МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ
УНИВЕРСИТЕТ
имени М.В.ЛОМОНОСОВА

Механико-математический факультет

В. А. МАЛЫШЕВ, М. В. МЕНЬШИКОВ,
Е. Н. ПЕТРОВА

ВВЕДЕНИЕ В ТЕОРИЮ ВЕРОЯТНОСТЕЙ

Москва 1997

УДК 519.2
М18

Малышев В. А., Меньшиков М. В., Петрова Е. Н.
Введение в теорию вероятностей. М.: Изд-во механико-математического факультета МГУ, 1997 г. —117 с.

Настоящая книга является записью основного курса лекций по теории вероятностей, прочитанного первыми двумя авторами на механико-математическом факультете МГУ для студентов второго курса в весеннем семестре 1995 года. Главной целью авторов является изложение на простейших примерах важнейших понятий и методов современной теории вероятностей наряду с классическими.

Выражаем благодарность Ахмитзянову Р.Р. и Маните А.Д. за полезные обсуждения и помощь в написании этой книги.

М $\frac{1602090000-005}{3Ш7(03)-97}$ Без объявл.

ISBN 5-87597-032-4

Оглавление

Лекция 1

Конечные и счетные вероятностные пространства

Пусть задано конечное множество $\Omega = \omega_1, \ldots, \omega_n$. Часто мы имеем дело с ситуацией, когда каждому элементу приписан вес (мера), т.е. неотрицательное число $w(\omega_i)$, характеризующее "частоту" появления элемента ω_i в некоторых обстоятельствах. Можно нормировать эти веса к единице (если не все они нули), т.е. ввести новые числа

$$\mathsf{P}(\omega_i) = Z^{-1} w(\omega_i), \quad Z = \sum_{i=1}^{n} w(\omega_i).$$

Отображение $\mathsf{P} : \Omega \to [0, 1]$ называется *вероятностной мерой* (или вероятностью). В статистической физике нормирующий множитель Z называется статистической суммой. Аналогичным образом можно ввести вероятностную меру на счетном множестве Ω, если

$$\sum_{i=1}^{\infty} w(\omega_i) < \infty.$$

Определение 1. Конечным (счетным) *вероятностным пространством* называется пара (Ω, P), где Ω — конечное (счетное) множество, а P вероятностная мера, то есть набор неотрицательных чисел $\mathsf{P}(\omega), \omega \in \Omega$, таких что $\sum_{\omega \in \Omega} \mathsf{P}(\omega) = 1$. Множество Ω называется *пространством элементарных событий*, его подмножества — *событиями*, а его элементы — *элементарными событиями*.

Если все $\mathsf{P}(\omega)$ равны, то говорят, что все элементарные события *равновероятны*. Вероятность события A определяется как сумма вероятностей входящих в него элементарных событий:

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

На практике часто говорят о случайном выборе элемента ω из множества Ω с вероятностью $\mathsf{P}(\omega)$. Мы говорим, что при этом выборе событие A осуществилось, если $\omega \in A$. Пересечение событий $A \cap B$ соответствует событию, что каждое из них осуществилось, объединение $A \cup B$ соответствует тому, что осуществилось хотя бы одно из них. Мы полагаем, что вероятность пустого события равна нулю.

Язык теории множеств	Язык теории вероятностей
Множество	Пространство элементарных событий
Точки	Элементарные события
Подмножества	События
Выбор точки $\omega \in A$	Произошло событие A
Объединение $A_1 \cup A_2$	Произошло хотя бы одно из A_1, A_2
Пересечение $A_1 \cap A_2$	Произошли оба события A_1, A_2
Дополнение $\bar{A} = \Omega \setminus A$	Событие A не произошло
Измеримая функция	Случайная величина
Взвешенная сумма (интеграл)	Математическое ожидание

Схема Бернулли

Приведем простой, но очень полезный пример вероятностного пространства, на котором в дальнейшем мы проиллюстрируем важные теоремы, верные и в более общей ситуации.

Рассмотрим результат n последовательных бросаний монеты; отождествим его с конечной последовательностью длины n из нулей и единиц: $\omega = (x_1, \ldots, x_n)$, где $x_i = 1$, если при i-м бросании появился "орел" (успех) и $x_i = 0$, если появилась "решка" (неудача). Таким образом, пространство элементарных событий

$$\Omega = \{\omega : \omega = (x_1, \ldots, x_n),\ x_i = 0, 1\}$$

состоит из 2^n точек. Обозначим

$$S_n = S_n(\omega) = \sum_{i=1}^{n} x_i$$

число успехов в последовательности ω. Зададим вероятность элементарного события

$$\mathsf{P}(\omega) = p^{S_n(\omega)} q^{n - S_n(\omega)}, \tag{1.1}$$

где p и q неотрицательные числа такие, что $p + q = 1$.

