t)}{n} \right)^4 dt.$$

Используя результат задачи 3 в конце гл. 1, мы без труда вычислим

$$\int_0^1 \left(\frac{r_1(t) + \dots + r_n(t)}{n} \right)^4 dt = \frac{n + \frac{4!}{2! 2!} \binom{n}{2}}{n^4}$$

и, следовательно,

$$\sum_{n=1}^{\infty} \int_0^1 f_n(t) dt < \infty.$$

Из этого следует, что ряд

$$\sum_{n=1}^{\infty} \left(\frac{r_1(t) + \dots + r_n(t)}{n} \right)^4$$

сходится почти всюду и тем более

$$\lim_{n \rightarrow \infty} \left(\frac{r_1(t) + \dots + r_n(t)}{n} \right)^4 = 0$$

почти всюду. Этим доказано равенство (2.1).

Так как

$$r_k(t) = 1 - 2\varepsilon_k(t),$$

то (2.1) эквивалентно тому, что для почти всех t

$$\lim_{n \rightarrow \infty} \frac{\varepsilon_1(t) + \dots + \varepsilon_n(t)}{n} = \frac{1}{2}. \quad (2.5)$$

Другими словами, почти каждое число t имеет (асимптотически) одинаковое число нулей и единиц в соответствующем ему двоичном разложении! Это арифметическое содержание теоремы Бореля. Что дает теорема в вероятностном плане? Пользуясь нашим словарем, мы приходим к следующему заключению. Если симметричная монета бросается неограниченно долго и если бросания независимы, то частота появления гербов

(решеток) равна $1/2$ с вероятностью 1 (в пределе, конечно). Это утверждение удовлетворяет нашим интуитивным представлениям о том, что должен говорить «закон средних», и вновь убеждает в хорошем выборе нашего словаря.

Читатель, несомненно, сознает, что основание 2 в знаменателях разложения не является чем-то «неприкосновенным».

Если g — большее единицы целое число, то можно записать

$$t = \frac{\omega_1(t)}{g} + \frac{\omega_2(t)}{g^2} + \dots, \quad 0 \leq t \leq 1, \quad (2.6)$$

где каждая цифра $\omega(t)$ принимает теперь значения $0, 1, \dots, g-1$. Мы предоставляем читателю доказать, что для почти всех t ($0 \leq t \leq 1$)

$$\lim_{n \rightarrow \infty} \frac{F_n^{(h)}(t)}{n} = \frac{1}{g}, \quad (2.7)$$

где $F_n^{(h)}(t)$ обозначает количество тех величин ω , из числа n первых, которые равны k , $0 \leq k \leq g-1$. (Это задача 1 на стр. 36.)

Из того факта, что счетное объединение множеств меры 0 снова имеет меру 0, вытекает следующее: почти все числа t , $0 \leq t \leq 1$, таковы, что в любой системе счисления (т. е. при любом $g > 1$) каждая допустимая цифра появляется с определенной (и ожидаемой!) частотой. Другими словами, почти все числа «нормальны»!

Как это часто бывает, доказать, что преобладающее большинство объектов обладает определенным свойством, много легче, чем *указать* хотя бы один такой объект. Данный случай не исключение. «Нормальные» числа строить довольно трудно! Простейшим примером служит число (записанное в десятичной системе)

$$0,123456789101112131415161718192021\dots,$$

где после запятой мы последовательно записываем все положительные целые числа. Доказательство нормальности этого числа никоим образом не является тривиальным.

ЗАДАЧИ

1. Доказать (2.7), показав сначала, что величины ω независимы, а затем обобщив результат задачи 3 гл. 1.

2. Пусть $f(t)$, $0 \leq t \leq 1$, — непрерывная функция. Доказать, что

$$\lim_{n \rightarrow \infty} \int_0^1 \dots \int_0^1 f\left(\frac{x_1 + \dots + x_n}{n}\right) dx_1 \dots dx_n = f\left(\frac{1}{2}\right).$$

Указание: сначала показать, подражая чебышевскому методу доказательства (1.4), что n -мерный объем множества, определенного неравенством

$$\left| \frac{x_1 + \dots + x_n}{n} - \frac{1}{2} \right| > \varepsilon, \quad 0 \leq x_i \leq 1, \quad i = 1, 2, \dots, n,$$

меньше $1/12\varepsilon^2 n$.

3. «Несимметричная» монета. Пусть $T_p(t)$, $0 < p < 1$, определено следующим образом:

$$T_p(t) = \begin{cases} \frac{t}{p}, & 0 \leq t \leq p, \\ \frac{t-p}{1-p}, & p < t \leq 1, \end{cases}$$

и пусть

$$\varepsilon_p(t) = \begin{cases} 0, & 0 \leq t \leq p, \\ 1, & p < t \leq 1. \end{cases}$$

Начертить графики функций

$$\varepsilon_1^{(p)}(t) = \varepsilon_p(t), \quad \varepsilon_2^{(p)}(t) = \varepsilon_p(T_p(t)), \quad \varepsilon_3^{(p)}(t) = \varepsilon_p(T_p(T_p(t))) \dots$$

и показать, что они независимы. Заметьте, что при $p = \frac{1}{2}$ получаются двоичные цифры.

4. Доказать, что мера множества, на котором

$$\varepsilon_1^{(p)}(t) + \dots + \varepsilon_n^{(p)}(t) = l, \quad 0 \leq l \leq n, \quad \text{равна}$$

$$\binom{n}{l} p^l (1-p)^{n-l}.$$

5. Объяснить, как функции $\varepsilon_n^{(p)}(t)$ могут быть использованы при построении модели для независимых бросаний несимметричной монеты, где вероятность H равна p и вероятность T равна $q = 1 - p$.

6. Показать, что если $f(t)$ непрерывна, то

$$\begin{aligned} & \int_0^1 f\left(\frac{\varepsilon_1^{(p)}(t) + \dots + \varepsilon_n^{(p)}(t)}{n}\right) dt = \\ & = \sum_{k=0}^n f\left(\frac{k}{n}\right) p^k (1-p)^{n-k} = B_n(p) \end{aligned}$$

[$B_n(p)$ — это известные полиномы Бернштейна.]

7. Используя чебышевский прием, оценить меру множества, на котором

$$\left| \frac{\varepsilon_1^{(p)}(t) + \dots + \varepsilon_n^{(p)}(t)}{n} - p \right| > \varepsilon,$$

и доказать, что

$$\lim_{n \rightarrow \infty} B_n(p) = f(p)$$

равномерно при $0 \leq p \leq 1$ [по определению $B_n(0) = f(0)$ и $B_n(1) = f(1)$]. (Это есть принадлежащее С. Бернштейну доказательство известной теоремы Вейерштрасса о приближении непрерывных функций полиномами.)

8. Пусть $f(t)$ удовлетворяет условию Липшица первого порядка, т. е.

$$|f(t_1) - f(t_2)| < M |t_1 - t_2|, \quad 0 \leq t_1, t_2 \leq 1,$$

где M — константа, не зависящая от t_1 и t_2 . Доказать, что

$$|f(p) - B_n(p)| \leq \frac{M}{2} \frac{1}{\sqrt{n}}.$$

9. Пусть

$$f(t) = \left| t - \frac{1}{2} \right|, \quad 0 \leq t \leq 1.$$

Эта функция удовлетворяет условию Липшица первого порядка. Использовать результат задачи 7 гл. 1 для оценки снизу

$$\left| f\left(\frac{1}{2}\right) - B_n\left(\frac{1}{2}\right) \right|$$

и таким образом показать, что порядок $1/\sqrt{n}$ оценки в задаче 8 наилучший из возможных.

10. Доказать, что для почти всех t

$$\lim_{n \rightarrow \infty} \frac{\varepsilon_1^{(p)}(t) + \dots + \varepsilon_n^{(p)}(t)}{n} = p.$$

11. Показать, что существует возрастающая функция $\varphi_p(t)$, такая, что

$$\varepsilon_k^{(p)}(t) = \varepsilon_k(\varphi_p(t)), \quad k=1, 2, \dots,$$

(величины ε_k — двоичные единицы). Показать далее, что при $p \neq \frac{1}{2}$ функция $\varphi_p(t)$ является «сингулярной», т. е. любое множество положительной меры E содержит подмножество E_1 , отличающееся от E на множество меры 0 и такое, что образ $\varphi_p(E_1)$ имеет меру 0. [См. Łomnicki Z., Ulam S., *Fund. Math.*, 23 (1934), 237—278, в частности 268—269.]

12. Доказать, что при любом $\varepsilon > 0$ ряд

$$\sum_{n=1}^{\infty} \frac{1}{n^{2+\varepsilon}} \exp \left\{ \frac{\sqrt{2 \log n}}{\sqrt{n}} |r_1(t) + \dots + r_n(t)| \right\}$$

сходится почти всюду и что, следовательно, неравенство

$$\limsup_{n \rightarrow \infty} \frac{|r_1(t) + \dots + r_n(t)|}{\sqrt{n \log n}} \leq \sqrt{2}$$

выполняется почти всюду.

Указание: следует заметить, что (ξ действительно)

$$\begin{aligned} \int_0^1 e^{\xi} |r_1(t) + \dots + r_n(t)| dt &< \int_0^1 e^{\xi} (r_1(t) + \dots + r_n(t)) dt + \\ &+ \int_0^1 e^{-\xi} (r_1(t) + \dots + r_n(t)) dt = 2(\operatorname{ch} \xi)^n. \end{aligned}$$

Замечание. Результат

$$\limsup_{n \rightarrow \infty} \frac{|r_1(t) + \dots + r_n(t)|}{\sqrt{n \log n}} \leq \sqrt{2}$$

был впервые получен Харди и Литлвудом в 1914 г. довольно сложным путем. Более сильный результат, утверждающий, что почти всюду

$$\limsup_{n \rightarrow \infty} \frac{|r_1(t) + \dots + r_n(t)|}{\sqrt{n \log \log n}} = \sqrt{2},$$

был доказан в 1922 г. Хинчиным. Это значительно труднее сделать.

3. «Герб или решетка» — более абстрактное изложение. Общепринятая структура статистических теорий (т. е. теорий, основанных на понятии вероятности) может быть кратко описана следующим образом.

Начинают с множества Ω («пространства выборов»), мера (вероятность) которого принимается равной 1. В Ω имеется совокупность подмножеств («элементарных множеств» или «элементарных событий»), меры (вероятности) которых заданы. Задача состоит в том, чтобы «продолжить» эту меру на возможно более широкую совокупность подмножеств Ω .

Правила продолжения меры следующие:

1. Если A_1, A_2, \dots — попарно непересекающиеся (попарно несовместимые) подмножества Ω (события) и если они измеримы (т. е. может быть определена их мера), то их объединение $\bigcup_{k=1}^{\infty} A_k$ также измеримо и

$$\mu \left\{ \bigcup_{k=1}^{\infty} A_k \right\} = \sum_{k=1}^{\infty} \mu \{ A_k \},$$

где $\mu \{ \}$ — это мера, приписанная множеству в скобках.

2. Если A измеримо, то также измеримо и его дополнение $\Omega - A$. (Из 1 и 2 следует, что $\mu\{\Omega - A\} = 1 - \mu(A)$, и, в частности, так как измеримость Ω постулирована, то мера пустого множества равна 0.)

3. Подмножество множества меры нуль измеримо.

Измеримые функции $f(\omega)$, $\omega \in \Omega$, определенные на Ω , называются «случайными величинами» (ужасная и вводящая в заблуждение терминология, к сожалению, непоправимо теперь закрепились). Посмотрим, как включается в эту схему «герб или решетка».

Пространство выборок Ω в этом случае — просто множество всех бесконечных последовательностей символов H и T , т. е. последовательностей, подобных

ω : $HTHTTTT\dots$

Что является элементарными событиями? Обычно это «цилиндрические множества», т. е. множества последовательностей, у каждой из которых на конечном числе выбранных мест стоят одни и те же символы. Например, множество всех последовательностей, у которых третий элемент H , седьмой — T и одиннадцатый — T , есть цилиндрическое множество. Какие значения меры должны быть приписаны этим множествам? Это зависит, конечно, от нематематических допущений относительно опытов с бросанием монеты. Эти допущения мы должны перевести на математический язык. Независимые бросания «симметричной монеты» переводятся на этот язык сопоставлением каждому цилиндрическому множеству меры

$$\left(\frac{1}{2}\right)^k,$$

где k — число фиксированных символов. Теперь возникает важная проблема, заключающаяся в доказательстве *единственности продолжения меры*. В нашем случае это может быть сделано очень просто посредством сведения к единственности меры Лебега. Последнее означает, что если мера μ , определенная на $(0,1)$, удовлетворяет 1, 2 и 3 и если μ -мера любого интервала равна его длине, то μ — обычная мера Лебега. Если мы будем писать 1 вместо H и 0 вместо T , то каждой последовательности символов H и T будет соответствовать (однозначно, исключая счетное множество двоично-рациональных чисел) число t , $0 \leq t \leq 1$, а именно число, двоичные цифры которого получаются заменой символов H и T данной последовательности единицей и нулем соответственно. Получившееся отображение имеет также то свойство, что оно переводит цилиндрические множества в объединения непересекающихся интервалов, концевые точки которых являются двоично-рациональными числами. Кроме того, мера, приписанная нами цилиндрическому множеству, равна лебеговой мере (длине) множества, полученного при его отображении. Теперь мы довели дело до конца!

Единственность продолжения может быть также доказана без использования отображения. Наиболее общая теорема такого рода была доказана Колмогоровым в его книге «Основные понятия теории вероятностей» в 1933 г.

Если мера на Ω введена, то можно стандартным путем построить теорию интегрирования, параллельную обычной теории Лебега.

Пусть $\omega \in \Omega$, т. е. ω — последовательность H и T .
Положим

$$X_k(\omega) = \begin{cases} +1, & \text{если } k\text{-й элемент } \omega = H, \\ -1, & \text{если } k\text{-й элемент } \omega = T. \end{cases}$$

Функции $X_k(\omega)$ — «независимые случайные величины» в том смысле, что для любой последовательности δ_j , где каждая из величин δ равна или 1, или -1 ,

$$\begin{aligned} \mu \{X_1(\omega) = \delta_1, X_2(\omega) = \delta_2, \dots, X_n(\omega) = \delta_n\} = \\ = \frac{1}{2^n} = \prod_{k=1}^n \mu \{X_k(\omega) = \delta_k\}. \end{aligned} \quad (3.1)$$

Ясно, что $X_k(\omega)$ дают нам модель независимых бросаний симметричной монеты.

4. В чем ценность абстракции? Абстрагировать — это, по-видимому, значит переходить к сути дела. Это значит освободиться от случайных черт и сосредотачивать внимание на особо важных свойствах. Абстрактно теория игры «герб или решетка» (симметричная монета, независимые бросания) сводится просто к изучению функций $X_k(\omega)$ со свойством (3.1), определенных на некотором пространстве Ω (с единичной мерой), где задана мера μ , удовлетворяющая требованиям 1, 2 и 3 предыдущего пункта. Не важно, чем является Ω — разрешено лишь пользоваться (3.1)

и элементарными свойствами 1, 2 и 3 меры. Следует, конечно, убедиться, что мы имеем дело не с математическим вакуумом, т. е. объекты, о которых мы говорим, могут быть определены. Это достигается тем, что в качестве Ω берут «пространство выборок» и требуемую меру μ строят, как это указывалось в п. 3. То обстоятельство, что *реализация* $X_h(\omega)$ возможна посредством функций Радемахера $r_h(t)$ (т. е. что в качестве Ω можно выбрать интервал $(0,1)$ с обычной мерой Лебега на нем), можно рассматривать как несущественное. Отметим, что, за исключением занимательного доказательства формулы Виета, где мы использовали весьма специальное свойство функций Радемахера, именно что

$$1 - 2t = \sum_{k=1}^{\infty} \frac{r_k(t)}{2^k},$$

мы никогда не обращались к чему-либо, кроме как к свойству (3.1) и *общим* свойствам меры. Однако цена, которой придется расплатиться за неограниченную абстракцию, в действительности значительно больше. Неограниченному абстрагированию свойственна тенденция отвлекать внимание от целых областей применения, самое открытие которых зависит от тех черт, какие абстрактная точка зрения исключает как случайные. Иллюстрации этого разбросаны по всей книге. Начнем с нескольких примеров из области, уже хорошо нам знакомой.

5. Пример 1. Сходимость ряда со случайными знаками. Какова вероятность того, что ряд

$$\sum_{k=1}^{\infty} \pm c_k \quad (c_k \text{ действительны})$$

со знаками, выбранными независимо и с вероятностью $\frac{1}{2}$ каждый, сходится? Эта задача была впервые поставлена в такой форме Г. Штейнгаузом в 1922 г. (и независимо от него Н. Винером), которому мы обязаны также постановкой проблем, изложенных в п. 3. Штейнгауз отметил, что задача эквивалентна нахождению меры множества точек t , в которых сходится ряд

$$\sum_1^{\infty} c_k r_k(t). \quad (5.1)$$

Этот вопрос был тогда уже разрешен Радемахером, который доказал, что если

$$\sum_1^{\infty} c_k^2 < \infty, \quad (5.2)$$

то ряд (5.1) сходится почти всюду. Мы могли, конечно, рассматривать вопрос сходимости ряда

$$\sum_{k=1}^{\infty} c_k X_k(\omega), \quad (5.3)$$

где $X_k(\omega)$ обладают свойством (3.1). Действительно, доказательство Колмогорова, который максимально

обобщил теорему Радемахера, использовало лишь (3.1). Существует, однако, замечательное доказательство, принадлежащее Палею и Зигмунду, которое существенно опирается на свойства функций Радемахера. Это доказательство мы воспроизведем здесь по причинам, которые станут понятными несколько позже. Оно основано на двух не совсем элементарных и очень важных теоремах.

1. Теорема Рисса — Фишера, которая утверждает, что если

$$\sum a_k^2 < \infty$$

и если $\varphi_1(t), \varphi_2(t), \dots$ — ортонормальная система, определенная на множестве E , т. е.

$$\int_E \varphi_i(t) \varphi_j(t) dt = \delta_{ij}, \quad (5.4)$$

то существует функция $f(t) \in L^2$ (т. е. $\int_E f^2(t) dt < \infty$),

такая, что

$$\lim_{n \rightarrow \infty} \int_E \left(f(t) - \sum_{k=1}^n a_k \varphi_k(t) \right)^2 dt = 0. \quad (5.5)$$

2. Фундаментальная теорема анализа, которая в своем «высшем» варианте устанавливает, что если

$$\int_0^1 |f(t)| dt < \infty, \quad (5.6)$$

то для почти каждого t_0

$$\lim_{m \rightarrow \infty} \frac{1}{\beta_m - \alpha_m} \int_{\alpha_m}^{\beta_m} f(t) dt = f(t_0) \quad (5.7)$$

при условии

$$\alpha_m < t_0 < \beta_m \text{ и } \lim_{m \rightarrow \infty} \alpha_m = \lim_{m \rightarrow \infty} \beta_m = t_0.$$

Далее, мы знаем, что функции Радемахера ортонормальны на интервале $(0,1)$:

$$\int_0^1 r_i(t) r_j(t) dt = \delta_{ij}.$$

Следовательно (по теореме Рисса — Фишера, сформулированной выше), существует функция $f(t)$, такая, что

$$\int_0^1 f^2(t) dt < \infty \quad (5.8)$$

и

$$\lim_{n \rightarrow \infty} \int_0^1 \left(f(t) - \sum_{k=1}^n c_k r_k(t) \right)^2 dt = 0. \quad (5.9)$$

(Вспомним наше допущение $\sum_1^{\infty} c_k^2 < \infty$.)

Пусть теперь t_0 таково, что выполняется (5.7) [из (5.8)

вытекает (5.6)!], и пусть

$$\alpha_m = \frac{k_m}{2^m} < t_0 < \frac{k_m + 1}{2^m} = \beta_m \quad (5.10)$$

(мы исключаем возможность, когда t_0 является двоично-рациональным). Мы имеем

$$\begin{aligned} \left| \int_{\alpha_m}^{\beta_m} \left(f(t) - \sum_1^n c_k r_k(t) \right) dt \right| &\leq \\ &\leq (\beta_m - \alpha_m)^{1/2} \left(\int_0^1 \left(f(t) - \sum_1^n c_k r_k(t) \right)^2 dt \right)^{1/2} \end{aligned}$$

и, следовательно, по (5.9)

$$\int_{\alpha_m}^{\beta_m} f(t) dt = \sum_1^{\infty} c_k \int_{\alpha_m}^{\beta_m} r_k(t) dt. \quad (5.11)$$

Заметим, что

$$\int_{\alpha_m}^{\beta_m} r_k(t) dt = 0, \quad k > m, \quad (5.12)$$

и

$$\int_{\alpha_m}^{\beta_m} r_k(t) dt = (\beta_m - \alpha_m) r_k(t_0), \quad k \leq m. \quad (5.13)$$

Таким образом, (5.11) превращается в

$$\frac{1}{\beta_m - \alpha_m} \int_{\alpha_m}^{\beta_m} f(t) dt = \sum_1^m c_k r_k(t_0),$$

и, следовательно, по (5.7) ряд

$$\sum_1^{\infty} c_k r_k(t_0)$$

сходится.

Вышеприведенные аргументы могут быть непосредственно применены к доказательству того, что ряд

$$\sum_{k=1}^{\infty} c_k \sin 2\pi 2^k t \quad (5.14)$$

сходится почти всюду, если

$$\sum c_k^2 < \infty. \quad (5.15)$$

Эта теорема естественным образом подсказывается, если заметим, что

$$r_k(t) = \operatorname{sgn} \sin 2\pi 2^{k-1} t.$$

Действительно, наше доказательство основывается на трех свойствах функций Радемахера:

- 1) ортонормальность,
- 2) (5.12),
- 3) (5.13).

При замене $r_k(t)$ на $\sin 2\pi 2^k t$ свойства 1 и 2 сохраняются. Последнее же не сохраняется в точности,

однако при $k \leq m$ мы имеем

$$\int_{\alpha_m}^{\beta_m} \sin 2\pi 2^k t \, dt = (\beta_m - \alpha_m) \sin 2\pi 2^k t_0 +$$

$$+ \int_{\alpha_m}^{\beta_m} (\sin 2\pi 2^k t - \sin 2\pi 2^k t_0) \, dt \quad (5.16)$$

и

$$\left| \int_{\alpha_m}^{\beta_m} (\sin 2\pi 2^k t - \sin 2\pi 2^k t_0) \, dt \right| \leq$$

$$\leq \int_{\alpha_m}^{\beta_m} |\sin 2\pi 2^k t - \sin 2\pi 2^k t_0| \, dt \leq 2\pi 2^k \int_{\alpha_m}^{\beta_m} |t - t_0| \, dt <$$

$$< 2\pi 2^k (\beta_m - \alpha_m)^2 = 2\pi \frac{2^k}{2^m} (\beta_m - \alpha_m).$$

Теперь вместо

$$\frac{1}{\beta_m - \alpha_m} \int_{\alpha_m}^{\beta_m} f(t) \, dt = \sum_1^m c_k r_k(t_0)$$

мы получим

$$\left| \frac{1}{\beta_m - \alpha_m} \int_{\alpha_m}^{\beta_m} f(t) \, dt - \sum_1^m c_k \sin 2\pi 2^k t_0 \right| \leq \sum_1^m |c_k| \frac{2\pi}{2^{m-k}},$$

а так как $c_n \rightarrow 0$ при $n \rightarrow \infty$ (вспомните, что $\sum c_n^2 < \infty$), то

$$\lim_{m \rightarrow \infty} \sum_1^m |c_k| \frac{2\pi}{2^{m-k}} = 0,$$

и этого достаточно для завершения доказательства.

Только что доказанная теорема о сходимости ряда

$$\sum_1^{\infty} c_k \sin 2\pi 2^k t$$

является фактически частным случаем известной теоремы Колмогорова о том, что сходимость

$$\sum_1^{\infty} c_k^2 < \infty$$

влечет за собой сходимость почти всюду ряда

$$\sum_{k=1}^{\infty} c_k \sin 2\pi n_k t$$

при условии существования числа q , для которого

$$\frac{n_{k+1}}{n_k} > q > 1.$$

Колмогоров при доказательстве использовал существенным образом то, что рассматриваемый ряд является тригонометрическим. Однако, обобщив аргументацию Палая — Зигмунда, можно доказать следующую значительно более общую теорему.

Если функция $g(t)$ — периодическая с периодом 1 и если

$$(a) \quad \int_0^1 g(t) dt = 0,$$

$$(b) \quad |g(t') - g(t'')| < M |t' - t''|^\alpha, \quad 0 < \alpha < 1,$$

то сходимость $\sum c_k^2$ влечет за собой сходимость почти всюду ряда

$$\sum_1^\infty c_k g(n_k t),$$

где целые числа n_k таковы, что

$$\frac{n_{k+1}}{n_k} > q > 1.$$

Доказательство этого утверждения из-за нескольких технических трудностей не будет здесь воспроизведено, хотя по существу никакой новой идеи, сверх принадлежащей Палею и Зигмунду, не требуется.

Какой вывод можно сделать из разобранного примера? На вид случайный факт, заключающийся в том, что

$$r_k(t) = \operatorname{sgn} \sin 2\pi 2^{k-1} t,$$

наводит на мысль, что здесь могут быть проведены аналогии между $r_k(t)$ и $\sin 2\pi 2^{k-1} t$. Так как $r_k(t)$ имеют определенную вероятностную интерпретацию, то этим открывается путь к установлению связи между игрой «гербы или решетки» и математической областью,

где нет понятий случая и вероятности, монет и чего бы то ни было подобного. Могло ли это быть достигнуто, если бы мы настояли на абстрактной трактовке игры «герб или решетка»? Может быть, но я сомневаюсь.

6. Пример 2. Расходимость ряда со случайными знаками. Что происходит с рядом

$$\sum_{k=1}^{\infty} \pm c_k, \quad (6.1)$$

если

$$\sum_1^{\infty} c_k^2 = \infty? \quad (6.2)$$

На это можно ответить, что ряд (6.1) расходится с вероятностью 1. Доказательство крайне просто. Заметим, во-первых, что наша задача состоит собственно в определении меры множества сходимости ряда

$$\sum_1^{\infty} c_k r_k(t) \quad (6.3)$$

при условии (6.2). Во-вторых, множество сходимости ряда (6.3) должно иметь меру или 0, или 1 (частный случай так называемого закона нуля или единицы). Вспомним, что

$$r_k(t) = r_1(2^{k-1}t)^1,$$

¹⁾ Здесь подразумевается, что $r_k(t)$ определены таким образом, чтобы они были периодичны с периодом 1. Другими словами, $r_k(t+1) = r_k(t)$.

и, следовательно, вместе с t множеству сходимости принадлежит и

$$t + \frac{1}{2^l}$$

при $l = 0, 1, 2, \dots$.

Действительно, если t заменить на $t + 2^{-l}$, то в (6.3) изменится лишь *конечное* число членов, а это не может повлиять на сходимость. Таким образом, характеристическая функция множества сходимости имеет произвольно малые периоды, и по известной теореме она должна быть постоянна почти всюду, причем эта константа равна или 0, или 1¹⁾.

¹⁾ Для ограниченной измеримой (следовательно, интегрируемой по Лебегу!) функции $\varphi(t)$ доказательство следующее: имеем

$$I = \int_0^1 \varphi(t) dt = \sum_{k=0}^{2^l-1} \int_{k/2^l}^{(k+1)/2^l} \varphi(t) dt = 2^l \int_{k/2^l}^{(k+1)/2^l} \varphi(t) dt.$$

Пусть t_0 таково, что при $k_l/2^l < t_0 < (k_l+1)/2^l$

$$\lim_{l \rightarrow \infty} 2^l \int_{k_l/2^l}^{(k_l+1)/2^l} \varphi(t) dt = \varphi(t_0).$$

По фундаментальной теореме анализа (см. п. 5) почти каждое t_0 имеет такое свойство. Таким образом, $\varphi(t_0) = I$ для почти каждого t_0 . Если $\varphi(t)$ не предполагается ограниченной, то те же аргументы применимы к $e^{i\varphi(t)}$. Это доказательство предложено Гартманом и Кершнером; теорема была доказана сначала более сложным способом Бурстиным. Характеристическая функция множества сходимости ряда из измеримых функций измерима, так как измеримо само это множество.

Мы можем предположить, что $c_n \rightarrow 0$, в противном случае утверждение нашей теоремы будет тривиальным.

Допустим теперь, что (6.2) выполняется, $c_n \rightarrow 0$ и ряд (6.3) сходится на множестве положительной меры. По вышеприведенному замечанию он должен сходиться почти всюду. Следовательно, существует измеримая функция $g(t)$, такая, что почти всюду

$$\lim_{n \rightarrow \infty} \sum_1^n c_k r_k(t) = g(t). \quad (6.4)$$

Теперь в силу (6.4) для каждого действительного $\xi \neq 0$ почти всюду

$$\lim_{n \rightarrow \infty} \exp \left[i\xi \sum_1^n c_k r_k(t) \right] = e^{i\xi g(t)}.$$

По теореме Лебега об ограниченной сходимости получаем

$$\lim_{n \rightarrow \infty} \int_0^1 \exp \left[i\xi \sum_1^n c_k r_k(t) \right] dt = \int_0^1 e^{i\xi g(t)} dt. \quad (6.5)$$

Однако известно, что

$$\lim_{n \rightarrow \infty} \int_0^1 \exp \left[i\xi \sum_1^n c_k r_k(t) \right] dt = \prod_{k=1}^{\infty} \cos \xi c_k. \quad (6.6)$$

Читателю предстоит самому вывести из (6.2) и условия $c_n \rightarrow 0$, что

$$\lim_{n \rightarrow \infty} \prod_{k=1}^n \cos \xi c_k = 0.$$

Тогда для каждого действительного $\xi \neq 0$

$$\int_0^1 e^{i\xi g(t)} dt = 0. \quad (6.7)$$

Выберем теперь последовательность $\xi_n \rightarrow 0$, но так, чтобы каждое $\xi_n \neq 0$ (например, $\xi_n = n^{-1}$), тогда для почти каждого t

$$\lim_{n \rightarrow \infty} \xi_n g(t) = 0$$

и, следовательно,

$$\lim_{n \rightarrow \infty} e^{i\xi_n g(t)} = 1$$

для почти каждого t .

Снова по теореме Лебега о предельном переходе под знаком интеграла

$$\lim_{n \rightarrow \infty} \int_0^1 e^{i\xi_n g(t)} dt = 1,$$

что ввиду (6.7) приводит к противоречию: $0 = 1$. Поэтому ряд (6.3) не может сходиться на множестве положительной меры. Следовательно, ряд должен расходиться почти всюду.

Этот метод доказательства существенным образом использует независимость $r_k(t)$ [см. (6.6)] и, по-видимому, непосредственно неприменим к изучению ряда

$$\sum_{h=1}^{\infty} c_h \sin 2\pi n_k t, \quad \frac{n_{k+1}}{n_k} > q > 1,$$

при условии

$$\sum_1^{\infty} c_k^2 = \infty.$$

Однако данный метод все же можно приспособить для этой цели, но мы отложим обсуждение этого вопроса до более позднего момента.

ЗАДАЧИ

1 Пусть $\sum_1^{\infty} c_k^2 = \infty$, $c_k \rightarrow 0$; рассмотрим ряд

$$\sum_{k=1}^{\infty} c_k \sin 2\pi 2^{k-1} t.$$

(а) Доказать, что

$$\lim_{n \rightarrow \infty} \int_0^1 \left(\frac{\sum_1^n c_k \sin 2\pi 2^{k-1} t}{\sqrt{\sum_1^n c_k^2}} \right)^4 dt$$

существует, и найти его значение.

(б) Пусть $F_n(t)$, $0 \leq t \leq 1$, последовательность функций, такая, что

$$\lim_{n \rightarrow \infty} \int_0^1 F_n^2(t) dt = \alpha, \quad \lim_{n \rightarrow \infty} \int_0^1 F_n^4(t) dt = \beta$$

Доказать тогда, что мера множества E , на котором $F_n(t)$ стремится к 0, не может превосходить

$$1 - \frac{\alpha^2}{\beta}$$

(в) Используя (а) и (б), доказать, что при условиях данной задачи ряд

$$\sum_{k=1}^{\infty} c_k \sin 2\pi 2^{k-1} t$$

расходится почти всюду.

2. Следующий пример показывает, что в теореме задачи 1 синус не может быть заменен «произвольной» периодической функцией $f(t)$ периода 1 (подчиненной, конечно, условию

$$\int_0^1 f(t) dt = 0$$

Пусть

$$f(t) = \sin 2\pi t - \sin 4\pi t;$$

показать, что

$$\sum_{k=1}^{\infty} \frac{1}{\sqrt{k}} f(2^{k-1}t)$$

сходится всюду.

ЛИТЕРАТУРА

Borel E., Les probabilités dénombrables et leurs applications arithmétiques, *Rend. Circ. Math. Palermo*, 27 (1909), 247 — 271.

- Champernowne D. G., The construction of decimals normal in the scale of ten, *J. London Math. Soc.*, 8 (1933), 254—260.
- Steinhaus H., Les probabilités dénombrables et leur rapport à la théorie de la mesure, *Fund. Math.*, 4 (1922), 286—310.
- Rademacher H., Einige Sätze über Reihen von allgemeinen Orthogonalfunktionen, *Math. Ann.*, 87 (1922), 112—138.
- Кас М., Convergence of certain gap series, *Ann. Math.*, 44 (1943). 411—415 (здесь дается ссылка на первоначальные работы Палея и Зигмунда).
- Hartman Ph. and Kirshner R., The structure of monotone functions, *Amer. J. Math.*, 59 (1937), 809—822.

НОРМАЛЬНЫЙ ЗАКОН

1. Муавр. В п. 1 гл. 2 мы обсуждали «слабый закон больших чисел». Муавром был доказан более точный результат, а именно что

$$\begin{aligned} \lim_{n \rightarrow \infty} \mu \{ \omega_1 \sqrt{n} < r_1(t) + \dots + r_n(t) < \omega_2 \sqrt{n} \} = \\ = \frac{1}{\sqrt{2\pi}} \int_{\omega_1}^{\omega_2} e^{-y^2/2} dy. \end{aligned} \quad (1.1)$$

Читатель не встретит затруднений при интерпретации этого результата в вероятностных терминах. Элементарное доказательство может быть основано на формуле (5.6) гл. 1. При этом (1.1) становится эквивалентным чисто комбинаторной формуле

$$\lim_{n \rightarrow \infty} \sum_{\frac{n}{2} + \omega_1 \sqrt{n} < l < \frac{n}{2} + \omega_2 \sqrt{n}} \frac{1}{2^n} \binom{n}{l} = \frac{1}{\sqrt{2\pi}} \int_{\omega_1}^{\omega_2} e^{-y^2/2} dy. \quad (1.2)$$

Прямое использование формулы Стирлинга даст (1.2), но в то же время это доказательство затемнит характер теоремы. Попытки обобщить (1.1) дали сильный толчок развитию аналитических методов теории вероят-

ностей. Мощный метод был предложен Марковым, но он не сумел сделать его строгим. Около двадцати лет спустя метод был обоснован П. Леви. Марковскому методу посвящены следующие два пункта.

2. Основная идея метода. Пусть

$$g(x) = \begin{cases} 1, & \omega_1 < x < \omega_2, \\ 0, & \text{в противном случае.} \end{cases} \quad (2.1)$$

Из элементарной теории интеграла Фурье известно, что

$$g(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{e^{i\omega_2\xi} - e^{i\omega_1\xi}}{i\xi} e^{-ix\xi} d\xi \quad (2.2)$$

при обычном условии, что для $x = \omega_1$ и $x = \omega_2$ получается значение, равное $\frac{1}{2}$. Далее, если только ω_1 и ω_2 не являются целыми кратными числа \sqrt{n} , то имеем

$$\begin{aligned} \mu \left\{ \omega_1 < \frac{r_1(t) + \dots + r_n(t)}{\sqrt{n}} < \omega_2 \right\} &= \\ &= \int_0^1 g \left(\frac{r_1(t) + \dots + r_n(t)}{\sqrt{n}} \right) dt = \\ &= \int_0^1 \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{e^{i\omega_2\xi} - e^{i\omega_1\xi}}{i\xi} \exp \left(-i\xi \frac{r_1(t) + \dots + r_n(t)}{\sqrt{n}} \right) d\xi dt. \end{aligned} \quad (2.3)$$

Изменяя порядок интегрирования в последнем выражении [в данном случае это легко обосновывается, так как сумма $r_1(t) + \dots + r_n(t)$ принимает только конечное число значений], мы получим

$$\begin{aligned} \mu \left\{ \omega_1 < \frac{r_1(t) + \dots + r_n(t)}{\sqrt{n}} < \omega_2 \right\} &= \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{i\omega_2 \xi} - e^{i\omega_1 \xi}}{i\xi} \left[\int_0^1 \exp \left(-i\xi \frac{r_1(t) + \dots + r_n(t)}{\sqrt{n}} \right) dt \right] d\xi = \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{i\omega_2 \xi} - e^{i\omega_1 \xi}}{i\xi} \left(\cos \frac{\xi}{\sqrt{n}} \right)^n d\xi. \end{aligned} \quad (2.4)$$

Для каждого действительного ξ

$$\lim_{n \rightarrow \infty} \left(\cos \frac{\xi}{\sqrt{n}} \right)^n = e^{-\xi^2/2}, \quad (2.5)$$

и напрашивается вывод, что

$$\begin{aligned} \lim_{n \rightarrow \infty} \mu \left\{ \omega_1 < \frac{r_1(t) + \dots + r_n(t)}{\sqrt{n}} < \omega_2 \right\} &= \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{i\omega_2 \xi} - e^{i\omega_1 \xi}}{i\xi} e^{-\xi^2/2} d\xi = \frac{1}{\sqrt{2\pi}} \int_{\omega_1}^{\omega_2} e^{-y^2/2} dy. \end{aligned} \quad (2.6)$$

Каковы затруднения с этим методом? Единственный шаг, требующий обоснования, — это переход к преде-

лу под знаком интеграла при $n \rightarrow \infty$. К сожалению, пределы интегрирования суть $-\infty$ и $+\infty$, а функция

$$\frac{e^{i\omega_2\xi} - e^{i\omega_1\xi}}{i\xi}$$

не является абсолютно интегрируемой.

Марков, который был великолепным математиком, не смог преодолеть эту трудность и отказался от этого метода!

Физики, понятие строгости у которых менее точно, чем наше, так и называют метод «марковским методом», тогда как математики едва ли знают о его происхождении.

3. Метод Маркова становится строгим. Обоснование марковского метода в действительности совсем просто. Оно основано на простой и широко применимой идее.

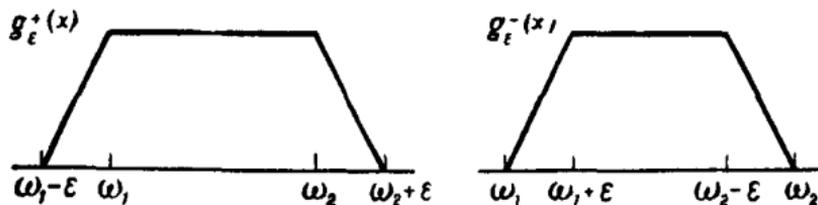
Исследуем сначала формулу (2.2). Это обычная формула Фурье:

$$g(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(y) e^{i\xi(y-x)} dy d\xi, \quad (3.1)$$

примененная к специальной функции (2.1).

Введем теперь две вспомогательные функции: $g_\varepsilon^+(x)$ и $g_\varepsilon^-(x)$, графики¹⁾ которых приведены ниже ($\varepsilon > 0$, $2\varepsilon < \omega_2 - \omega_1$).

1) Высота обоих графиков равна 1.