Замечание 1. Вероятностная мера P задана с помощью (1.1) корректно. В самом деле, нам необходимо проверить, что

$$\sum_{\omega \in \Omega} \mathsf{P}(\omega) = 1.$$

Поскольку число элементарных событий, содержащих ровно k успехов, равно C_n^k, имеем по биному Ньютона:

$$\sum_{\omega \in \Omega} \mathsf{P}(\omega) = \sum_{k=0}^{n} C_n^k p^k q^{n-k} = (p+q)^n = 1.$$

Замечание 2. Как видно из замечания 1, вероятность события, состоящего в том, что орел выпал ровно k раз, равна $\mathsf{P}\{S_n = k\} = C_n^k p^k q^{n-k}$. Набор вероятностей $P_k = \mathsf{P}\{S_n = k\}, k = 0, \ldots, n$, называется *биномиальным распределением.*

Упражнение 1. Вероятность того, что при i-м бросании выпадет орел, равна p. Доказать.

Описанная нами вероятностная модель называется схемой Бернулли, или моделью n независимых испытаний с двумя исходами (позднее мы поймем, почему в названии присутствует слово "независимые").

Докажем теперь, что для типичной последовательности число единиц примерно равно np. Что это означает точно, мы сейчас сформулируем.

Теорема 1 (Закон больших чисел для схемы Бернулли). Рассмотрим вероятностное пространство (Ω, P), соответствующее схеме Бернулли с n испытаниями. Тогда для любого фиксированного $\varepsilon > 0$

$$\mathsf{P}\left\{\omega : \left|\frac{S_n(\omega)}{n} - p\right| \geq \varepsilon\right\} \to 0 \quad \text{при } n \to \infty. \tag{1.2}$$

Не будем забывать, что Ω и P различны для разных n. Для доказательства этой теоремы нам понадобятся новые понятия: случайной величины, математического ожидания и дисперсии, к определению которых мы и переходим.

Определение 2. Пусть имеется конечное вероятностное пространство (Ω, P) . *Случайной величиной* называется любая вещественная функция $\xi : \Omega \to \mathbf{R}$. *Математическим ожиданием*, или средним значением случайной величины называется число

$$\mathsf{E}\xi = \sum_{\omega \in \Omega} \xi(\omega) \mathsf{P}(\omega).$$

Дисперсией случайной величины называется число

$$\mathsf{D}\xi = \mathsf{E}(\xi - \mathsf{E}\xi)^2 = \sum_{\omega \in \Omega} (\xi(\omega) - \mathsf{E}\xi)^2 \mathsf{P}(\omega).$$

Упражнение 2. Доказать, что

$$\mathsf{E}\xi = \sum_{a} a \mathsf{P}\{\xi = a\},$$

где сумма берется по всем значениям случайной величины ξ, а событие $\{\xi = a\} = \{\omega : \xi(\omega) = a\}$. Таким образом, если случайная величина принимает одно единственное значение, ее математическое ожидание совпадает с этим значением.

Упражнение 3. Доказать следующие свойства математического ожидания и дисперсии.
1. Математическое ожидание линейной комбинации случайных величин равно линейной комбинации их математических ожиданий:

$$\mathsf{E}(c_1\eta_1 + \ldots + c_m\eta_m) = c_1\mathsf{E}\eta_1 + \ldots + c_m\mathsf{E}\eta_m.$$

2. Если $\xi(\omega) \le \eta(\omega)$ для всех $\omega \in \Omega$, то $\mathsf{E}\xi \le \mathsf{E}\eta$. В частности, если константа c такова, что $0 \le \xi(\omega) \le c$, то $\mathsf{E}\xi \le c$.
3. $\mathsf{D}(c\xi) = c^2\mathsf{D}\xi$ для любой константы c.

Нам понадобится также следующее утверждение. Оно легко доказывается, но, как мы увидим в дальнейшем, играет важную роль во многих рассуждениях.