Мы имеем

$$g_{\varepsilon}^{-}(x) \leq g(x) \leq g_{\varepsilon}^{+}(x) \quad (3.2)$$

и, следовательно,

$$\begin{aligned} \int_0^1 g_{\varepsilon}^{-} \left(\frac{r_1(t) + \dots + r_n(t)}{\sqrt{n}} \right) dt &\leq \\ &\leq \mu \left\{ \omega_1 < \frac{r_1(t) + \dots + r_n(t)}{\sqrt{n}} < \omega_2 \right\} \leq \\ &< \int_0^1 g_{\varepsilon}^{+} \left(\frac{r_1(t) + \dots + r_n(t)}{\sqrt{n}} \right) dt \end{aligned} \quad (3.3)$$

Теперь функции

$$G_{\varepsilon}^{-}(\xi) = \int_{-\infty}^{\infty} g_{\varepsilon}^{-}(y) e^{iy\xi} dy \quad \text{и} \quad G_{\varepsilon}^{+}(\xi) = \int_{-\infty}^{\infty} g_{\varepsilon}^{+}(y) e^{iy\xi} dy$$

абсолютно интегрируемы по ξ в интервале $(-\infty, \infty)$ и вследствие этого аргументация п 2 приводит к строгому выводу

$$\begin{aligned} \lim_{n \rightarrow \infty} \int_0^1 g_{\varepsilon}^{-} \left(\frac{r_1(t) + \dots + r_n(t)}{\sqrt{n}} \right) dt &= \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\xi^2/2} \int_{-\infty}^{\infty} g_{\varepsilon}^{-}(y) e^{i\xi y} dy d\xi = \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g_{\varepsilon}^{-}(y) e^{-y^2/2} dy \end{aligned} \quad (3.4)$$

и

$$\begin{aligned} \lim_{n \rightarrow \infty} \int_0^1 g_{\varepsilon}^{+} \left(\frac{r_1(t) + \dots + r_n(t)}{\sqrt{n}} \right) dt &= \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\xi^2/2} \int_{-\infty}^{\infty} g_{\varepsilon}^{+}(y) e^{i\xi y} dy d\xi = \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g_{\varepsilon}^{+}(y) e^{-y^2/2} dy. \end{aligned} \quad (3.5)$$

Комбинируя (3.4) и (3.5) с (3.3), получаем

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g_{\varepsilon}^{-}(y) e^{-y^2/2} dy \leq$$

$$\begin{aligned}
&\leq \liminf_{n \rightarrow \infty} \mu \left\{ \omega_1 < \frac{r_1(t) + \dots + r_n(t)}{\sqrt{n}} < \omega_2 \right\} \leq \\
&\leq \limsup_{n \rightarrow \infty} \mu \left\{ \omega_1 < \frac{r_1(t) + \dots + r_n(t)}{\sqrt{n}} < \omega_2 \right\} \leq \\
&\leq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g_{\varepsilon}^+(y) e^{-y^2/2} dy. \tag{3.6}
\end{aligned}$$

Так как (3.6) справедливо при *любом* $\varepsilon > 0$, то мы сразу получим

$$\begin{aligned}
\lim_{n \rightarrow \infty} \mu \left\{ \omega_1 < \frac{r_1(t) + \dots + r_n(t)}{\sqrt{n}} < \omega_2 \right\} = \\
= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(y) e^{-y^2/2} dy = \frac{1}{\sqrt{2\pi}} \int_{\omega_1}^{\omega_2} e^{-y^2/2} dy \tag{3.7}
\end{aligned}$$

ЗАДАЧИ

1. В 1917 г. Г. Вейль доказал, что при *любом* иррациональном α последовательность $a_n = n\alpha - [n\alpha]$, $n = 1, 2, \dots$, равномерно распределена на $(0, 1)$. Другими словами, если $0 \leq \omega_1 < \omega_2 \leq 1$ и $k_n(\omega_1, \omega_2)$ обозначает число величин a_j , $1 \leq j \leq n$, которые попали в интервал (ω_1, ω_2) , то

$$\lim_{n \rightarrow \infty} \frac{k_n(\omega_1, \omega_2)}{n} = \omega_2 - \omega_1.$$

Вводя периодическую функцию $g(x)$ с периодом 1, заданную на $(0, 1)$ соотношением (2.1), и используя ряд Фурье вместо интеграла Фурье, доказать теорему Вейля.

2. Использовать метод Маркова для доказательства формулы Лапласа

$$\lim_{x \rightarrow \infty} e^{-x} \sum_{x+\omega_1 \sqrt{x} < k < x+\omega_2 \sqrt{x}} \frac{x^k}{k!} = \frac{1}{\sqrt{2\pi}} \int_{\omega_1}^{\omega_2} e^{-y^2/2} dy$$

4. Более внимательный взгляд на метод. Анализ материала п. 3 обнаруживает, что в действительности доказана следующая теорема.

Пусть $f_n(t)$, $0 \leq t \leq 1$, — последовательность измеримых функций, такая, что для каждого действительного ξ

$$\lim_{n \rightarrow \infty} \int_0^1 e^{i\xi f_n(t)} dt = e^{-\xi^2/2}. \quad (4.1)$$

Тогда

$$\lim_{n \rightarrow \infty} \mu \{ \omega_1 \leq f_n(t) < \omega_2 \} = \frac{1}{\sqrt{2\pi}} \int_{\omega_1}^{\omega_2} e^{-y^2/2} dy. \quad (4.2)$$

Пусть

$$\sigma_n(\omega) = \mu \{ f_n(t) < \omega \}, \quad (4.3)$$

тогда $\sigma_n(\omega)$ обладает следующими свойствами:

- 1) $\sigma_n(-\infty) = 0$, $\sigma_n(+\infty) = 1$;
- 2) $\sigma_n(\omega)$ — неубывающая;
- 3) $\sigma_n(\omega)$ непрерывна слева

(Отметим, что свойство 3 есть следствие полной аддитивности меры Лебега.) Функция $\sigma(\omega)$, имеющая

свойства 1, 2 и 3, называется *функцией распределения*. Тогда

$$\int_0^1 e^{i\xi f_n(t)} dt = \int_{-\infty}^{+\infty} e^{i\xi\omega} d\sigma_n(\omega) \quad (4.4)$$

и наша теорема может быть сформулирована так.

Если последовательность функций распределения $\sigma_n(\omega)$ такова, что для каждого действительного ξ

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{+\infty} e^{i\xi\omega} d\sigma_n(\omega) = e^{-\xi^2/2}, \quad (4.5)$$

то

$$\sigma_n(\omega_2) - \sigma_n(\omega_1) \rightarrow G(\omega_2) - G(\omega_1), \quad (4.6)$$

где

$$G(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\omega} e^{-y^2/2} dy. \quad (4.7)$$

Внимательный читатель заметит небольшой логический пробел. Если нам задана какая-либо последовательность функций распределения $\sigma_n(\omega)$, то последнее утверждение следует из предыдущего только тогда, когда мы можем подобрать последовательность функций $f_n(t)$, $0 \leq t \leq 1$, таких, что

$$\mu \{f_n(t) < \omega\} = \sigma_n(\omega). \quad (4.8)$$

В сущности, это затруднение можно обойти, пользуясь аргументами п. 3. Однако конструкция функций $f_n(t)$ чрезвычайно проста. Действительно, мы можем взять

в качестве $f_n(t)$ функции, обратные к функциям $\sigma_n(\omega)$, имея в виду, что интервалы постоянства $\sigma_n(\omega)$ переходят в разрывы $f_n(t)$ и разрывы $\sigma_n(\omega)$ — в интервалы постоянства $f_n(t)$. Мы оставляем детали читателю. Заключение, что (4.5) влечет за собой (4.6), является специальным случаем важной общей теоремы, известной под названием теоремы непрерывности для преобразований Фурье—Стилтьеса. Эта теорема может быть сформулирована следующим образом.

Если $\sigma_n(\omega)$ — последовательность функций распределения, такая, что для каждого действительного ξ

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{+\infty} e^{i\xi\omega} d\sigma_n(\omega) = c(\xi) \quad (4.9)$$

и $c(\xi)$ непрерывна в точке $\xi = 0$, то существует единственная функция распределения $\sigma(\omega)$, такая, что

$$\int_{-\infty}^{+\infty} e^{i\xi\omega} d\sigma(\omega) = c(\xi) \quad (4.10)$$

и

$$\lim_{n \rightarrow \infty} \sigma_n(\omega) = \sigma(\omega) \quad (4.11)$$

при каждом ω , для которого $\sigma(\omega)$ непрерывна.

Доказательство в дополнение к уже изложенным идеям использует так называемый принцип выбора Хелли и ввиду некоторых технических трудностей не может быть здесь представлено. Поэтому мы пропустим его, но в то же время будем свободно использовать теорему впоследствии.

ЗАДАЧИ

1. Пусть функции $f_n(t)$, $0 \leq t \leq 1$, таковы, что при $k=0, 1, 2, \dots$ мы имеем

$$\lim_{n \rightarrow \infty} \int_0^1 f_n^k(t) dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} y^k e^{-y^2/2} dy = \begin{cases} 0, & k \text{ нечетно,} \\ \frac{k!}{2^{k/2} \left(\frac{k}{2}\right)!}, & k \text{ четно.} \end{cases}$$

Доказать, что для каждого действительного ξ .

$$\lim_{n \rightarrow \infty} \int_0^1 e^{i\xi f_n(t)} dt = e^{-\xi^2/2}$$

и, следовательно, что выполняется (4.2).

2 Пусть $\{n_m\}$ — последовательность целых чисел, такая, что

$$\lim_{m \rightarrow \infty} \frac{n_{m+1}}{n_m} = \infty.$$

Доказать, что при $k=0, 1, 2, \dots$

$$\begin{aligned} \lim_{n \rightarrow \infty} \int_0^1 \left(\frac{\sqrt{2} \cos 2\pi n_1 t + \cos 2\pi n_2 t + \dots + \cos 2\pi n_m t}{\sqrt{m}} \right) dt &= \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^k e^{-y^2/2} dy \end{aligned}$$

и, следовательно,

$$\begin{aligned} \lim_{m \rightarrow \infty} \mu \left\{ \omega_1 < \frac{\sqrt{2} \cos 2\pi n_1 t + \cos 2\pi n_2 t + \dots + \cos 2\pi n_m t}{\sqrt{m}} < \omega_2 \right\} &= \\ &= \frac{1}{\sqrt{2\pi}} \int_{\omega_1}^{\omega_2} e^{-y^2/2} dy \end{aligned}$$

Замечание: тем же самым способом, но используя более хитрые комбинаторные соображения, можно доказать, что если

$$\sum_{k=1}^{\infty} c_k^2 = \infty, \quad |c_k| < M,$$

и

$$\frac{n_{k+1}}{n_k} > q > 1,$$

то

$$\lim_{n \rightarrow \infty} \mu \left\{ \omega_1 < \sqrt{2} \frac{\sum_1^n c_k \cos 2\pi n_k t}{\sqrt{\sum_1^n c_k^2}} < \omega_2 \right\} = \frac{1}{\sqrt{2\pi}} \int_{\omega_1}^{\omega_2} e^{-y^2/2} dy.$$

В частности, отсюда следует, что равенство $\sum_1^{\infty} c_k^2 = \infty$ влечет за собой расходимость почти всюду $\sum_1^{\infty} c_k \cos 2\pi n_k t$ (аргументы, конечно, применимы и в том случае, когда косинус заменяется на синус).

Как читатель может видеть, это близко связано с методом, использованным в примере 2 п. 5 гл. 2.

3 Пусть $\sigma(\omega)$ — функция распределения, и пусть

$$c(\xi) = \int_{-\infty}^{\infty} e^{i\xi\omega} d\sigma(\omega).$$

Доказать, что

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T |c(\xi)|^2 d\xi \text{ равен сумме квадратов скачков функции } \sigma(\omega)$$

(Эта простая, но замечательная теорема принадлежит Н. Винеру.)

Доказательство может основываться на том, что

$$c(\xi) = \int_0^1 e^{i\xi f(t)} dt,$$

где $f(t)$ — функция, обратная $\sigma(\omega)$ в том смысле, как это описано выше. Таким образом,

$$\frac{1}{T} \int_0^T |c(\xi)|^2 d\xi = \int_0^1 \int_0^1 \frac{1}{T} \int_0^T e^{i\xi(f(s)-f(t))} d\xi ds dt$$

и

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T e^{i\xi(f(t)-f(s))} d\xi = \begin{cases} 0, & f(t) \neq f(s), \\ 1, & f(t) = f(s). \end{cases}$$

Отсюда по теореме об ограниченной сходимости следует, что

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T |c(\xi)|^2 d\xi$$

существует и равен мере множества точек (t, s) ($0 \leq t, s \leq 1$) плоскости, для которых $f(t) = f(s)$. Это эквивалентно нашей теореме.

4. Доказать, что

$$f(t) = \sum_{k=1}^{\infty} c_k r_k(t), \quad \sum_{k=1}^{\infty} c_k^2 < \infty$$

не может быть константой на множестве положительной меры, за исключением случая, когда все величины c_k , кроме конечного числа их, равны 0

5. Закон природы или математическая теорема? В заключение главы мы рассмотрим пример, который поучителен с точки зрения как используемых понятий, так и технических приемов.

Сначала нам понадобятся три определения.

1. *Относительная мера.* Пусть A — множество действительных чисел. Рассмотрим часть A , которая лежит в интервале $(-T, T)$, т. е. $A \cap (-T, T)$. Относительная мера множества $\mu_R \{A\}$ определяется как предел

$$\mu_R \{A\} = \lim_{T \rightarrow \infty} \frac{1}{2T} \mu \{A \cap (-T, T)\}, \quad (5.1)$$

если он существует. Относительная мера не является *вообще аддитивной*; так, если $A_i = (i, i+1)$, $i = 0, \pm 1, \pm 2, \dots$, то

$$\mu_R \left\{ \bigcup_{i=-\infty}^{\infty} A_i \right\} = 1,$$

в то время как

$$\sum_{i=-\infty}^{\infty} \mu_R \{A_i\} = 0.$$

2. *Среднее значение функции.* Среднее значение $M \{f(t)\}$ функции $f(t)$, $-\infty < t < \infty$, определяется как предел

$$M \{f(t)\} = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T f(t) dt, \quad (5.2)$$

если он существует.

3. *Линейная независимость действительных чисел.* Действительные числа $\lambda_1, \lambda_2, \dots$ называются линейно независимыми (или независимыми над полем рациональных чисел), если единственным решением (k_1, k_2, \dots) (в целых числах) уравнения

$$k_1\lambda_1 + k_2\lambda_2 + \dots = 0 \quad (5.3)$$

является

$$k_1 = k_2 = k_3 = \dots = 0.$$

Наиболее известный пример линейно независимых чисел представляет собой последовательность

$$\log p_1, \log p_2, \log p_3, \dots \quad (5.4)$$

логарифмов простых чисел ($p_1 = 2, p_2 = 3, \dots$). Как читатель, несомненно, отметит, линейная независимость величин (5.4) эквивалентна теореме о единственности разложения чисел на простые сомножители. Это простое, но замечательное наблюдение было сделано в 1910 г. Г. Бором, для которого оно стало отправной точкой нового подхода к проблемам, связанным со знаменитой ζ -функцией Римана.

Пусть теперь $\lambda_1, \lambda_2, \dots$ линейно независимы. Рассмотрим функцию

$$\sqrt{2} \frac{\cos \lambda_1 t + \dots + \cos \lambda_n t}{\sqrt{n}}. \quad (5.5)$$

Обозначим через $A_n(\omega_1, \omega_2)$ множество, на котором

$$\omega_1 < \sqrt{2} \frac{\cos \lambda_1 t + \dots + \cos \lambda_n t}{\sqrt{n}} < \omega_2. \quad (5.6)$$

Мы можем теперь доказать, что $\mu_R \{A_n(\omega_1, \omega_2)\}$ опре

делена и, более того, что

$$\lim_{n \rightarrow \infty} \mu_R \{A_n(\omega_1, \omega_2)\} = \frac{1}{\sqrt{2\pi}} \int_{\omega_1}^{\omega_2} e^{-y^2/2} dy. \quad (5.7)$$

Используя систему обозначений п. 3 этой главы, будем иметь

$$\begin{aligned} \frac{1}{2T} \int_{-T}^T g_{\varepsilon}^{-} \left(\sqrt{2} \frac{\cos \lambda_1 t + \dots + \cos \lambda_n t}{\sqrt{n}} \right) dt &\leq \\ &\leq \frac{1}{2T} \int_{-T}^T g \left(\sqrt{2} \frac{\cos \lambda_1 t + \dots + \cos \lambda_n t}{\sqrt{n}} \right) dt \leq \\ &\leq \frac{1}{2T} \int_{-T}^T g_{\varepsilon}^{+} \left(\sqrt{2} \frac{\cos \lambda_1 t + \dots + \cos \lambda_n t}{\sqrt{n}} \right) dt \end{aligned} \quad (5.8)$$

и

$$\begin{aligned} \frac{1}{2T} \int_{-T}^T g_{\varepsilon}^{\pm} \left(\sqrt{2} \frac{\cos \lambda_1 t + \dots + \cos \lambda_n t}{\sqrt{n}} \right) dt &= \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} G_{\varepsilon}^{\pm}(\xi) \left[\frac{1}{2T} \int_{-T}^T \exp \left(i\xi \sqrt{2} \times \right. \right. \\ &\times \left. \left. \frac{\cos \lambda_1 t + \dots + \cos \lambda_n t}{\sqrt{n}} \right) dt \right] d\xi^1, \end{aligned} \quad (5.9)$$

¹⁾ Напомним, что мы воспользовались сокращением

$$G_{\varepsilon}^{\pm}(\xi) = \int_{-\infty}^{\infty} g_{\varepsilon}^{\pm}(x) e^{i\xi x} dx.$$

где обе функции $G_E^+(\xi)$ и $G_E^-(\xi)$ абсолютно интегрируемы в интервале $(-\infty, \infty)$. (Таким образом, изменение порядка интегрирования легко обосновывается.)

Докажем теперь, что

$$\begin{aligned} \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \exp \left(i\xi \sqrt{2} \frac{\cos \lambda_1 t + \dots + \cos \lambda_n t}{\sqrt{n}} \right) dt = \\ = J_0^n \left(\sqrt{2} \frac{\xi}{\sqrt{n}} \right), \end{aligned} \quad (5.10)$$

где J_0 — известная функция Бесселя.

Мы проведем доказательство для $n=2$, так как доказательство для произвольного n точно такое же.

Мы имеем (обозначая $\eta = \xi \sqrt{2}/\sqrt{n}$)

$$\begin{aligned} \frac{1}{2T} \int_{-T}^T e^{i\eta(\cos \lambda_1 t + \cos \lambda_2 t)} dt = \\ = \sum_{k, l=0}^{\infty} \frac{(i\eta)^k (i\eta)^l}{k!l!} \frac{1}{2T} \int_{-T}^T \cos^k \lambda_1 t \cos^l \lambda_2 t dt \end{aligned} \quad (5.11)$$

Задача состоит в том, чтобы найти

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \cos^k \lambda_1 t \cos^l \lambda_2 t dt = M \{ \cos^k \lambda_1 t \cos^l \lambda_2 t \}.$$

Вычисляем

$$\begin{aligned} \cos^k \lambda_1 t \cos^l \lambda_2 t &= \frac{1}{2^k} \frac{1}{2^l} (e^{i\lambda_1 t} + e^{-i\lambda_1 t})^k (e^{i\lambda_2 t} + e^{-i\lambda_2 t})^l = \\ &= \frac{1}{2^k} \frac{1}{2^l} \sum_{r=0}^k \sum_{s=0}^l \binom{k}{r} \binom{l}{s} e^{i[(2r-k)\lambda_1 + (2s-l)\lambda_2]t} \end{aligned}$$

и

$$M \{ e^{i\alpha t} \} = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T e^{i\alpha t} dt = \begin{cases} 1, & \alpha = 0, \\ 0, & \alpha \neq 0. \end{cases}$$

Вследствие линейной независимости λ_1 и λ_2 выражение

$$(2r - k)\lambda_1 + (2s - l)\lambda_2$$

может быть равно нулю только тогда, когда $2r = k$ и $2s = l$, отсюда почти немедленно следует, что

$$M \{ \cos^k \lambda_1 t \cos^l \lambda_2 t \} = \frac{1}{2^k} \binom{k}{\frac{k}{2}} \frac{1}{2^l} \binom{l}{\frac{l}{2}}, \quad (5.12)$$

если одновременно k и l четны, и равны 0 во всех других случаях. Мы можем записать (5.12) в форме

$$M \{ \cos^k \lambda_1 t \cos^l \lambda_2 t \} = M \{ \cos^k \lambda_1 t \} M \{ \cos^l \lambda_2 t \}. \quad (5.13)$$

Теперь, сопоставляя это с (5.11), мы получаем

$$M \{ e^{i\eta(\cos \lambda_1 t + \cos \lambda_2 t)} \} = M \{ e^{i\eta \cos \lambda_1 t} \} M \{ e^{i\eta \cos \lambda_2 t} \}. \quad (5.14)$$

Ясно, что

$$M \{ e^{i\eta \cos \lambda t} \} = \frac{1}{2\pi} \int_0^{2\pi} e^{i\eta \cos \theta} d\theta = J_0(\eta) \quad (5.15)$$

и, следовательно, [по (5.14)]

$$M \{ e^{i\eta(\cos \lambda_1 t + \cos \lambda_2 t)} \} = J_0^2(\eta).$$

Таким образом, формула (5.10) доказана. Полагая теперь в (5.8) $T \rightarrow \infty$, мы с помощью (5.9) и (5.10) получаем

$$\begin{aligned} & \frac{1}{2\pi} \int_{-\infty}^{\infty} G_{\varepsilon}^{-}(\xi) J_0^n \left(V\sqrt{2} \frac{\xi}{\sqrt{n}} \right) d\xi \leq \\ & \leq \liminf_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T g \left(V\sqrt{2} \frac{\cos \lambda_1 t + \dots + \cos \lambda_n t}{\sqrt{n}} \right) dt \leq \\ & \leq \limsup_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T g \left(V\sqrt{2} \frac{\cos \lambda_1 t + \dots + \cos \lambda_n t}{\sqrt{n}} \right) dt \leq \\ & \leq \frac{1}{2\pi} \int_{-\infty}^{\infty} G_{\varepsilon}^{+}(\xi) J_0^n \left(V\sqrt{2} \frac{\xi}{\sqrt{n}} \right) d\xi. \end{aligned} \quad (5.16)$$

Хорошо известно, что при $\eta \rightarrow \pm \infty$

$$J_0(\eta) = O \left(\frac{1}{\sqrt{|\eta|}} \right)$$

и, следовательно, при $n \geq 3$ функция

$$J_0^n \left(\sqrt{2} \frac{\xi}{\sqrt{n}} \right)$$

абсолютно интегрируема по ξ . Из этого вытекает, что (при $n \geq 3$)

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \frac{1}{2\pi} \int_{-\infty}^{\infty} G_{\varepsilon}^{-}(\xi) J_0^n \left(\sqrt{2} \frac{\xi}{\sqrt{n}} \right) d\xi &= \\ &= \lim_{\varepsilon \rightarrow 0} \frac{1}{2\pi} \int_{-\infty}^{\infty} G_{\varepsilon}^{+}(\xi) J_0^n \left(\sqrt{2} \frac{\xi}{\sqrt{n}} \right) d\xi \end{aligned}$$

и поэтому предел

$$\begin{aligned} \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T g \left(\sqrt{2} \frac{\cos \lambda_1 t + \dots + \cos \lambda_n t}{\sqrt{n}} \right) dt &= \\ &= \mu_R \{A_n(\omega_1, \omega_2)\} \end{aligned}$$

существует!¹⁾. Теперь (5.16) может быть записано в форме

$$\begin{aligned} \frac{1}{2\pi} \int_{-\infty}^{\infty} G_{\varepsilon}^{-}(\xi) J_0^n \left(\sqrt{2} \frac{\xi}{\sqrt{n}} \right) d\xi &\leq \mu_R \{A_n(\omega_1, \omega_2)\} \leq \\ &\leq \frac{1}{2\pi} \int_{-\infty}^{\infty} G_{\varepsilon}^{+}(\xi) J_0^n \left(\sqrt{2} \frac{\xi}{\sqrt{n}} \right) d\xi \end{aligned}$$

¹⁾ Для $n=1$ и $n=2$ это также верно, однако доказательство изменяется.

и легко проверяется, что

$$\lim_{n \rightarrow \infty} J_0^n \left(\sqrt{2} \frac{\xi}{\sqrt{n}} \right) = e^{-\xi^2/2}.$$

Доказательство (5.7) может быть теперь закончено, как в п. 3. Если мы взглянем на

$$q_n(t) = \sqrt{2} \frac{\cos \lambda_1 t + \dots + \cos \lambda_n t}{\sqrt{n}},$$

как на результат суперпозиции колебаний с несоизмеримыми частотами, то теорема, содержащаяся в (5.7), дает точную информацию об относительном времени, которое $q_n(t)$ проводит между ω_1 и ω_2 . То, что мы пришли здесь к нормальному закону

$$\frac{1}{\sqrt{2\pi}} \int_{\omega_1}^{\omega_2} e^{-y^2/2} dy,$$

связанному обычно со случайными явлениями, может служить указанием на то, что детерминистская и вероятностная точки зрения не такие уж непримиримые, какими они кажутся с первого взгляда. Далее мы не будем останавливаться на этом вопросе, ибо это увело бы нас слишком далеко в сторону. Однако может быть уместной ссылка на Пуанкаре, который сказал (отчасти, несомненно, в шутку), что в нормальном законе должно быть что-то таинственное, так как математики считают его законом природы, тогда как физики убеждены в том, что он является математической теоремой.

ЗАДАЧИ

1. Доказать, что если $\lambda_1, \dots, \lambda_n$ линейно независимы, то функции $\cos \lambda_1 t, \dots, \cos \lambda_n t$ статистически независимы, т. е. для всех действительных $\alpha_1, \dots, \alpha_n$

$$\mu_R \{ \cos \lambda_1 t < \alpha_1, \dots, \cos \lambda_n t < \alpha_n \} = \prod_{k=1}^n \mu_R \{ \cos \lambda_k t < \alpha_k \}.$$

[Это, конечно, то свойство, которое лежит в основе доказательства утверждения (5.7).]

2. Пусть $s = \sigma + it$, $\sigma > 1$. Рассмотрим ζ -функцию Римана

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_p \frac{1}{1 - \frac{1}{p^s}}.$$

Доказать, что при $l > 0$

$$M \{ |\zeta(\sigma + it)|^l \} = M \left\{ \frac{1}{|\zeta(\sigma + it)|^{l-2}} \right\} \zeta^{l-1}(2\sigma).$$

ЛИТЕРАТУРА

Марков А., Исчисление вероятностей, М., 1924.

Лоэв М., Теория вероятностей, ИЛ, М., 1962. Эта книга содержит в полном объеме теорию функций распределения и, в частности, указанные выше результаты П. Леви.

Кас М., Steinhaus Н., Sur les fonctions indépendantes IV, *Studia Math.*, 7 (1938), 1—15.

ПРОСТЫЕ ЧИСЛА «ИГРАЮТ В АЗАРТНУЮ ИГРУ»

1. Теоретико-числовые функции, плотность, независимость. Теоретико-числовая функция $f(n)$ — это функция, заданная на множестве чисел натурального ряда $1, 2, 3, \dots$. Среднее значение $M\{f(n)\}$ функции $f(n)$ определяется как предел (если он существует)

$$M\{f(n)\} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f(n). \quad (1.1)$$

Пусть A — некоторое множество положительных целых чисел. Обозначим $A(N)$ количество тех его элементов, которые содержатся среди первых N чисел натурального ряда. Если существует предел

$$\lim_{N \rightarrow \infty} \frac{A(N)}{N} = D(A), \quad (1.2)$$

то он называется *плотностью* A . Плотность аналогична относительной мере (см. п. 5 гл. 3) и, подобно ей, не является вполне аддитивной. Рассмотрим целые числа, делящиеся на простое число p . Плотность множества таких чисел, очевидно, равна $1/p$. Возьмем

теперь множество целых чисел, которые делятся одновременно на p и q (q — другое простое число). Делимость на p и q эквивалентна делимости на pq , и, следовательно, плотность нового множества равна $1/pq$. Так как

$$\frac{1}{pq} = \frac{1}{p} \cdot \frac{1}{q}, \quad (1.3)$$

то мы можем истолковать это так: «события», заключающиеся в делимости на p и q , независимы. Это, конечно, выполняется для любого количества простых чисел, и мы можем сказать, употребляя образное, но не очень точное выражение, что простые числа играют в азартную игру! Это простое, почти тривиальное наблюдение является источником нового направления, которое существенным образом связывает теорию чисел, с одной стороны, и теорию вероятностей — с другой.

Мы проиллюстрируем в деталях некоторые элементарные аспекты этого направления и кратко обсудим более сложные.

2. Статистика значений ϕ -функций Эйлера. Обозначим через $\phi(n)$ количество целых чисел, не превосходящих n и взаимно простых с ним. Эта теоретико-числовая функция, впервые введенная Эйлером, имеет много применений и представляет значительный интерес сама по себе.

Сразу же проверяется, что если

$$(m, n) = 1$$

(т. е. m и n взаимно просты), то

$$\varphi(mn) = \varphi(m)\varphi(n) \quad (2.1)$$

и

$$\varphi(p^\alpha) = p^\alpha - p^{\alpha-1}. \quad (2.2)$$

Таким образом,

$$\varphi(n) = \prod_{\substack{p^\alpha | n \\ p^{\alpha+1} \nmid n}} (p^\alpha - p^{\alpha-1}) \quad (2.3)$$

или так как

$$n = \prod_{\substack{p^\alpha | n \\ p^{\alpha+1} \nmid n}} p^\alpha \quad (2.4)$$

(единственность разложения на простые множители!), то

$$\frac{\varphi(n)}{n} = \prod_{p|n} \left(1 - \frac{1}{p}\right). \quad (2.5)$$

Введем теперь функции $\varrho_p(n)$ следующим образом:

$$\varrho_p(n) = \begin{cases} 1, & p | n, \\ 0, & p \nmid n. \end{cases} \quad (2.6)$$

В терминах этих функций мы можем записать

$$\frac{\varphi(n)}{n} = \prod_p \left(1 - \frac{\varrho_p(n)}{p}\right). \quad (2.7)$$

Заметим теперь, что если ε_j равны или 0, или 1, то

$$\begin{aligned} D \{ \varrho_{p_1}(n) = \varepsilon_1, \varrho_{p_2}(n) = \varepsilon_2, \dots, \varrho_{p_k}(n) = \varepsilon_k \} = \\ = D \{ \varrho_{p_1}(n) = \varepsilon_1 \} D \{ \varrho_{p_2}(n) = \varepsilon_2 \} \dots D \{ \varrho_{p_k}(n) = \varepsilon_k \}. \end{aligned} \quad (2.8)$$

Это — просто другой способ записи того факта, что «события», заключающиеся в делимости на p_1, p_2, \dots, p_k , независимы (или что функции $\varrho_p(n)$ независимы).

Из свойства (2.8) следует, что

$$\begin{aligned} M \left\{ \prod_{p \leq p_k} \left(1 - \frac{\varrho_p(n)}{p} \right) \right\} = \prod_{p \leq p_k} M \left\{ \left(1 - \frac{\varrho_p(n)}{p} \right) \right\} = \\ = \prod_{p \leq p_k} \left(1 - \frac{1}{p^2} \right). \end{aligned} \quad (2.9)$$

Поэтому мы можем предполагать, что

$$\begin{aligned} M \left\{ \frac{\varphi(n)}{n} \right\} = M \left\{ \prod_p \left(1 - \frac{\varrho_p(n)}{p} \right) \right\} = \\ = \prod_p M \left\{ \left(1 - \frac{\varrho_p(n)}{p} \right) \right\} = \prod_p \left(1 - \frac{1}{p^2} \right) = \frac{1}{\zeta(2)} = \frac{6}{\pi^2}. \end{aligned} \quad (2.10)$$

К сожалению, (2.10) невозможно вывести прямо из (2.9), так как плотность D не является вполне аддитивной.

Но, с другой стороны, (2.10) легко получается следующим образом.

Из (2.5) вытекает, что

$$\frac{\varphi(n)}{n} = \sum_{d|n} \frac{\mu(d)}{d}, \quad (2.11)$$

где $\mu(d)$ — функция Мёбиуса, определяемая свойствами:

1) $\mu(1) = 1$;

2) $\mu(m) = 0$, если m делится на квадрат простого числа;

3) $\mu(m) = (-1)^v$, если m является произведением v различных простых чисел.

После этого получаем

$$\frac{1}{N} \sum_{n=1}^N \frac{\varphi(n)}{n} = \frac{1}{N} \sum_{d=1}^N \frac{\mu(d)}{d} \left[\frac{N}{d} \right], \quad (2.12)$$

и, следовательно,

$$M \left\{ \frac{\varphi(n)}{n} \right\} = \sum_{d=1}^{\infty} \frac{\mu(d)}{d^2} = \prod_p \left(1 - \frac{1}{p^2} \right) = \frac{1}{\zeta(2)} = \frac{6}{\pi^2}. \quad (2.13)$$

Обозначим

$$f_k(n) = \prod_{p \leq p_k} \left(1 - \frac{\varrho_p(n)}{p} \right) \quad (2.14)$$

и рассмотрим

$$f_k(n) - \frac{\varphi(n)}{n}.$$

Мы, очевидно, имеем

$$0 \leq f_k(n) - \frac{\varphi(n)}{n} \leq 1, \quad (2.15)$$

и, кроме того, согласно (2.13) и (2.9),

$$\begin{aligned} M \left\{ f_k(n) - \frac{\varphi(n)}{n} \right\} &= \\ &= \prod_{p \leq p_k} \left(1 - \frac{1}{p^2} \right) - \prod_p \left(1 - \frac{1}{p^2} \right). \end{aligned} \quad (2.16)$$

Далее, при $l > 1$

$$0 \leq f_k^l(n) - \left(\frac{\varphi(n)}{n} \right)^l \leq l \left(f_k(n) - \frac{\varphi(n)}{n} \right) \quad (2.17)$$

и потому

$$\begin{aligned} \frac{1}{N} \sum_{n=1}^N f_k^l(n) &\geq \frac{1}{N} \sum_{n=1}^N \left(\frac{\varphi(n)}{n} \right)^l \geq \\ &\geq \frac{1}{N} \sum_{n=1}^N f_k^l(n) - \frac{l}{N} \sum_{n=1}^N \left(f_k(n) - \frac{\varphi(n)}{n} \right). \end{aligned}$$

Полагая $N \rightarrow \infty$, получаем

$$\begin{aligned}
 M \{f_k^l(n)\} &\geq \\
 &\geq \limsup_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \left(\frac{\varphi(n)}{n}\right)^l \geq \liminf_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \left(\frac{\varphi(n)}{n}\right)^l \geq \\
 &\geq M \{f_k^l(n)\} - LM \left\{f_k(n) - \frac{\varphi(n)}{n}\right\}. \quad (2.18)
 \end{aligned}$$

В то же время

$$\begin{aligned}
 M \{f_k^l(n)\} &= \\
 &= M \left\{ \prod_{p \leq p_k} \left(1 - \frac{\varrho_p(n)}{p}\right)^l \right\} = \prod_{p \leq p_k} M \left\{ \left(1 - \frac{\varrho_p(n)}{p}\right)^l \right\} = \\
 &= \prod_{p \leq p_k} \left[1 - \frac{1}{p} + \frac{1}{p} \left(1 - \frac{1}{p}\right)^l \right].
 \end{aligned}$$

Объединяя последний результат с (2.16) и (2.18), мы получим при $k \rightarrow \infty$ формулу

$$M \left\{ \left(\frac{\varphi(n)}{n}\right)^l \right\} = \prod_p \left[1 - \frac{1}{p} + \frac{1}{p} \left(1 - \frac{1}{p}\right)^l \right], \quad (2.19)$$

предложенную И. Шуром,

Формально (2.19) следует из цепочки равенств

$$\begin{aligned} M \left\{ \left(\frac{\varphi(n)}{n} \right)^l \right\} &= M \left\{ \prod_p \left(1 - \frac{\varrho_p(n)}{p} \right)^l \right\} = \\ &= \prod_p M \left\{ \left(1 - \frac{\varrho_p(n)}{p} \right)^l \right\} = \\ &= \prod_p \left[1 - \frac{1}{p} + \frac{1}{p} \left(1 - \frac{1}{p} \right)^l \right], \end{aligned}$$

но так как D не является вполне аддитивной, то необходимо обоснование, данное выше.

Из равенства (2.7) мы имеем

$$\begin{aligned} \log \frac{\varphi(n)}{n} &= \sum_p \log \left(1 - \frac{\varrho_p(n)}{p} \right) = \\ &= \sum_p \varrho_p(n) \log \left(1 - \frac{1}{p} \right), \end{aligned} \quad (2.20)$$

и опять-таки, формально,

$$\begin{aligned} M \left\{ \exp \left(i\xi \log \frac{\varphi(n)}{n} \right) \right\} &= \\ &= \prod_p M \left\{ \exp \left(i\xi \varrho_p(n) \log \left(1 - \frac{1}{p} \right) \right) \right\} = \\ &= \prod_p \left(1 - \frac{1}{p} + \frac{1}{p} \exp \left(i\xi \log \left(1 - \frac{1}{p} \right) \right) \right) = c(\xi) \end{aligned} \quad (2.21)$$

для каждого действительного ξ .

Строгое обоснование (2.21) почти совпадает с тем, что дано для (2.19) и предоставляется читателю.

Пусть теперь $K_N(\omega)$ — количество целых n , не превосходящих N , для которых

$$\log \frac{\varphi(n)}{n} < \omega.$$

Положим

$$\sigma_N(\omega) = \frac{K_N(\omega)}{N}. \quad (2.22)$$

Заметим, что $\sigma_N(\omega)$ есть функция распределения и что

$$\int_{-\infty}^{\infty} e^{i\xi\omega} d\sigma_N(\omega) = \frac{\exp\left(i\xi \log \frac{\varphi(1)}{1}\right) + \dots + \exp\left(i\xi \log \frac{\varphi(N)}{N}\right)}{N}. \quad (2.23)$$

Из равенства (2.21) следует, что

$$\lim_{N \rightarrow \infty} \int_{-\infty}^{\infty} e^{i\xi\omega} d\sigma_N(\omega) = M \left\{ \exp\left(i\xi \log \frac{\varphi(n)}{n}\right) \right\} = c(\xi). \quad (2.24)$$

Кроме того, легко показать, что $c(\xi)$ непрерывна при $\xi = 0$. Тогда по теореме, сформулированной в конце п. 4 гл. 3, существует функция распределения $\sigma(\omega)$,

такая, что

$$\int_{-\infty}^{\infty} e^{i\xi\omega} d\sigma(\omega) = c(\xi) = \prod_p \left(1 - \frac{1}{p} + \frac{1}{p} \exp \left(i\xi \log \left(1 - \frac{1}{p} \right) \right) \right) \quad (2.25)$$

и в каждой точке непрерывности $\sigma(\omega)$

$$\lim_{N \rightarrow \infty} \sigma_N(\omega) = \sigma(\omega). \quad (2.26)$$

Теперь легко доказать, что $\sigma(\omega)$ непрерывна при каждом ω . Для этого мы используем результат задачи 3 (стр. 71, гл. 3).

Имеем

$$\begin{aligned} |c(\xi)|^2 &= \\ &= \prod_p \left[\left(1 - \frac{1}{p} \right)^2 + \frac{2}{p} \left(1 - \frac{1}{p} \right) \cos \left(\xi \log \left(1 - \frac{1}{p} \right) \right) + \frac{1}{p^2} \right] \ll \\ &\ll \prod_{p \leq p_h} \left[\left(1 - \frac{1}{p} \right)^2 + \frac{2}{p} \left(1 - \frac{1}{p} \right) \cos \left(\xi \log \left(1 - \frac{1}{p} \right) \right) + \frac{1}{p^2} \right]. \end{aligned} \quad (2.27)$$

Можно показать (см. следующую за этим пунктом задачу 1), что числа

$$\log \left(1 - \frac{1}{p} \right)$$

линейно независимы,

Согласно рассуждениям п. 5 гл. 3, имеем

$$\begin{aligned} \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \prod_{p \leq p_k} \left[\left(1 - \frac{1}{p}\right)^2 + \right. \\ \left. + \frac{2}{p} \left(1 - \frac{1}{p}\right) \cos \left(\xi \log \left(1 - \frac{1}{p}\right) \right) + \frac{1}{p^2} \right] d\xi = \\ = \prod_{p \leq p_k} \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \left[\left(1 - \frac{1}{p}\right)^2 + \right. \\ \left. + \frac{2}{p} \left(1 - \frac{1}{p}\right) \cos \left(\xi \log \left(1 - \frac{1}{p}\right) \right) + \frac{1}{p^2} \right] d\xi = \\ = \prod_{p \leq p_k} \left[\left(1 - \frac{1}{p}\right)^2 + \frac{1}{p^2} \right], \end{aligned}$$

а из элементарных сведений о простых числах мы знаем, что

$$\lim_{k \rightarrow \infty} \prod_{p \leq p_k} \left[\left(1 - \frac{1}{p}\right)^2 + \frac{1}{p^2} \right] = \prod_p \left[\left(1 - \frac{1}{p}\right)^2 + \frac{1}{p^2} \right] = 0.$$

Таким образом, получаем

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T |c(\xi)|^2 d\xi = 0, \quad (2.28)$$

и, следовательно, $\sigma(\omega)$ непрерывна при всех ω . Суммируем все сказанное.