Лемма 1 (Неравенство Чебышева). Для любого $\varepsilon > 0$ и любой неотрицательной случайной величины u имеет место следующее неравенство:

$$\mathsf{P}\{u \ge \varepsilon\} \le \frac{\mathsf{E}u}{\varepsilon}. \tag{1.3}$$

Доказательство. По определению математического ожидания

$$\mathsf{E}u = \sum_{\omega \in \Omega} u(\omega)\mathsf{P}(\omega) \ge \sum_{\omega:\, u(\omega) \ge \varepsilon} u(\omega)\mathsf{P}(\omega) \ge$$

$$\ge \sum_{\omega:\, u(\omega) \ge \varepsilon} \varepsilon \mathsf{P}(\omega) = \varepsilon \mathsf{P}\{\omega : u(\omega) \ge \varepsilon\},$$

откуда и следует (1.3).

Следствие. Если u — произвольная случайная величина, то для любого $\varepsilon > 0$

$$\mathsf{P}\{| u - \mathsf{E}u |\geq \varepsilon\} \leq \frac{\mathsf{D}u}{\varepsilon^2}. \tag{1.4}$$

Обозначим ξ_i случайную величину, равную единице, если при i-м бросании выпал орел, и нулю в противном случае: $\xi_i(\omega) = x_i$. Будем называть такую случайную величину *бернуллиевской*, как и меру P для случая одного бросания (определенную на пространстве $\Omega = \{0,1\}$). Очевидно,

$$S_n(\omega) = x_1 + \ldots + x_n = \xi_1(\omega) + \ldots + \xi_n(\omega).$$

Заметим, что для любого i

$$\mathsf{P}(\xi_i = 1) = p, \;\; \mathsf{P}(\xi_i = 0) = q,$$

то есть мера P не зависит от номера i случайной величины. В таком случае случайные величины ξ_i называются *одинаково распределенными*.

Упражнение 4. Доказать, что $\mathsf{E}\xi_i = p, \;\; \mathsf{D}\xi_i = pq$.

Для доказательства закона больших чисел мы хотим использовать неравенство Чебышева, следовательно, нам необходимо подсчитать дисперсию случайной величины S_n. Для этого нам понадобится понятие независимости.

Определение 3. Случайные величины ξ и η называются *независимыми*, если для любых $a, b \in \mathbf{R}$

$$\mathsf{P}\{\xi = a, \eta = b\} = \mathsf{P}\{\xi = a\}\mathsf{P}\{\eta = b\}. \tag{1.5}$$

Предложение 1. Для любых $i, j = 1, \ldots, n, \;\; i \neq j$, определенные выше случайные величины ξ_i и ξ_j независимы.

Доказательство. Подсчитаем, вероятность того, что при i-м бросании выпал орел, а при j-м бросании выпала решка:

$$\mathsf{P}\{\xi_i = 1, \xi_j = 0\} = \sum_{\omega\in\Omega: x_i=1, x_j=0} p^{S_n(\omega)} q^{n-S_n(\omega)} = pq.$$

В упражнении 1 мы видели, что

$$\mathsf{P}\{\xi_i = 1\} = p, \;\; \mathsf{P}\{\xi_j = 0\} = q.$$

Таким образом,

$$\mathsf{P}\{\xi_i = 1, \xi_j = 0\} = \mathsf{P}\{\xi_i = 1\}\mathsf{P}\{\xi_j = 0\}.$$

Проверяя аналогично для других значений x_i и x_j, получим (1.5).

Предложение 2. Если случайные величины ξ и η независимы, то $\mathsf{E}(\xi\eta) = \mathsf{E}\xi\mathsf{E}\eta$.

Доказательство. По определению математического ожидания имеем:

$$\mathsf{E}(\xi\eta) = \sum_{\omega} \xi(\omega)\eta(\omega)P(\omega) =$$

$$\sum_{a,b} ab \sum_{\omega:\xi(\omega)=a,\eta(\omega)=b} \mathsf{P}(\omega) = \sum_{a,b} ab\mathsf{P}\{\xi = a, \eta = b\}.$$

В силу независимости ξ и η последнее выражение равно

$$\sum_{a} a\mathsf{P}\{\xi = a\} \sum_{b} b\mathsf{P}\{\eta = b\} = \mathsf{E}\xi\mathsf{E}\eta.$$

Упражнение 5. Если случайные величины η_1 и η_2 независимы, а c_1 и c_2 произвольные константы, то случайные величины $\eta_1 - c_1$ и $\eta_2 - c_2$ также независимы. Доказать.