Плотность

$$D \left\{ \log \frac{\varphi(n)}{n} < \omega \right\} = \sigma(\omega)$$

существует при каждом ω , $\sigma(\omega)$ непрерывна и

$$\int_{-\infty}^{\infty} e^{i\xi\omega} d\sigma(\omega) = \prod_p \left[\left(1 - \frac{1}{p}\right) + \frac{1}{p} \exp\left(i\xi \log\left(1 - \frac{1}{p}\right)\right) \right].$$

Этот результат (впервые полученный Шёнбергом) мог быть выведен более элементарным путем и был в значительной степени обобщен Эрдёшем¹⁾. Мы выбрали более длинный путь, чтобы выявить своеобразие вероятностного оттенка результата и продемонстрировать взаимодействие ряда идей и методов исследования.

Формула (2.21) является очевидным аналогом формулы

$$\frac{\sin \xi}{\xi} = \prod_{k=1}^{\infty} \cos \frac{\xi}{2^k},$$

с которой мы начинали. Это, так сказать, вариации одной и той же темы, и то, что тема допускает столь разнообразные вариации, есть очевидное подтверждение богатства ее «мелодического» содержания.

¹⁾ Эрдёш доказал также замечательную теорему о том, что наша $\sigma(\omega)$ сингулярна, т. е. $\sigma'(\omega) = 0$ почти всюду.

ЗАДАЧИ

1. Доказать, что величины $\log(1 - (1/p))$, а также и $\log(1 + (1/p))$ линейно независимы.

2. Статистика значений функции $\sigma(n)$ (суммы делителей числа n).

(а) Пусть $\alpha_p(n)$ обозначает показатель степени, в которой простое число p появляется (в единственном) представлении n в виде произведения степеней его простых делителей, т. е.

$$n = \prod_p p^{\alpha_p(n)}.$$

Доказать, что функции $\alpha_p(n)$ статистически независимы.

(б) Показать, что если $\sigma(n)$ обозначает сумму всех делителей n , то

$$\frac{\sigma(n)}{n} = \prod_p \left(1 + \frac{1}{p} + \dots + \frac{1}{p^{\alpha_p(n)}} \right).$$

(в) Используя то, что

$$\frac{\sigma(n)}{n} = \sum_{k|n} \frac{1}{k},$$

доказать равенство

$$M \left\{ \frac{\sigma(n)}{n} \right\} = \frac{\pi^2}{6}.$$

(г) Показать, что

$$M \left\{ \prod_{p \leq p_k} \left(1 + \frac{1}{p} + \dots + \frac{1}{p^{\alpha_p(n)}} \right) \right\} = \prod_{p \leq p_k} \frac{1}{1 - \frac{1}{p^2}}.$$

(д) Положим

$$f_k(n) = \prod_{p \leq p_k} \left(1 + \frac{1}{p} + \dots + \frac{1}{p^{\alpha_p(n)}} \right).$$

Исходя из того, что

$$\frac{f_k(n)}{\sigma(n)} = \prod_{p > p_k} \frac{1}{1 + \frac{1}{p} + \dots + \frac{1}{p^{\alpha_p(n)}}},$$

вывести неравенство

$$\prod_{p > p_k} \left(1 - \frac{Q_p(n)}{p}\right) \leq \frac{f_k(n)}{\sigma(n)} \leq \prod_{p > p_k} \frac{1}{1 + \frac{Q_p(n)}{p}}.$$

(е) Показать, что

$$\begin{aligned} M \{ e^{i\xi \log \frac{\sigma(n)}{n}} \} &= \prod_p M \left\{ \exp \left[i\xi \log \left(1 + \frac{1}{p} + \dots + \frac{1}{p^{\alpha_p(n)}} \right) \right] \right\} = \\ &= \prod_p \left[1 - \frac{1}{p} + \sum_{\alpha=1}^{\infty} \left(\frac{1}{p^\alpha} - \frac{1}{p^{\alpha+1}} \right) \times \right. \\ &\quad \left. \times \exp \left[i\xi \log \left(1 + \frac{1}{p} + \dots + \frac{1}{p^\alpha} \right) \right] \right] = c(\xi). \end{aligned}$$

(ж) Используя то, что

$$\begin{aligned} \left| 1 - \frac{1}{p} + \sum_{\alpha=1}^{\infty} \left(\frac{1}{p^\alpha} - \frac{1}{p^{\alpha+1}} \right) \exp \left[i\xi \log \left(1 + \frac{1}{p} + \dots + \right. \right. \right. \\ \left. \left. \left. + \frac{1}{p^\alpha} \right) \right] \right| \leq \left| 1 - \frac{1}{p} + \frac{1}{p} \left(1 - \frac{1}{p} \right) \exp \left[i\xi \log \left(1 + \frac{1}{p} \right) \right] \right| + \\ + \frac{1}{p^2} = \left(1 - \frac{1}{p} \right) \sqrt{1 + \frac{2}{p} \cos \left[\xi \log \left(1 + \frac{1}{p} \right) \right] + \frac{1}{p^2} + \frac{1}{p^2}}, \end{aligned}$$

а также факт линейной независимости величин $\log(1+(1/p))$, доказать, что плотность

$$D \left\{ \frac{\sigma(n)}{n} < \omega \right\} = \tau(\omega)$$

существует и непрерывна по ω .

Этот результат, полученный впервые Г. Давенпортом, содержится в более общей теореме Эрдёша.

Случай $\omega=2$ представляет собой интерес, так как показывает, что «избыточные числа» (т. е. числа, для которых $\sigma(n) > 2n$), а также и «дефектные числа» (т. е. те, для которых $\sigma(n) < 2n$) имеют плотность. Из этого также следует, что «совершенные числа» (для которых $\sigma(n) = 2n$) имеют плотность, равную 0. Предполагают, что существует лишь конечное число «совершенных чисел».

3. *Формула обращения.* Пусть

$$\int_{-\infty}^{\infty} e^{i\xi\omega} d\sigma(\omega) = c(\xi),$$

где $\sigma(\omega)$ — функция распределения. Доказать, что если ω_1 и ω_2 — точки непрерывности функции σ , то

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{i\omega_2\xi} - e^{i\omega_1\xi}}{i\xi} c(\xi) d\xi = \sigma(\omega_2) - \sigma(\omega_1).$$

Если же или ω_1 , или ω_2 (или и та, и другая) являются точками разрыва, то величины $\sigma(\omega_1)$ или $\sigma(\omega_2)$ (или обе) следует заменить на

$$\frac{\sigma(\omega_1-0) + \sigma(\omega_1+0)}{2} \quad \text{или} \quad \frac{\sigma(\omega_2-0) + \sigma(\omega_2+0)}{2}$$

соответственно. В частности, показать, что

$$D \left\{ \omega_1 < \log \frac{\varphi(n)}{n} < \omega_2 \right\} = \\ = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{i\omega_2 \xi} - e^{i\omega_1 \xi}}{i\xi} \prod_p \left(1 - \frac{1}{p} + \frac{1}{p} \exp \left[i\xi \log \left(1 - \frac{1}{p} \right) \right] \right) d\xi.$$

Точная, но почти бесполезная формула!

3. Другое применение. Пусть $\omega(n)$ обозначает количество простых делителей n с учетом их кратности, т. е.

$$\omega(n) = \sum_p \alpha_p(n), \quad (3.1)$$

где величины $\alpha_p(n)$ определяются, как в задаче 2 п. 2 этой главы.

Пусть $\nu(n)$ обозначает количество простых делителей n без учета кратности, т. е.

$$\nu(n) = \sum_p \varrho_p(n). \quad (3.2)$$

Назовем разность $\omega(n) - \nu(n)$ эксцессом и найдем плотность целых чисел, для которых эксцесс равен k ($k \geq 0$, целое), т. е. найдем

$$d_k = D \{ \omega(n) - \nu(n) = k \}. \quad (3.3)$$

Излишне говорить, что существование плотности не очевидно и должно быть доказано.

Мы начнем с формулы (5.3) гл. 1

$$\frac{1}{2\pi} \int_0^{2\pi} e^{imx} dx = \begin{cases} 1, & m=0, \\ 0, & m \neq 0, \end{cases} \quad (3.4)$$

где m целое, и рассмотрим

$$\begin{aligned} \frac{1}{N} \sum_{n=1}^N \frac{1}{2\pi} \int e^{i(\omega(n)-\nu(n)-k)x} dx &= \\ &= \frac{1}{2\pi} \int_0^{2\pi} e^{-ikx} \frac{1}{N} \sum_{n=1}^N e^{i(\omega(n)-\nu(n))x} dx. \end{aligned} \quad (3.5)$$

Левая часть равенства (3.5) представляет собой [ввиду (3.4)] долю целых чисел $n \leq N$, для которых эксцесс равен k . Таким образом,

$$d_k = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \frac{1}{2\pi} \int_0^{2\pi} e^{i(\omega(n)-\nu(n)-k)x} dx, \quad (3.6)$$

если предел существует.

Обращаясь еще раз к принципу ограниченной сходимости, мы видим из равенства (3.5), что достаточно будет доказать следующее: предел

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N e^{i(\omega(n)-\nu(n))x} = M \{e^{i(\omega(n)-\nu(n))x}\} \quad (3.7)$$

существует для каждого действительного x .

Далее

$$\omega(n) - \nu(n) = \sum_p (\alpha_p(n) - \varrho_p(n)),$$

и легко заметить, что функции $\alpha_p(n) - \varrho_p(n)$ независимы. Это позволяет предположить, что предел (3.7) не только существует, но также что

$$\begin{aligned} & M \{ e^{i(\omega(n) - \nu(n))x} \} = \\ & = M \left\{ \exp \left[ix \sum_p (\alpha_p(n) - \varrho_p(n)) \right] \right\} = \\ & = \prod_p M \{ e^{ix(\alpha_p(n) - \varrho_p(n))} \} = \\ & = \prod_p \left[1 - \frac{1}{p} + \sum_{\alpha=1}^{\infty} \left(\frac{1}{p^\alpha} - \frac{1}{p^{\alpha+1}} \right) e^{ix(\alpha-1)} \right] = \\ & = \prod_p \left(1 - \frac{1}{p} \right) \left(1 + \frac{1}{p - e^{ix}} \right). \end{aligned} \quad (3.8)$$

Строгое обоснование этого вывода несложно и подобно приведенному в п. 2. Возьмем сначала

$$\sum_{n=1}^N (\alpha_p(n) - \varrho_p(n))$$

и рассмотрим целые n , $1 \leq n \leq N$, для которых

$$\alpha_p(n) = \beta.$$

Эти целые числа делятся на p^β , но не делятся на $p^{\beta+1}$, следовательно, их количество равно

$$\left[\frac{N}{p^\beta} \right] - \left[\frac{N}{p^{\beta+1}} \right].$$

Отсюда, таким образом, следует, что

$$\sum_{n=1}^N (\alpha_p(n) - \varrho_p(n)) = \sum_{\beta \geq 2} (\beta - 1) \left\{ \left[\frac{N}{p^\beta} \right] - \left[\frac{N}{p^{\beta+1}} \right] \right\}. \quad (3.9)$$

Пусть теперь

$$g_k(n) = \sum_{p > p_k} (\alpha_p(n) - \varrho_p(n)). \quad (3.10)$$

Заметим, что из (3.9) вытекает следующее:

$$\begin{aligned} M \{g_k(n)\} &= \sum_{p > p_k} \sum_{\beta \geq 2} (\beta - 1) \left(\frac{1}{p^\beta} - \frac{1}{p^{\beta+1}} \right) < \\ &< \sum_{p > p_k} \sum_{\beta \geq 2} (\beta - 1) \frac{1}{p^\beta} = \sum_{p > p_k} \frac{1}{(p-1)^2}. \end{aligned} \quad (3.11)$$

Тогда

$$\begin{aligned} &\frac{1}{N} \sum_{n=1}^N e^{ix(\omega(n) - \nu(n))} = \\ &= \frac{1}{N} \sum_{n=1}^N \exp \left[ix \sum_{p \leq p_k} (\alpha_p(n) - \varrho_p(n)) \right] e^{ixg_k(n)} \end{aligned} \quad (3.12)$$

И ПОТОМУ

$$\begin{aligned}
 & \left| \frac{1}{N} \sum_{n=1}^N e^{ix(\omega(n) - \nu(n))} - \right. \\
 & \left. - \frac{1}{N} \sum_{n=1}^N \exp \times \left[ix \sum_{p \leq p_k} (\alpha_p(n) - \varrho_p(n)) \right] \right| = \\
 & = \left| \frac{1}{N} \sum_{n=1}^N \exp \left[ix \sum_{p \leq p_k} (\alpha_p(n) - \varrho_p(n)) \right] (e^{ixg_k(n)} - 1) \right| \ll \\
 & \ll \frac{1}{N} \sum_{n=1}^N |e^{ixg_k(n)} - 1| \ll \frac{|x|}{N} \sum_{n=1}^N g_k(n).
 \end{aligned}$$

Так как

$$\begin{aligned}
 & \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \exp \left[ix \sum_{p \leq p_k} (\alpha_p(n) - \varrho_p(n)) \right] = \\
 & = M \left\{ \exp \left[ix \sum_{p \leq p_k} (\alpha_p(n) - \varrho_p(n)) \right] \right\} = \\
 & = \prod_{p \leq p_k} M \{ e^{ix(\alpha_p(n) - \varrho_p(n))} \} = \\
 & = \prod_{p \leq p_k} \left(1 - \frac{1}{p} \right) \left(1 + \frac{1}{p - e^{ix}} \right),
 \end{aligned}$$

то в силу (3.14) заключаем, что расстояние каждой предельной точки последовательности

$$\frac{1}{N} \sum_{n=1}^N e^{ix(\omega(n) - \nu(n))}$$

от

$$\prod_{p \leq p_k} \left(1 - \frac{1}{p}\right) \left(1 + \frac{1}{p - e^{ix}}\right)$$

меньше, чем

$$|x| \sum_{p > p_k} \frac{1}{(p-1)^2}.$$

Так как k произвольно, отсюда сразу следует, что

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N e^{ix(\omega(n) - \nu(n))} &= M \{e^{ix(\omega(n) - \nu(n))}\} = \\ &= \prod_p \left(1 - \frac{1}{p}\right) \left(1 + \frac{1}{p - e^{ix}}\right) \end{aligned} \quad (3.13)$$

и, таким образом, (3.8) обосновано.

Возвращаясь назад к (3.5) и (3.6), мы получаем

$$\begin{aligned} d_k &= D \{\omega(n) - \nu(n) = k\} = \\ &= \frac{1}{2\pi} \int_0^{2\pi} e^{ikx} \prod_p \left(1 - \frac{1}{p}\right) \left(1 + \frac{1}{p - e^{ix}}\right) dx. \end{aligned} \quad (3.14)$$

Рассмотрим теперь функцию

$$F(z) = \prod_p \left(1 - \frac{1}{p}\right) \left(1 + \frac{1}{p-z}\right). \quad (3.15)$$

Она аналитическая во всей плоскости, исключая простые полюсы в точках $z = 2, 3, 5, \dots$. В частности, $F(z)$ аналитична в круге $|z| < 2$, и ее можно разложить там в степенной ряд

$$F(z) = \sum_{k=0}^{\infty} a_k z^k,$$

радиус сходимости которого равен 2.

Каковы коэффициенты a_k ? Если мы используем известную формулу

$$a_k = \frac{1}{2\pi i} \int \frac{F(z)}{z^{k+1}} dz,$$

где интеграл берется по контуру $|z| = 1$, то, подставляя $z = e^{ix}$, получаем

$$a_k = d_k.$$

Или в других терминах

$$\sum_{k=0}^{\infty} d_k z^k = \prod_p \left(1 - \frac{1}{p}\right) \left(1 + \frac{1}{p-z}\right). \quad (3.16)$$

Эта прекрасная формула была открыта (другим способом) А. Реньи.

Хотя вывод точных формул для d_k затруднителен, совсем просто определяется асимптотическое поведение d_k при больших k .

Действительно, $F(z)$ может быть записана в форме

$$F(z) = \frac{A}{z-2} + G(z),$$

где $G(z)$ — функция, аналитическая в круге $|z| < 3$, и A (вычет в полюсе $z=2$) дается формулой

$$A = -\frac{1}{2} \prod_{p>2} \left(1 - \frac{1}{p}\right) \left(1 + \frac{1}{p-2}\right).$$

Таким образом,

$$F(z) = \frac{1}{4} \prod_{p>2} \left(1 - \frac{1}{p}\right) \left(1 + \frac{1}{p-2}\right) \sum_{k=0}^{\infty} \frac{z^k}{2^k} + \sum_{k=0}^{\infty} b_k z^k,$$

где радиус сходимости ряда $\sum b_k z^k$ равен 3. Вследствие того что

$$d_k = \frac{1}{4} \prod_{p>2} \left(1 - \frac{1}{p}\right) \left(1 + \frac{1}{p-2}\right) \frac{1}{2^k} + b_k$$

и

$$\limsup_{k \rightarrow \infty} |b_k|^{1/k} = \frac{1}{3},$$

имеем при $k \rightarrow \infty$,

$$d_k \sim \frac{1}{2^{2(k+1)}} \prod_{p>2} \left(1 - \frac{1}{p}\right) \left(1 + \frac{1}{p-2}\right) \quad (3.17)$$

ИЛИ

$$\lim_{k \rightarrow \infty} 2^{2(k+1)} d_k = \prod_{p>2} \left(1 - \frac{1}{p}\right) \left(1 + \frac{1}{p-2}\right). \quad (3.18)$$

Два частных случая (3.16) заслуживают внимания. Полагая $z = 0$, мы получим

$$d_0 = \prod_p \left(1 - \frac{1}{p^2}\right) = \frac{1}{\zeta(2)} = \frac{6}{\pi^2}.$$

Это хорошо известный результат, означающий, что плотность чисел, не делящихся на полный квадрат, равна $6/\pi^2$.

Полагая $z = 1$, получаем

$$\sum_{k=0}^{\infty} d_k = \prod_p \left(1 - \frac{1}{p}\right) \left(1 + \frac{1}{p-1}\right) = 1.$$

В силу того что множества целых чисел, для которых $\omega(n) - \nu(n) = k$, взаимно не пересекаются и вместе исчерпывают множество всех целых чисел, то этот результат мог быть совершенно тривиальным, если бы плотность была вполне аддитивной. Так как этого нет, то тот факт, что мы тем не менее получили

$$\sum_{k=0}^{\infty} d_k = 1,$$

является по крайней мере приятно удивляющим.

4. Почти каждое целое m имеет приблизительно $\log \log m$ простых делителей. Рассмотрим целые m , $1 \leq m \leq n$, для которых выполняется или

$$v(m) < \log \log n - g_n \sqrt{\log \log n} \quad (4.1)$$

или

$$v(m) > \log \log n + g_n \sqrt{\log \log n},$$

где g_n — последовательность такая, что

$$\lim_{n \rightarrow \infty} g_n = \infty. \quad (4.2)$$

Обозначим число таких m через K_n и попробуем оценить эту величину. Используем для этого чебышевский метод, объясненный в п. 1 гл. 2.

Имеем

$$\sum_{m=1}^n (v(m) - \log \log n)^2 \geq \sum' (v(m) - \log \log n)^2, \quad (4.3)$$

где штрих у символа суммы указывает на то, что суммирование распространяется только на целые m , удовлетворяющие (4.1).

Очевидно,

$$\sum' (v(m) - \log \log n)^2 \geq K_n g_n^2 \log \log n \quad (4.4)$$

и, следовательно, в силу (4.3)

$$\frac{K_n}{n} \leq \frac{1}{ng_n^2 \log \log n} \sum_{m=1}^n (v(m) - \log \log n)^2. \quad (4.5)$$

Остается оценить

$$\begin{aligned} & \sum_{m=1}^n (v(m) - \log \log n)^2 = \\ &= \sum_{m=1}^n v^2(m) - 2 \log \log n \sum_{m=1}^n v(m) + n (\log \log n)^2. \end{aligned} \quad (4.6)$$

Теперь имеем

$$v(m) = \sum_p \varrho_p(m)$$

и

$$v^2(m) = \sum_p \varrho_p(m) + 2 \sum_{p < q} \varrho_p(m) \varrho_q(m)$$

($\varrho_p^2 = \varrho_p$); следовательно,

$$\sum_{m=1}^n v(m) = \sum_p \left[\frac{n}{p} \right] \quad (4.7)$$

и

$$\sum_{m=1}^n v^2(m) = \sum_p \left[\frac{n}{p} \right] + 2 \sum_{p < q} \left[\frac{n}{pq} \right]. \quad (4.8)$$

Суммирование в (4.7) и (4.8) происходит только по простым p и q , которые меньше или равны n . Таким образом,

$$\sum_{m=1}^n v(m) \geq n \sum_{p \leq n} \frac{1}{p} - \pi(n), \quad (4.9)$$

где $\pi(n)$ обозначает количество простых чисел, не превосходящих n ; аналогично

$$\begin{aligned} \sum_{m=1}^n v^2(m) &\leq n \sum_{p \leq n} \frac{1}{p} + 2n \sum_{p < q \leq n} \frac{1}{pq} < \\ &< n \sum_{p \leq n} \frac{1}{p} + n \left(\sum_{p \leq n} \frac{1}{p} \right)^2. \end{aligned} \quad (4.10)$$

Известно, что

$$\sum_{p \leq n} \frac{1}{p} = \log \log n + e_n, \quad (4.11)$$

где e_n ограничено, и потому

$$\begin{aligned} \sum_{m=1}^n v^2(m) &\leq n (\log \log n)^2 + 2n \log \log n e_n + \\ &+ ne_n^2 + n \log \log n + ne_n \end{aligned}$$

и

$$\sum_{m=1}^n v(m) \geq n \log \log n + ne_n - \pi(n).$$

Наконец, с помощью (4.6) получаем

$$\begin{aligned} \sum_{m=1}^n (v(m) - \log \log n)^2 &\leq ne_n^2 + \\ &+ n \log \log n + ne_n + 2 \log \log n \pi(n) \end{aligned}$$

и, следовательно,

$$\frac{K_n}{n} \leq \frac{1}{g_n^2} + \frac{e_n^2}{g_n^2 \log \log n} + \frac{e_n}{g_n^2 \log \log n} + 2 \frac{\pi(n)}{n} \frac{1}{g_n^2}.$$

Так как e_n ограничено, $\pi(n) < n$ и $g_n \rightarrow \infty$, то отсюда следует, что

$$\lim_{n \rightarrow \infty} \frac{K_n}{n} = 0. \quad (4.12)$$

Ввиду того что величина $\log \log m$ меняется очень медленно, полученный результат (4.12) влечет за собой следующее.

Если l_n обозначает количество целых чисел m , $1 \leq m \leq n$, для которых имеет место или

$$v(m) < \log \log m - g_m \sqrt{\log \log m} \quad (4.13)$$

или

$$v(m) > \log \log m + g_m \sqrt{\log \log m},$$

то

$$\lim_{n \rightarrow \infty} \frac{l_n}{n} = 0. \quad (4.14)$$

Доказательство предоставляется читателю (см. задачу 1 в конце этого раздела). Утверждение (4.14) было впервые доказано Харди и Рамануджаном в 1917 г. Именно они сформулировали его очень образно, в такой форме: почти каждое целое m имеет приближенно $\log \log m$ простых делителей. Доказательство, приведенное выше, предложено П. Тураном, и оно гораздо проще первоначального доказательства Харди — Рамануджана. Как читатель может заметить, прием Турана является прямым аналогом доказательства слабого закона больших чисел, которое мы дали в п. 1 гл. 2. Это еще

один пример, когда идеи, заимствованные из одной области, приводят к плодотворным применениям в другой.

ЗАДАЧИ

1. Доказать (4.14). *Указание:* пусть $0 < \alpha < 1$; рассмотреть лишь целые числа из отрезка $n^\alpha \leq m \leq n$ и показать, что в этих пределах каждое целое m , удовлетворяющее

$$|v(m) - \log \log m| > g_m \sqrt{\log \log m},$$

удовлетворяет также неравенству

$$|v(m) - \log \log n| > h_n \sqrt{\log \log n}$$

при соответствующе выбранном $h_n \rightarrow \infty$.

2. Доказать (4.12) для $\omega(m)$.

5. Нормальный закон в теории чисел. То обстоятельство, что $v(m)$, число простых делителей m , равно сумме

$$\sum_p \varrho_p(m) \tag{5.1}$$

независимых функций, подсказывает нам, что распределение величин $v(m)$ может быть дано нормальным законом. Это действительно имеет место, и в 1939 г. Эрдёш и Кац доказали следующую теорему.

Пусть $K_n(\omega_1, \omega_2)$ — количество целых чисел m , $1 \leq m \leq n$, для которых

$$\begin{aligned} \log \log n + \omega_1 \sqrt{\log \log n} < \\ < v(m) < \log \log n + \omega_2 \sqrt{\log \log n}. \end{aligned} \tag{5.2}$$

Тогда

$$\lim_{n \rightarrow \infty} \frac{K_n(\omega_1, \omega_2)}{n} = \frac{1}{\sqrt{2\pi}} \int_{\omega_1}^{\omega_2} e^{-y^2/2} dy. \quad (5.3)$$

Ввиду того, что величина $\log \log n$ изменяется медленно (см. задачу 1 в конце п. 4), результат (5.3) эквивалентен утверждению

$$D \{ \log \log n + \omega_1 \sqrt{\log \log n} < v(n) < \log \log n + \omega_2 \sqrt{\log \log n} \} = \frac{1}{\sqrt{2\pi}} \int_{\omega_1}^{\omega_2} e^{-y^2/2} dy. \quad (5.4)$$

Теперь известно несколько различных доказательств этого результата (наилучшим, по моему мнению, является недавнее доказательство, принадлежащее Реньи и Турану), но ни одно из них, к сожалению, из-за недостаточной краткости и элементарности не может быть воспроизведено здесь. Поэтому мы будем довольствоваться эвристическими аргументами, основанными на следующей классической теореме Ландау.

Если $\pi_k(n)$ обозначает количество целых чисел, не превосходящих числа n , которое имеет ровно k простых делителей, то

$$\pi_k(n) \sim \frac{1}{(k-1)!} \frac{n}{\log n} (\log \log n)^{k-1}. \quad (5.5)$$

При $k=1$ это известный закон простых чисел: при $k > 1$ (5.5) может быть выведено из закона простых чисел с помощью совершенно элементарных соображений.

Имеем

$$K_n(\omega_1, \omega_2) = \sum_{\log \log n + \omega_1 \sqrt{\log \log n} < k < \log \log n + \omega_2 \sqrt{\log \log n}} \pi_k(n), \quad (5.6)$$

и, следовательно, можно предполагать, что

$$\begin{aligned} \frac{K_n(\omega_1, \omega_2)}{n} &\sim \\ &\sim \frac{1}{\log n} \sum_{\log \log n + \omega_1 \sqrt{\log \log n} < k < \log \log n + \omega_2 \sqrt{\log \log n}} \frac{(\log \log n)^{k-1}}{(k-1)!}. \end{aligned} \quad (5.7)$$

Если мы вспомним задачу 2 (п. 3 гл. 3) и положим

$$x = \log \log n \quad \left(e^{-x} = \frac{1}{\log n} \right), \quad (5.8)$$

то получим

$$\frac{K_n(\omega_1, \omega_2)}{n} \sim \frac{1}{\sqrt{2\pi}} \int_{\omega_1}^{\omega_2} e^{-y^2/2} dy,$$

или (5.3).

К несчастью, трудно сделать строгими эти весьма привлекательные доводы, так как необходимы равномерные оценки остаточного члена в теореме Ландау (5.5), а их нелегко получить. Интересно то, что пер-

воначальное доказательство Харди и Рамануджана теоремы п. 4 существенно опиралось на (5.5), хотя они нуждались лишь в оценках, а не в точном асимптотическом результате. Теория, развитая в гл. 3, подсказывает метод доказательства (5.3). Пусть $K_n(\omega)$ обозначает количество целых m , $1 \leq m \leq n$, для которых

$$v(m) < \log \log n + \omega \sqrt{\log \log n}.$$

Положим

$$\sigma_n(\omega) = \frac{K_n(\omega)}{n}. \quad (5.9)$$

Ясно, что $\sigma_n(\omega)$ является функцией распределения и что выполняется

$$\frac{1}{n \log \log n} \sum_{m=1}^n (v(m) - \log \log n)^2 = \int_{-\infty}^{\infty} \omega^2 d\sigma_n(\omega). \quad (5.10)$$

Если мы используем точную оценку

$$\sum_{p \leq n} \frac{1}{p} = \log \log n + C + \varepsilon_n, \quad \varepsilon_n \rightarrow 0, \quad (5.11)$$

то в силу аргументов п. 4

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} \omega^2 d\sigma_n(\omega) = 1 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^2 e^{-y^2/2} dy. \quad (5.12)$$

Мы имеем (почти тривиально!)

$$\lim_{n \rightarrow \infty} \frac{1}{n \sqrt{\log \log n}} \sum_{m=1}^n (v(m) - \log \log n) = 0$$

и потому

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} \omega d\sigma_n(\omega) = 0 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} ye^{-y^2/2} dy. \quad (5.13)$$

Если бы мы смогли доказать, что для каждого целого $k > 2$

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} \omega^k d\sigma_n(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^k e^{-y^2/2} dy, \quad (5.14)$$

то получили бы, что

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} e^{i\xi\omega} d\sigma_n(\omega) = e^{-\xi^2/2}$$

для каждого действительного ξ и, следовательно,

$$\lim_{n \rightarrow \infty} \sigma_n(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\omega} e^{-y^2/2} dy. \quad (5.15)$$

Это в силу обозначения (5.9) — не что иное, как наша теорема (5.3). Доказательство (5.14), безусловно, эквивалентно доказательству равенства

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n (\log \log n)^{k/2}} \sum_{m=1}^n (v(m) - \log \log n)^k &= \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^k e^{-y^2/2} dy, \end{aligned} \quad (5.16)$$

а это в свою очередь зависит от асимптотических оценок сумм

$$\sum_{p_{l_1} \dots p_{l_k} < n} \frac{1}{p_{l_1} \dots p_{l_k}}.$$

(Вспомните, что в п. 4 доказательство Турана было связано с оценкой

$$\sum_{pq \leq n} \frac{1}{pq}.)$$

Этот довольно заманчивый путь оказывается совсем не легким, однако недавно Халберстаму удалось получить доказательство этим методом. Данный подход, без сомнения, наиболее понятен и ближе всего по духу традиционным направлениям теории вероятностей. Наивысшим торжеством вероятностных методов в теории чисел явилось данное Реньи и Тураном доказательство того, что остаточный член

$$\frac{K_n(\omega)}{n} - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\omega} e^{-y^2/2} dy$$

имеет порядок

$$\frac{1}{\sqrt{\log \log n}}.$$

То, что ошибка имеет порядок $(\log \log n)^{-1/2}$, было предположено Леваком по аналогии с подобной оценкой в теории вероятностей — простые числа действительно играют в азартную игру!

ЗАДАЧИ

1. Показать, что (5.4) выполняется, если $\nu(n)$ заменить на $\omega(n)$ (число простых делителей с учетом кратности). (Указание: из того, что $M\{\omega(n) - \nu(n)\} < \infty$, вывести сначала, что плотность множества целых чисел, для которых $\omega(n) - \nu(n) > g_n$, $g_n \rightarrow \infty$, равна 0.)

2. Пусть $d(n)$ обозначает число делителей n .

(а) Показать, что

$$d(n) = \prod_p (\alpha_p(n) + 1).$$

(Определение $\alpha_p(n)$ см. в задаче 2 п. 3 этой главы.)

(б) Показать, что

$$M \left\{ \frac{d(n)}{2^{\nu(n)}} \right\} = \prod_p \left(1 + \frac{1}{2^p(p-1)} \right) < \infty.$$

(в) Используя (5.4) и указание к задаче 1, помещенное выше, доказать, что

$$\begin{aligned} D \left\{ 2^{\log \log n + \omega_1} \sqrt{\log \log n} < d(n) < 2^{\log \log n + \omega_2} \sqrt{\log \log n} \right\} = \\ = \frac{1}{\sqrt{2\pi}} \int_{\omega_1}^{\omega_2} e^{-y^2/2} dy. \end{aligned}$$

ЛИТЕРАТУРА

Ссылки на работы Давенпорта, Эрдёша, Эрдёша и Каца, Халберстама, Шёнберга и Турана см. в обзорных статьях: К а с М., Probability methods in some problems of analysis and number theory, *Bull. Amer. Math. Soc.*, 55 (1949), 641—665.

- К у б и л ю с И. П., Вероятностные методы в теории чисел,
Успехи матем. наук, 68 (1956), 31—66.
- R é n y i А., On the density of certain sequences of integers,
Publ. Inst. Math. Belgrade, 8 (1955), 157—162.
- R é n y i А., Turán P., On a theorem of Erdős—Kac, *Acta
Arith.*, 4 (1958), 71—84.

ОТ КИНЕТИЧЕСКОЙ ТЕОРИИ К НЕПРЕРЫВНЫМ ДРОБЯМ

1. Парадоксы кинетической теории. В середине девятнадцатого столетия были предприняты попытки объединения дисциплин механики и термодинамики.

Основной была проблема получения второго закона термодинамики из представлений о веществе, как состоящем из частиц (атомов и молекул), которые подвержены действию сил и подчиняются законам механики.

В работах Максвелла и Больцмана (и позднее Дж. У. Гиббса) при помощи этого кинетического подхода было получено одно из самых замечательных и многообещающих достижений науки.

Однако вначале этот подход был поколеблен двумя парадоксами. Первый, высказанный в 1876 г. Лошмидтом, заключался в том, что законы механики обратимы во времени (т. е. инвариантны относительно замены t на $-t$).

С другой стороны, второй закон термодинамики постулирует типично необратимый процесс.

Таким образом, казалось невозможным когда-либо получить второй закон, исходя из чисто механистических соображений.

Второй парадокс, связанный с именем Цермело, является даже более решающим.

Цермело использовал простую, но фундаментальную теорему Пуанкаре о том, что консервативная динамическая система, удовлетворяющая некоторым широким условиям, имеет следующее свойство: «почти каждое» (в некотором специальном смысле, поясняемом ниже) первоначальное положение системы обязательно будет повторено с любой степенью точности.

Это также противоречит необратимому поведению.

Чтобы полностью понять эти парадоксы, рассмотрим два резервуара, один с газом, а другой пустой.

Через некоторое время соединим резервуары. Второй закон предсказывает, что газ будет переходить из первого резервуара во второй и что количество газа в первом резервуаре будет *уменьшаться монотонно по времени*. Такое поведение газа устанавливает определенное *направление времени*.

С кинетической (механистической) точки зрения мы имеем дело с динамической системой, которая неспособна выявить направление времени и которая, кроме того, будет вести себя квазипериодически, как указывается теоремой Пуанкаре. Очевидно, что мы имеем здесь парадоксальное явление.

2. Предварительные сведения. Чтобы понять ответ Больцмана, потребуется небольшой экскурс в классическую механику.

Система с n степенями свободы описывается в терминах n обобщенных координат q_1, q_2, \dots, q_n и соот-

ветствующих им импульсов p_1, p_2, \dots, p_n . Для консервативной динамической системы существует функция $H(q_1, \dots, q_n; p_1, \dots, p_n)$, называемая функцией Гамильтона, которая описывает полную энергию системы.

Уравнения движения записываются в форме

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}, \quad i = 1, 2, \dots, n, \quad (2.1)$$

$$\frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}, \quad i = 1, 2, \dots, n. \quad (2.2)$$

Причем если мы знаем начальное положение $q_i(0)$ и начальные импульсы $p_i(0)$, то движение [т. е. функции $q_i(t)$ и $p_i(t)$] определяется однозначно.

Обычно принято изображать систему в виде точки $2n$ -мерного евклидова пространства (фазового или Γ -пространства), имеющей координаты $q_1, \dots, q_n, p_1, \dots, p_n$.

Таким образом, в момент времени t динамическая система изображается точкой

$$P_t = (q_1(t), \dots, q_n(t), p_1(t), \dots, p_n(t)).$$

Теперь движение нашей системы с помощью соотношения

$$T_t(P_0) = P_t \quad (2.3)$$

определяет однопараметрическое семейство отображений.

Допустим, что мы имеем множество A точек P_0 , и обозначим $T_t(A)$ множество соответствующих точек P_t .

Лиувилль заметил (доказательство крайне просто и может быть основано на обобщении для $2n$ -мерного пространства хорошо известной теоремы о дивергенции), что из гамильтоновых уравнений движения (2.1) и (2.2) следует замечательный факт: меры Лебега множества A и $T_t(A)$ в $2n$ -мерном пространстве равны!

Другими словами, отображения T_t сохраняют меру, обычную меру Лебега в Γ -пространстве.

Уравнения (2.1) и (2.2) имеют еще одно важное следствие, а именно что

$$\begin{aligned} H(q_1(t), \dots, q_n(t), p_1(t), \dots, p_n(t)) = \\ = H(q_1(0), \dots, q_n(0), p_1(0), \dots, p_n(0)) \end{aligned}$$

(сохранение энергии) и, следовательно, точка, изображающая нашу динамическую систему, обязана лежать на «поверхности постоянной энергии» Ω

$$H(q_1, \dots, q_n, p_1, \dots, p_n) = \text{const.} \quad (2.4)$$

Предположим, что поверхность постоянной энергии является компактом и достаточно «регулярна», так что применима элементарная теория поверхностных интегралов. Предположим также, что на Ω

$$\|\nabla H\|^2 = \sum_{i=1}^n \left(\frac{\partial H}{\partial p_i} \right)^2 + \left(\frac{\partial H}{\partial q_i} \right)^2 > c > 0. \quad (2.5)$$

Пусть $B \subset \Omega$ — множество на поверхности энергии такое, что интеграл

$$\int_B \frac{d\sigma}{\|\nabla H\|}$$

($d\sigma$ — элемент поверхности) определен. Мы определим меру множества B формулой

$$\mu\{B\} = \frac{\int_B \frac{d\sigma}{\|\nabla H\|}}{\int_{\Omega} \frac{d\sigma}{\|\nabla H\|}} \quad (2.6)$$

так, что

$$\mu\{\Omega\} = 1. \quad (2.7)$$

Теперь из вышеприведенной теоремы Лиувилля и простых геометрических соображений следует, что

$$\mu\{T_t(B)\} = \mu\{B\}. \quad (2.8)$$

Другими словами, T_t сохраняет меру μ на Ω .

Формула (2.6) приписывает меру только некоторым элементарным множествам (к которым применима элементарная теория интегрирования по поверхности). Однако мера может быть продолжена на более широкий класс множеств таким же образом, как, начиная от интервалов на действительной прямой и *полагая* меру интервала равной его длине, можно построить вполне аддитивную меру Лебега.

В частности, множество C имеет μ -меру 0, если при любом $\varepsilon > 0$ существует конечная или счетная

совокупность элементарных множеств B_i , такая, что

$$C \subset \bigcup_i B_i \text{ и } \sum_i \mu \{B_i\} < \varepsilon.$$

Мы можем теперь сформулировать в точных терминах теорему Пуанкаре, использованную Цермело.

Если B μ -измеримо, то почти все $P_0 \in B$ (т. е. исключая множество μ -меры 0) таковы, что для некоторого t (зависящего, возможно, от P_0) $T_t(P_0) \in B$.

3. Ответ Больцмана. Чтобы понять ответ Больцмана, вернемся назад к нашему примеру с двумя резервуарами. Допустим, что нам известен точный вид функции Гамильтона

$$H(q_1, \dots, q_n, p_1, \dots, p_n) \quad (3.1)$$

и ее значение C при $t=0$. Таким образом, мы знаем поверхность постоянной энергии (ее уравнение $H=C$).

Существует, очевидно, множество B точек из Ω , соответствующее условию, что при $t=0$ все частицы находятся в одном из двух резервуаров, и нам известно, что наша система начинает движение из множества B .

Первое утверждение Больцмана о том, что μ -мера $\mu\{B\}$ множества B «крайне» мала, соответствует нашему интуитивному представлению о том, что мы начинаем из весьма необыкновенного или редкого состояния. С другой стороны, множество R точек Ω , соответствующее состояниям, при которых число частиц в двух резервуарах «почти пропорционально» объемам этих резервуаров, таково, что мера $\mu\{R\}$ «крайне» близка к 1.

Конечно, эти утверждения в большой степени зависят от значения выражений «крайне» и «почти», но достаточно будет сказать, что вследствие необъятности числа атомов в кубическом сантиметре (порядка 10^{20}) вполне надежно интерпретировать «крайне», как меньше, чем 10^{-10} , и «почти пропорционально», как то, что отклонение от соответствующего отношения меньше 10^{-10} .

Второе утверждение было гораздо смелее. Больцман считал, что первое утверждение *влечет* за собой следующее: относительное время, которое фактическая кривая, описывающая движение системы, находится в B и R соответственно, «крайне» мало и «крайне» велико.

Другими словами, если система находится в необыкновенном состоянии, то почти немедленно оставит его (в то время как по теореме Пуанкаре она в конце концов почти наверное вернется обратно), а если она находится в множестве, соответствующем «почти нормальному» состоянию, то она останется здесь «в сущности» навсегда.

Больцман получил первое утверждение правдоподобными, но не очень строгими оценками. Для обоснования второго утверждения он высказал *гипотезу* о том, что кривая, изображающая движение системы, проходит через *каждую* точку поверхности постоянной энергии.

Эта гипотеза, которую Больцман назвал *эргодической* (Ergodenhypothese), ошибочна (исключая случай $n = 1$, когда она тривиальна).

Больцман старался спасти свое объяснение, заменив неверную эргодическую гипотезу тем, что он назвал «квазиэргодической гипотезой». Новая гипотеза постулировала, что кривая движения подходит произвольно близко к каждой точке на поверхности постоянной энергии. Это хотя и очень правдоподобно, но недостаточно для установления связи между относительным временем, проведенным во множестве $A \subset \Omega$, и его μ -мерой $\mu\{A\}$.

Понятно, что суть дела заключается в связи между относительным временем, проведенным в A , и $\mu\{A\}$.

Но что мы подразумеваем под относительным временем, проведенным в A ? Определение подсказывается почти непосредственно. Пусть $t(\tau, P_0, A)$ обозначает время, которое кривая движения, начавшегося из P_0 , пробыла в A вплоть до момента времени τ . Относительное время есть тогда предел

$$\lim_{\tau \rightarrow \infty} \frac{t(\tau, P_0, A)}{\tau}, \quad (3.2)$$

если он, конечно, существует.

Оказывается, доказательство существования этого предела представляет реальное затруднение. Если же это сделано, то потребуется лишь дополнительное предположение относительно T_t , чтобы вывести, что предел равен $\mu\{A\}$.

4. Абстрактное изложение. После того как я столь подробно остановился на основах статистической механики, я оставлю в стороне большую ее часть и выделю из нее чисто математическое содержание.

Вместо поверхности постоянной энергии я возьму множество Ω (с полной мерой 1), на котором задана вполне аддитивная мера μ .

Я допускаю теперь, что существует однопараметрическое семейство отображений T_t множества Ω на себя, сохраняющее μ -меру. Это утверждение требует некоторого разъяснения. В динамике отображения T_t являются взаимно однозначными (это непосредственное следствие единственности решения гамильтоновых уравнений движения). Однако нет необходимости предполагать отображения T_t взаимно однозначными, если надлежащим образом определить, что значит «сохранять меру».

Соответствующее определение выглядит следующим образом. Пусть $T_t^{-1}(A)$ — прообраз множества A , т. е.

$$T_t(T_t^{-1}(A)) = A. \quad (4.1)$$

Отображение T_t называется сохраняющим меру, если

$$\mu\{T_t^{-1}(A)\} = \mu\{A\}. \quad (4.2)$$

Понятно, что для взаимно однозначных отображений соотношение (4.2) эквивалентно обычному определению инвариантности меры, т. е.

$$\mu\{T_t(A)\} = \mu\{A\}. \quad (4.3)$$

Пусть теперь $P_0 \in \Omega$, и $g(P)$ обозначает характеристическую функцию измеримого множества A , т. е.

$$g(P) = \begin{cases} 1, & P \in A, \\ 0, & P \notin A. \end{cases} \quad (4.4)$$

Теперь ясно, что $t(\tau, P_0, A)$ дается формулой

$$t(\tau, P_0, A) = \int_0^\tau g(T_t(P_0)) dt, \quad (4.5)$$

и проблема заключается в доказательстве существования предела

$$\lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau g(T_t(P_0)) dt. \quad (4.6)$$

Вместе с этим вариантом, где время изменяется непрерывно, удобно рассмотреть дискретный случай. Пусть T — сохраняющее меру отображение, т. е.

$$\mu\{T^{-1}(A)\} = \mu\{A\}. \quad (4.7)$$

Рассмотрим его степени (повторные отображения) T^2, T^3, \dots .

Аналогом предела (4.6) служит теперь

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n g(T^k(P_0)). \quad (4.8)$$

В 1931 г. Биркгофу удалось доказать, что пределы (4.6) и (4.8) существуют для почти всех P_0 (в смысле μ -меры). Несколько ранее фон Нейман показал, что пределы (4.6) и (4.8) существуют в смысле среднего квадратического.

Сейчас имеются разнообразные доказательства этих теорем, кратчайшее из них принадлежит Ф. Риссу. Мы опустим доказательство, отослав читателя к пре-

восходной брошюре Халмоша «Лекции по эргодической теории».

Что можно сказать относительно предела (4.8) [или (4.6)]?

Обозначая этот предел через $h(P_0)$, мы непосредственно видим, что эта функция μ -измерима, ограниченная (действительно, $0 \leq h(P_0) \leq 1$) и такая, что для почти каждого P_0

$$h(T(P_0)) = h(P_0). \quad (4.9)$$

Пусть теперь H_α — множество тех точек P_0 , для которых

$$h(P_0) < \alpha,$$

и пусть $Q \in T^{-1}(H_\alpha)$. Тогда $T(Q) \in H_\alpha$ и, следовательно

$$h(T(Q)) < \alpha.$$

Так как для почти каждого Q $h(T(Q)) = h(Q)$, то для всех Q , за исключением множества меры 0, выполняется $h(Q) < \alpha$. Поэтому, исключая множество μ -меры 0, имеем

$$T^{-1}(H_\alpha) = H_\alpha$$

для каждого α (исключительное множество может, конечно, зависеть от α).

Другими словами, множества H_α инвариантны (с точностью до множеств меры 0).

Отображение называется «метрически транзитивным», если инвариантными являются единственно множества или нулевой, или единичной меры.

Если мы предположим, что наше отображение метрически транзитивно, то все множества H_α будут иметь меру или 0, или 1 и, следовательно, $h(P_0)$ постоянна почти всюду.

Значение этой постоянной легко определяется, если заметить, что из равенства

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n g(T^k(P_0)) = h(P_0) \quad (\text{почти всюду})$$

следует (по теореме об ограниченной сходимости)

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \int_{\Omega} g(T^k(P_0)) d\mu = \int_{\Omega} h(P_0) d\mu. \quad (4.10)$$

Действительно,

$$\int_{\Omega} g(T^k(P_0)) d\mu = \int_{\Omega} g(P_0) d\mu = \mu\{A\}$$

(это прямое следствие того, что T сохраняет меру), и потому

$$\int_{\Omega} h(P_0) d\mu = \mu\{A\}.$$

Таким образом, константа равна $\mu\{A\}$.

Объединяя все это, мы можем сказать, что если T метрически транзитивно, то для почти всех P_0

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n g(T^k(P_0)) = \mu\{A\}. \quad (4.11)$$

Сказанное легко обобщается следующим образом. Если $f(P_0)$ μ -интегрируема, т. е.

$$\int_{\Omega} |f(P_0)| d\mu < \infty,$$

и если T метрически транзитивно, то для почти всех P_0

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n f(T^k(P_0)) = \int_{\Omega} f(P_0) d\mu. \quad (4.12)$$

Можно подумать, что доказанная теорема (4.12) полностью реабилитирует точку зрения Больцмана. К сожалению, отображения T_t , к которым мы пришли в динамике, столь сложны, что, за исключением некоторых очень простых случаев, не известно, будут ли они метрически транзитивными или нет. Это, однако, никоим образом не умаляет красоты и значительности эргодической теоремы (4.12).

5. Эргодическая теорема и непрерывные дроби. Пусть x , $0 < x \leq 1$, — действительное число. Представим его простой непрерывной дробью

$$x = \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \dots}}}, \quad (5.1)$$

где a_1, a_2, \dots — положительные целые числа. Для величин a легко вывести формулы.

Мы имеем

$$a_1(x) = \left[\frac{1}{x} \right], \quad a_2(x) = \left[\frac{1}{\frac{1}{x} - \left[\frac{1}{x} \right]} \right], \dots,$$

где $[y]$, как обычно, обозначает наибольшее целое число, меньшее или равное y .

Формулы для a постепенно все более и более усложняются, однако небольшое размышление обнаруживает, что их вычисление может быть подогнано под следующую схему.

Пусть

$$T(x) = \frac{1}{x} - \left[\frac{1}{x} \right], \quad (5.2)$$

тогда

$$a_2(x) = a_1(T(x)), \quad (5.3)$$

$$a_3(x) = a_2(T(x)) = a_1(T^2(x)) \quad (5.4)$$

и т. д.

Возможность применения эргодической теоремы становится теперь очевидной, так как мы здесь получили степени отображения $T(x)$, заданного (5.2).

Что служит здесь пространством Ω ? Очевидно, интервал $(0,1)$ с исключенной точкой 0.

Что служит инвариантной мерой? На это ответить труднее, однако ответ в сущности был найден уже Гауссом.

Можно поступить следующим образом: пусть $\varrho(x)$, $0 < x \leq 1$, таково, что

$$(a) \varrho(x) \geq 0, \quad (b) \int_0^1 \varrho(x) dx = 1. \quad (5.5)$$

Определим $\mu\{A\}$ формулой

$$\mu\{A\} = \int_A \varrho(x) dx. \quad (5.6)$$

Возьмем теперь интервал (α, β) , $0 < \alpha < \beta < 1$, и рассмотрим его прообраз при отображении $T(x)$.

Имеем

$$T^{-1}(\alpha, \beta) = \bigcup_{k=1}^{\infty} \left(\frac{1}{k+\beta}, \frac{1}{k+\alpha} \right) \quad (5.7)$$

и, следовательно,

$$\mu\{T^{-1}(\alpha, \beta)\} = \sum_{k=1}^{\infty} \int_{1/(k+\beta)}^{1/(k+\alpha)} \varrho(x) dx. \quad (5.8)$$

Если μ сохраняется, то мы должны получить

$$\int_{\alpha}^{\beta} \varrho(x) dx = \sum_{k=1}^{\infty} \int_{1/(k+\beta)}^{1/(k+\alpha)} \varrho(x) dx \quad (5.9)$$

при всех α и β .

Мы не знаем, как «планомерно» подойти к решению уравнения (5.9). Однако легко проверить, что

$$\varrho(x) = \frac{1}{\log 2} \frac{1}{1+x} \quad (5.10)$$

является решением и удовлетворяет условиям (5.5).

Это все, что нам требовалось, исключая проверку метрической транзитивности $T(x)$, а это совершенно тривиально.

Если $f(x)$ μ -интегрируема, т. е.

$$\frac{1}{\log 2} \int_0^1 |f(x)| \frac{dx}{1+x} < \infty, \quad (5.11)$$

то по (4.12)

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^n f(T^k(x)) = \frac{1}{\log 2} \int_0^1 f(x) \frac{dx}{1+x} \quad (5.12)$$

для почти каждого x (отметим, что множества μ -меры 0 совпадают с множествами обычной лебеговой меры 0).

Пусть теперь

$$f(x) = \log a_1(x). \quad (5.13)$$

Из (5.12) получаем, что для почти всех x

$$\lim_{n \rightarrow \infty} (a_1 a_2 \dots a_n)^{1/n} = C, \quad (5.14)$$

где

$$\begin{aligned} C &= \exp \left(\frac{1}{\log 2} \int_0^1 \log a_1(x) \frac{dx}{1+x} \right) = \\ &= \exp \left(\frac{1}{\log 2} \sum_{k=1}^{\infty} \log k \log \frac{(k+1)^2}{k(k+2)} \right). \end{aligned} \quad (5.15)$$

Эта замечательная теорема была впервые доказана (другим путем) Хинчиным в 1935 г. Настоящее доказательство предложено К. Рыл-Нарджевским.

Я легко бы мог освободить читателя от первых трех пунктов этой главы. Я мог бы начать с абстрактного изложения п. 4 и тем самым избежать упоминания о динамике и кинетической теории.

Но, сделав это, я бы уничтожил наиболее восхитительную и, по моему мнению, наиболее поучительную часть рассказа, так как путь от кинетической теории, как представлял ее себе Больцман и другие, к непрерывным дробям являет собой великолепный пример часто забываемого факта, что математика не есть нечто обособленное, но что она своей мощью и красотой в значительной степени обязана другим дисциплинам.

ЗАДАЧИ

1. Пусть $B \subset \Omega$ μ -измеримо, и пусть $\mu\{B\} \neq 0$. Если T сохраняет меру (но не обязательно метрически транзитивно), то доказать, что для почти каждого $P_0 \in B$ существует целое $n \geq 1$, так что $T^n(P_0) \in B$. (Это дискретный вариант теоремы Пуанкаре; чтобы это показать, надо рассмотреть множество $C \subset B$, такое, что при $P_0 \in C$ $T^n(P_0) \notin B$ для $n=1, 2, \dots$. Доказать, что C μ -измеримо и множества $C, T^{-1}(C), T^{-2}(C), \dots$ все попарно не пересекаются.)

2. Пусть $n(P_0), P_0 \in B$, — первое положительное целое число, такое, что $T^{n(P_0)}(P_0) \in B$. Доказать, что если T (в добавление к свойству сохранять меру) метрически транзитивно, то

$$\int_B n(P_0) d\mu = 1.$$

3. Пусть x , $0 < x \leq 1$, представлено непрерывной дробью:

$$x = \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \dots}}},$$

и пусть B — множество, на котором $a_1(x) = k$ (т. е. $1/(k+1) < x \leq 1/k$). Пусть далее $n(x, k)$ обозначает наименьшее целое число, большее 1 и такое, что $a_n(x, k) = k$.

Показать, что

$$\frac{1}{\log 2} \int_{1/k}^{1/(k+1)} (n(x, k) - 1) \frac{dx}{1+x} = 1.$$

4. Пусть $0 \leq x \leq 1$ и $T(x) = 2x - [2x]$. Применяя эргодическую теорему, получить теорему Бореля гл. 2.

ЛИТЕРАТУРА

1. Ryll-Nardzewski C., On the ergodic theorems II, *Studia Math.*, 12 (1951), 74—79.

ДОБАВЛЕНИЕ

Ю. В. Прохоров

Здесь собраны различные замечания к отдельным частям книги, которые могут оказаться полезными читателю, заинтересовавшемуся ее предметом. Литературные ссылки при этом ни в коем случае не претендуют на полноту.

К главе 1. Именно широкое использование понятия «независимости» выделяет теорию вероятностей из чистой теории меры (см. по этому поводу А. Н. Колмогоров, Основные понятия теории вероятностей, п. 5, гл. 1 или его же статью о теории вероятностей в книге «Математика, ее содержание, методы и значение», т. 2, М., 1956, 252 — 284).

К главе 2, п. 1. Неравенство Чебышева обладает большой общностью (например, оно использует лишь попарную независимость случайных величин — слагаемых), но дает слабую оценку для вероятностей событий типа (1.2). С. Н. Бернштейном и А. Н. Колмогоровым было показано, что полное использование факта независимости случайных величин позволяет получить значительно более сильные оценки (см., например, С. Н. Бернштейн, Курс теории вероятностей, изд. 4, М., 1946). В специальном случае величин r_k

рассуждение может быть проведено следующим образом. Во-первых (ср. 1.7), при любом действительном ξ

$$\int_0^1 \exp \xi [r_1(t) + \dots + r_n(t)] dt = \prod_{k=0}^n \int_0^1 e^{\xi r_k(t)} dt = (\operatorname{ch} \xi)^n,$$

(именно здесь используется факт независимости, точнее — его следствие: математическое ожидание произведения независимых случайных величин равно произведению их математических ожиданий). Далее (ср. 1.5) при $\xi > 0$

$$\begin{aligned} \mu [r_1(t) + \dots + r_n(t) > \varepsilon_n] &< e^{-\xi \varepsilon n} (\operatorname{ch} \xi)^n < \\ &< \exp n \left[-\xi \varepsilon + \frac{\xi^2}{2} \right]. \end{aligned}$$

Выбирая $\xi = \varepsilon$, мы придаем правой части последнего неравенства минимальное значение, равное $e^{-n\varepsilon^2}$ (см. задачу 12 в конце п. 2, гл. 2).

К п. 2. Теория вероятностей устанавливает ряд теорем о свойствах последовательностей

$$X_1(\omega), X_2(\omega), \dots, X_n(\omega), \dots \quad (1)$$

независимых случайных величин, имеющих место с вероятностью единица, т. е. для «почти всех» реализаций $x_1, x_2, \dots, x_n, \dots$ последовательности (1). Так, например, если $X_k = \varepsilon_k$, то для любой s -членной цепочки Δ_s , состоящей из нулей и единиц с вероятностью единица

$$F_n^{(\Delta_s)}(t) - \frac{1}{2^s} \cdot n = o(n) \quad (2)$$

и

$$\limsup_{n \rightarrow \infty} \frac{1}{\sqrt{c_s n \ln \ln n}} \left| F_n^{(\Delta_s)}(t) - \frac{1}{2^s} \cdot n \right| = 1, \quad (3)$$

где $F_n^{(\Delta_s)}(t)$ обозначает число появлений указанной цепочки в последовательности $\varepsilon_1(t), \dots, \varepsilon_{n+s-1}(t)$.

Многие математики, начиная с Бореля, занимались построением и исследованием числовых последовательностей, в том или ином смысле напоминающих «типичную реализацию» последовательности (1). В каком именно смысле — должно определяться используемыми методами и кругом возможных применений. Свойство (2) характеризует так называемые «нормальные» последовательности двоичных знаков. Примером может служить последовательность

$$0\ 1\ 00\ 01\ 10\ 11\ 000\ 001\ 010\ 011\ 100\dots,$$

в которой подряд выписываются все комбинации двоичных знаков в порядке возрастания их длины. Заметим, что проверка свойства (2) в этом и аналогичных примерах облегчается при использовании следующего критерия: если существует такая константа C , что при всех s и Δ_s

$$\limsup_{n \rightarrow \infty} \frac{F_n^{(\Delta_s)}}{n} < C \cdot \frac{1}{2^s},$$

то имеет место (2).

«Нормальные» последовательности знаков могут, однако, не обладать более тонкими свойствами последовательностей независимых случайных величин. Напри-

мер, доказано, что для любой функции $\psi(n) \uparrow \infty$ при $n \rightarrow \infty$ существует «нормальная» последовательность двоичных знаков, для которой

$$\limsup_{n \rightarrow \infty} \frac{1}{\psi(n)} \left| F_n^{(\Delta_s)} - \frac{1}{2^s} \cdot n \right| \leq 1.$$

Сведения о результатах, полученных в указанных направлениях, и литературные ссылки можно найти в работе А. Г. Постникова, Арифметическое моделирование случайных процессов, *Тр. Матем. ин-та АН СССР*, т. 57, 1960.

К п. 3. Уместно заметить, что значительная часть теории вероятностей, касающаяся распределений в пространствах конечного числа измерений и используемая, например, в классических вопросах математической статистики, может быть изложена без обращения к общей теории меры. Привлечение методов общей теории меры и интеграла Лебега становится неизбежным, когда рассматриваются распределения в бесконечных произведениях пространств.

К п. 4. Формулировки многих предельных теорем теории вероятностей начинаются обычно словами: рассмотрим последовательность независимых случайных величин X_k , имеющих функции распределения $F_k \dots$. Как отмечает автор, естественно возникает вопрос о существовании такой последовательности, т. е. требуется указать такую пару: пространство Ω и определенную на нем меру μ и такую последовательность функций $X_k(\omega)$, $\omega \in \Omega$, что $n \mu \{X_n(\omega) < \lambda_n\}$ при любом

равно $F_n(\lambda_n)$ и

$$\begin{aligned} \mu \{X_1(\omega) < \lambda_1, \dots, X_n(\omega) < \lambda_n\} = \\ = \mu \{X_1(\omega) < \lambda_1\} \dots \mu \{X_n(\omega) < \lambda_n\}. \end{aligned}$$

Вопрос о существовании подобных последовательностей решается положительно теоремой Колмогорова, упомянутой в тексте (Основные понятия теории вероятностей, § 4, гл. 3). При этом в качестве Ω берется пространство всех последовательностей $\omega = (x_1, x_2, \dots, x_n, \dots)$ действительных чисел. Любопытно отметить, что в качестве (Ω, μ) можно взять и отрезок $0 \leq t \leq 1$ с обычной лебеговой мерой. При этом функции $X_k(t)$ будут непременно разрывны (как и $r_k(t)$). Несколько сложнее обстоит дело с континуальными системами независимых случайных величин. Возьмем следующий пример: для описания процесса изменения координаты малой частицы, движущейся под влиянием молекулярных толчков (броуновское движение), теория вероятностей предлагает следующую модель: приращения координаты за непересекающиеся промежутки времени представляют собой независимые случайные величины с нормальным законом распределения. Математическое ожидание смещения за время Δt равно 0, а дисперсия $b \cdot \Delta t$. Предположим дополнительно, что $t \geq 0$ и что в начальный момент времени координата была равна 0. Из упомянутой теоремы Колмогорова вытекает существование соответствующих Ω и μ . При этом в качестве Ω берется множество всех действительных функций аргумента $t \geq 0$. Однако из физических соображений желательно, чтобы реализации процесса были непрерывными

функциями, т. е. чтобы в качестве Ω выступало множество всех непрерывных функций аргумента t . Дополнительное исследование показывает, что замена множества произвольных действительных функций множеством непрерывных функций возможна. При этом почти все реализации оказываются недифференцируемыми ни в одной точке (типа известной функции Вейерштрасса). См. Дж. Дуб, Вероятностные процессы, ИЛ, М., 1956.

К п. 5. С последовательностями независимых случайных величин связан один замечательный результат (так называемый закон нуля или единицы, он упоминается в тексте, в начале п. 6). Допустим, что некоторое событие таково, что его наступление определяется по значениям случайных величин X_k с $k \geq n$, каково бы ни было n . Примером может служить событие: ряд из случайных величин сходится. Тогда вероятность такого события равна нулю или единице. В частности, как показано в пп. 5 и 6,

ряд $\sum_1^{\infty} c_k r_k(t)$ сходится при почти всех t , если $\sum_1^{\infty} c_k^2 < \infty$

и расходится при почти всех t , если $\sum_1^{\infty} c_k^2 = \infty$. Необ-

ходимые и достаточные условия сходимости рядов из независимых случайных величин были найдены А. Я. Хинчиным и А. Н. Колмогоровым в 1925 г. Доказательство можно найти, например, в книге М. Лозева «Теория вероятностей», ИЛ, М., 1962. Типичными средствами доказательства в общем случае являются леммы,

подобные следующей: положим $s_0 = 0$, $s_n = \sum_1^n c_k r_k(t)$; тогда при любом $\varepsilon > 0$

$$\mu \{s_N \geq \varepsilon\} \leq \mu \left\{ \max_{1 \leq n \leq N} s_n \geq \varepsilon \right\} \leq 2\mu \{s_N \geq \varepsilon\}. \quad (4)$$

Грубо говоря, это означает, что колебания максимальной из нарастающих сумм независимых случайных величин имеют тот же порядок, что и колебания последней суммы. Левая часть неравенства (4) очевидна. Правая доказывается следующим образом. Обозначим рассматриваемые события буквами A и B , так что неравенство примет вид $\mu(A) \leq 2\mu(B)$. Пусть $k = 1, 2, \dots, n$ и

$$E_k = \{s_1(t) < \varepsilon, \dots, s_{k-1}(t) < \varepsilon, s_k(t) \geq \varepsilon\}.$$

Очевидно,

$$\begin{aligned} \mu(B) &= \sum_1^n \mu(B \cap E_k) \geq \\ &\geq \sum_1^n \mu \{E_k \cap (s_n(t) - s_k(t) \geq 0)\} = \\ &= \sum_1^n \mu(E_k) \cdot \mu(s_n(t) - s_k(t) \geq 0) \geq \frac{1}{2} \sum_1^n \mu(E_k) = \frac{1}{2} \mu(A), \end{aligned}$$

что и требовалось доказать.

Допустим теперь, что $\sum_1^{\infty} c_k^2 < \infty$. Применяя доказанное неравенство к $r_m(t)$, $r_{m+1}(t)$, ... и используя неравенство Чебышева, получим

$$\mu \left\{ \sup_{n \geq 1} \sum_m^{m+n} c_k r_k(t) \geq \varepsilon \right\} \leq \frac{1}{\varepsilon^2} \sum_m^{\infty} c_k^2 \rightarrow 0$$

при $m \rightarrow \infty$. Из этого неравенства и аналогичного, с заменой $\geq \varepsilon$ на $\leq -\varepsilon$, вытекает утверждение о сходимости ряда $\sum_1^{\infty} c_k r_k(t)$ с вероятностью единица.

Как отмечает автор, ряды по некоторым системам ортогональных функций аналогичны рядам из независимых случайных величин. Для произвольных ортонормальных систем

$$\int_{\Omega} X_i(\omega) \cdot X_j(\omega) d\mu = \delta_{ij}, \quad \mu(\Omega) = 1,$$

с $\int X_k(\omega) d\mu = 0$ верно следующее утверждение: из $\sum_1^{\infty} c_k^2 \log^2 k < \infty$ вытекает сходимость с вероятностью 1 ряда $\sum_1^{\infty} c_k X_k(\omega)$ (см. М. Лозэв, § 33).

В связи с некоторыми вопросами, возникающими, в частности, в теории случайных процессов, изучались

аналитические свойства сумм функциональных рядов со случайными коэффициентами. Например, функция

$$\sum_1^{\infty} \frac{r_k(t)}{k} z^k$$

будет с вероятностью единица непрерывна в замкнутом круге [ср. Г. А. Хант, Случайные преобразования Фурье, *Математика*, 2:6 (1958), 87—114].

К главе 3. Теоремы о сходимости распределений сумм случайных величин к нормальному закону объединяются в теории вероятностей под общим названием «центральной предельной теоремы». Название достаточно полно характеризует место, занимаемое этими теоремами. Возникновение нормального закона связывают обычно со схемой сложения большого числа «равномерно малых» случайных величин. В этой связи уместно упомянуть одну теорему А. Я. Хицчина. Рассмотрим последовательность серий

$$\begin{aligned} &X_{1,1}, \\ &X_{2,1}, X_{2,2}, \\ &\dots \\ &X_{n,1}, \dots, X_{n,n}, \end{aligned}$$

независимых внутри каждой серии случайных величин (в п. 1 и 2 можно принять $X_{n,k} = \frac{r_k(t)}{\sqrt{n}}$), подчиненных условию «предельной пренебрегаемости»:

при любом $\varepsilon > 0$ и $n \rightarrow \infty$

$$\max_{1 \leq k \leq n} \mu(|X_{n,k}(\omega)| \geq \varepsilon) \rightarrow 0. \quad (5)$$

Пусть распределения сумм $S_n = X_{n,1} + \dots + X_{n,n}$ сходятся к предельному. Тогда этот предельный закон будет нормальным в том и только том случае, когда при $n \rightarrow \infty$

$$\mu\left(\max_{1 \leq k \leq n} |X_{n,k}(\omega)| \geq \varepsilon\right) \rightarrow 0.$$

Последнее условие сильнее (5) и, как легко показать, равносильно следующему:

$$\sum_1^n \mu(|X_{n,k}(\omega)| \geq \varepsilon) \rightarrow 0 \text{ при } n \rightarrow \infty$$

(см. Б. В. Гнеденко и А. Н. Колмогоров, Предельные теоремы для сумм независимых случайных величин, М., 1949, п. 26).

К п. 4. Преобразования Фурье — Стильтьеса, упоминаемые в этом пункте, используются в теории вероятностей под названием характеристических функций. Детальное изложение свойств характеристических функций имеется, например, в книге М. Лозва. Заметим, что теорема непрерывности в несколько более узкой, чем приводимая автором формулировке, может быть доказана методом Маркова (именно, если предположить, что

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{+\infty} e^{itx} d\sigma_n(x) = c(t),$$

где

$$c(t) = \int_{-\infty}^{+\infty} e^{itx} d\sigma(x).$$

Ниже приводится вариант доказательства для этого случая, изложенный в вероятностных терминах. Пусть $\sigma_X(x)$ — функция распределения, $p_X(x)$ — плотность вероятности, $c_X(t)$ — характеристическая функция случайной величины X . Заметим, что

$$c_X(t) = \int_{-\infty}^{+\infty} e^{itx} p_X(x) dx$$

и, если $c_X(t)$ абсолютно интегрируема, то

$$p_X(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} c_X(t) dt$$

и

$$\sigma_X(x_2) - \sigma_X(x_1) = \int_{x_1}^{x_2} p_X(x) dx = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{e^{-itx_2} - e^{-itx_1}}{-it} c_X(t) dt.$$

Буквой Y с теми или иными индексами будут обозначаться случайные величины с $p_Y(x) = \frac{1}{2h} \left(1 - \frac{x}{2h}\right)$ при $|x| \leq 2h$ и $p_Y(x) = 0$ в других случаях. Тогда $|Y| = 2h$ и $c_Y(t) = \left(\frac{\sin th}{th}\right)^2$. Пусть теперь при $n \rightarrow \infty$

и каждом t функции $c_{X_n}(t) \rightarrow c_{X_0}(t)$ и Y_n не зависит от X_n ($n \geq 0$).

Мы имеем

$$c_{X_n+Y_n}(t) = c_{X_n}(t) \cdot c_{Y_n}(t) \rightarrow c_{X_0}(t) \cdot c_{Y_0}(t) = c_{X_0+Y_0}(t)$$

и по теореме об ограниченной сходимости

$$\sigma_{X_n+Y_n}(x_2) - \sigma_{X_n+Y_n}(x_1) \rightarrow \sigma_{X_0+Y_0}(x_2) - \sigma_{X_0+Y_0}(x_1).$$

Выберем теперь в качестве x_1 и x_2 ($x_1 < x_2$) точки непрерывности σ_{X_0} . Из неравенства

$$\sigma_{X_n}(x_2) - \sigma_{X_n}(x_1) \leq \sigma_{X_n+Y_n}(x_2+2h) - \sigma_{X_n+Y_n}(x_1-2h)$$

видим, что

$$\begin{aligned} \limsup_{n \rightarrow \infty} [\sigma_{X_n}(x_2) - \sigma_{X_n}(x_1)] &\leq \\ &\leq \sigma_{X_0+Y_0}(x_2+2h) - \sigma_{X_0+Y_0}(x_1-2h) \leq \\ &\leq \sigma_{X_0}(x_2+4h) - \sigma_{X_0}(x_1-4h) \rightarrow \sigma_{X_0}(x_2) - \sigma_{X_0}(x_1) \end{aligned}$$

при $h \rightarrow 0$. Отсюда и из обратного неравенства для \liminf вытекает, что

$$\sigma_{X_n}(x_2) - \sigma_{X_n}(x_1) \rightarrow \sigma_{X_0}(x_2) - \sigma_{X_0}(x_1).$$

Использованный здесь прием «сглаживания» распределений путем добавления к случайным величинам малых независимых «поправок» впервые был использован Ляпуновым при оценке остаточного члена в центральной предельной теореме.

К задаче 1 в п. 4. Утверждение является частным случаем следующей более общей теоремы (Фреше

и Шохат): пусть при каждом $k = 0, 1, 2, \dots$ и $n \rightarrow \infty$

$$\int_{-\infty}^{+\infty} x^k d\sigma_n(x) \rightarrow m_k.$$

Тогда существует такая функция $\sigma(x)$, что $m_k = \int_{-\infty}^{+\infty} x^k d\sigma(x)$. Если такая функция единственна, то

$$\sigma_n(x) \rightarrow \sigma(x)$$

в каждой точке непрерывности $\sigma(x)$. Достаточное

условие единственности: ряд $\sum_1^{\infty} m_k \frac{z^k}{k!}$ имеет ненулевой

радиус сходимости (см. М. Лозв).

К главе 4, п. 1. а). Интересующемуся читателю можно рекомендовать решение задач о распределении значений функций натурального аргумента из книги Г. Поля и Г. Сегё, Задачи и теоремы из анализа, ч. I, изд. 2, Гостехиздат, 1956 г.

б). Одно из наиболее полных изложений вопросов так называемой «вероятностной теории чисел» имеется в монографии И. П. Кубилюса «Вероятностные методы в теории чисел», 2 изд., Вильнюс, 1962. В частности, там приведены и замечательные результаты самого И. П. Кубилюса.

К п. 2. а). Для счетно-аддитивных мер имеется ряд теорем о переходе к пределу под знаком интеграла.

Например, если μ -почти всюду $f_n(\omega) \downarrow$ и $f_n \geq 0$, то

$$\lim_{n \rightarrow \infty} \int_{\Omega} f_n(\omega) d\mu = \int_{\Omega} \lim_{n \rightarrow \infty} f_n(\omega) d\mu.$$

Эти теоремы, вообще говоря, неприменимы в случае конечно-аддитивных мер ($\lim f_n$ может даже не быть измеримым!). Этим и объясняется, как подчеркивает автор, невозможность сразу получить равенства, подобные (2.10).

б). Непрерывность функции $\sigma(\omega) = D\left(\log \frac{\varphi(n)}{n} < \omega\right)$ (как и функции $\tau(\omega) = D\left(\frac{\sigma(n)}{n} < \omega\right)$) вытекает из одной общей теоремы Леви.

Пусть $X_1, X_2, \dots, X_n, \dots$ — последовательность независимых дискретно распределенных случайных величин со сходящейся с вероятностью 1 суммой

$$\sum_1^{\infty} X_k = X.$$

Пусть

$$d_k = \sup_x P\{X_k = x\}.$$

Функция распределения X будет непрерывна в том и только том случае, когда $\prod_k^{\infty} d_k = 0$. В случае функции $\varphi(n)$ можно принять

$$X_k = \begin{cases} 0 & \text{с вероятностью } 1 - \frac{1}{p}, \\ \log\left(1 - \frac{1}{p}\right) & \text{с вероятностью } \frac{1}{p}. \end{cases}$$

Остается принять во внимание хорошо известное соотношение

$$\prod_1^{\infty} \left(1 - \frac{1}{p}\right) = 0.$$

Можно добавить, что распределение X может быть только или чисто абсолютно непрерывным, или чисто дискретным, или чисто сингулярным (Иессен и Винтнер); этот результат и теоремы Леви приведены в обширном мемуаре Эссеена: С. G. Esseen. Fourier analysis of distribution functions. A mathematicae study of the Laplace — Gaussian law. *Acta. Math.*, 77 (1945), 1 — 125).

Существование распределения функций, подобных $\frac{\varphi(n)}{n}$ и $\frac{\sigma(n)}{n}$, может быть установлено единообразным методом, в основу которого может быть положено следующее вспомогательное замечание¹⁾.

Назовем функцию $f(n)$ периодической, если при некотором N и всех n

$$f(n + N) = f(n).$$

Легко установить наличие у всякой периодической функции распределения $D(f(n) < \omega)$.

1) Нижеследующие замечания обязаны моим беседам с А. Г. Постниковым по поводу работ Новоселова, в которых задачи о распределении значений определенного класса функций натурального аргумента изучаются с помощью топологизации множества натуральных чисел и пополнения его до топологически полного кольца. При этом плотность продолжается в (счетно-аддитивную) меру в указанном кольце.

Допустим теперь, что функция $f(n)$ такова, что при любом $\varepsilon > 0$ можно подобрать периодическую функцию $f_\varepsilon(n)$, для которой

$$\overline{\lim} \frac{1}{N} \sum_{n=1}^N |f(n) - f_\varepsilon(n)| \leq \varepsilon.$$

Тогда при любом ω , кроме, быть может, счетного числа, существует

$$D(f(n) < \omega).$$

Доказательство можно провести методом моментов, применяя его к

$$\tilde{f}(n) = \operatorname{arctg} f(n).$$

Для функций $f(n)$, удовлетворяющих условиям

$$f(n) = \sum_{d|n} \psi(d),$$

$$\sum_d \left| \frac{\psi(d)}{d} \right| < \infty$$

в качестве аппроксимирующих функций можно брать

$$f_\varepsilon(n) = \sum_{\substack{d=p_1^{\alpha_1} \dots p_k^{\alpha_k} \\ \alpha_1, \dots, \alpha_k \leq s_k}} \psi(d).$$

где p_1, \dots, p_k — первые k простых чисел, а k и s_k выбираются в зависимости от ε .

К главе 5, п. 4 и 5. О редукции задач кинетической теории к задачам теории вероятностей см. А. Я. Хинчин, «Математические основания статистической механики», ГИТТЛ, М. — Л., 1943, см. также М. Кас, Probability and related topics in physical sciences. N. Y. 1959.

Для метрически транзитивных T мы имеем сходимость средних вдоль траекторий к среднему по пространству

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \xi_k(P_0) = a,$$

где $\xi_k(P_0) = f(T^k(P_0))$, $a = \int_{\Omega} f(P_0) d\mu$ (формула 4.12).

В ряде случаев удается доказать, что отклонение средних «по времени» от a подчиняется в пределе нормальному закону: если f не равна μ -почти всюду постоянной, то найдется $\sigma = \sigma(f) > 0$, такая, что

$$\begin{aligned} \mu \left(P_0 : \omega_1 < \frac{\xi_1(P_0) + \dots + \xi_n(P_0) - na}{\sigma \sqrt{n}} < \omega_2 \right) &\rightarrow \\ &\rightarrow \frac{1}{\sqrt{2\pi}} \int_{\omega_1}^{\omega_2} e^{-y^2/2} dy. \end{aligned}$$

$$\log \frac{1}{\xi(s)} = \sum \left[\log \left(1 - \frac{1}{p^s} \right) + \frac{1}{p^s} \right] + \sum \frac{1}{p^s}.$$

В частности, это справедливо для следующих динамических систем, встречающихся в теории чисел. В обоих случаях Ω — отрезок $(0, 1]$.

$$T_1(x) = \{2x\}$$

(фигурные скобки — знак «дробной части»)

$$T_2(x) = \frac{1}{x} - \left[\frac{1}{x} \right]$$

(последнее преобразование как раз и рассматривается в п. 5 гл. 5).

Результаты М. Каца (*Bull. Amer. math. Soc.*, 55 (1949), №7 и *Ann. Math.*, 47 (1946), №1, 33—49), относящиеся к этому случаю, были обобщены И. А. Ибрагимовым. [*Теория вероятностей и ее применения*, т. V, вып. 2 (1960), 256—257 и *ДАН СССР*, 125:4 (1959), 714—714.] Согласно И. А. Ибрагимову от $f(X)$, $0 < X \leq 1$, достаточно требовать выполнения двух свойств

$$\text{а) } f(x) \in L_2[0, 1]$$

$$\text{б) } \int_0^1 |f(x+h) - f(x)|^2 dt \leq C \frac{1}{\left(\log \frac{1}{h}\right)^{2+\varepsilon}}.$$

Некоторые вопросы метрической теории цепных дробей хорошо изложены в книге А. Я. Хинчина «Цепные дроби», Физматгиз М., 1961.

ОГЛАВЛЕНИЕ

Глава 1. ОТ ВЬЕТА К ПОНЯТИЮ СТАТИСТИЧЕСКОЙ НЕЗАВИСИМОСТИ	13
1. Формула Виета	13
2. Другой взгляд на формулу Виета	14
3. Случайность или начало чего-либо более глубокого?	17
4. $\left(\frac{1}{2}\right)^n = \frac{1}{2} \cdot \dots \cdot \frac{1}{2}$ (n раз)	19
5. Герб или решетка?	21
6. Независимость и «независимость» Задачи	24 26
Глава 2. БОРЕЛЬ И ПОСЛЕ НЕГО	29
1. «Законы больших чисел»	29
2. Борель и «нормальные числа» Задачи	32 36
3. «Герб или решетка» — более абстрактное изложение	40
4. В чем ценность абстракции?	43
5. Пример 1. Сходимость ряда со случайными знаками	45

6. Пример 2. Расходимость ряда со случайными знаками	53
Задачи	57
Литература	58

Глава 3. НОРМАЛЬНЫЙ ЗАКОН 60

1. Муавр	60
2. Основная идея метода	61
3. Метод Маркова становится строгим	63
Задачи	66
4. Более внимательный взгляд на метод	67
Задачи	70
5. Закон природы или математическая теорема?	73
Задачи	81
Литература	81

Глава 4. ПРОСТЫЕ ЧИСЛА «ИГРАЮТ В АЗАРТНУЮ ИГРУ» 82

1. Теоретико-числовые функции, плотность, независимость	82
2. Статистика значений φ -функций Эйлера	83
Задачи	94
3. Другое применение	97
4. Почти каждое целое m имеет приближенно $\log \log m$ простых делителей	106
Задачи	110
5. Нормальный закон в теории чисел	110
Задачи	116
Литература	116

Глава 5. ОТ КИНЕТИЧЕСКОЙ ТЕОРИИ К НЕПРЕРЫВНЫМ ДРОБЯМ	118
1. Парадоксы кинетической теории	118
2. Предварительные сведения	119
3. Ответ Больцмана	123
4. Абстрактное изложение	125
5. Эргодическая теорема и непрерывные дроби	130
З а д а ч и	134
Л и т е р а т у р а	135
ДОБАВЛЕНИЕ Ю. В. Прохоров	136