Доказательство теоремы 1. Из свойства линейности математического ожидания имеем: $\mathsf{E}S_n/n = p$. Воспользовавшись неравенством (1.4), получим

$$\mathsf{P}\left\{\left|\frac{S_n}{n} - p\right| \geq \varepsilon\right\} \leq \frac{\mathsf{D}(S_n/n)}{\varepsilon^2}. \tag{1.6}$$

Вновь по линейности:

$$\mathsf{D}\left(\frac{S_n}{n}\right) = \mathsf{E}\left(\frac{S_n}{n} - p\right)^2 = \frac{\mathsf{E}(S_n - np)^2}{n^2}.$$

Обозначим $u_i = \xi_i - p$, тогда

$$\mathsf{D}\left(\frac{S_n}{n}\right) = \frac{1}{n^2}\mathsf{E}\left(\sum_{i=1}^{n} u_i\right)^2 = \frac{1}{n^2}\left(\sum_{i=1}^{n} \mathsf{E}u_i^2 + \sum_{i\neq j} \mathsf{E}u_i u_j\right). \tag{1.7}$$

Заметим (см. упражнение 4), что

$$\mathsf{E}u_i^2 = \mathsf{D}\xi_i = pq. \tag{1.8}$$

Случайные величины u_i и u_j независимы для $i \neq j$ (см. предложение 1 и упражнение 5). Поэтому

$$\mathsf{E} u_i u_j = \mathsf{E} u_i \mathsf{E} u_j = 0. \tag{1.9}$$

Подставляя (1.9) и (1.8) в (1.7), получаем

$$\mathsf{D}\left(\frac{S_n}{n}\right) = \frac{pq}{n}. \tag{1.10}$$

Подставляя теперь (1.10) в (1.6), получаем утверждение теоремы.

Из закона больших чисел следует, что

$$\mathsf{P}\{n(p-\varepsilon) \leq S_n \leq n(p+\varepsilon)\} \to 1, \quad n \to \infty, \tag{1.11}$$

то есть при больших n вероятность множества последовательностей с таким числом успехов близка к единице, и в этом смысле последовательности, входящие в это множество, являются типичными. Зададимся вопросом: насколько мала вероятность того, что число успехов равно $[\alpha n]$, где $\alpha \neq p$? Подобного рода задачи называются задачами о больших уклонениях.

Теорема 2 (большие уклонения). Пусть $0 < \alpha < 1, \alpha \neq p$. Тогда для достаточно больших n

$$\mathsf{P}\{S_n = [\alpha n]\} \sim \frac{1}{\sqrt{2\pi\alpha(1-\alpha)n}} \exp(-n\hat{H}(p,\alpha)), \tag{1.12}$$

где

$$\hat{H}(p,\alpha) = \alpha \ln \frac{\alpha}{p} + (1-\alpha) \ln \frac{1-\alpha}{1-p},$$

а символ $\sim$ означает, что отношение величин стремится к единице при $n \to \infty$.

Замечание 3. Обозначим через P бернуллиевскую меру с вероятностью успеха p, а через π бернуллиевскую меру с вероятностью успеха α. Функция $\hat{H}(p,\alpha)$ называется *энтропией* меры π относительно меры P.

Доказательство теоремы 2. Для простоты докажем результат для случая $p = 1/2$, когда все последовательности равновероятны: $\mathsf{P}(\omega) = 2^{-n}$ и

$$\mathsf{P}\{S_n = [\alpha n]\} \sim \frac{1}{\sqrt{2\pi\alpha(1-\alpha)n}} \exp(-n(\ln 2 - H(\alpha)),$$

$$H(\alpha) = -\alpha \ln \alpha - (1-\alpha)\ln(1-\alpha).$$

Функция $H(\alpha)$ называется *энтропией меры* π. Она неотрицательна, выпукла кверху, равна нулю при $\alpha = 0, \alpha = 1$ и достигает максимума при $\alpha = 1/2$. Ее можно интерпретировать как степень неопределенности результата при бросании монеты с вероятностью успеха α. Действительно, при $\alpha = 0$ появляется только решка, а при $\alpha = 1$ только орел и, следовательно, никакой неопределенности нет. Если же $\alpha = 1/2$, появление орла и решки одинаково вероятно, и в этом смысле неопределенность результата максимальна.

Имеем:

$$\mathsf{P}\{S_n = [\alpha n]\} = 2^{-n} C_n^{[\alpha n]} = 2^{-n} \frac{n!}{[\alpha n]!(n-[\alpha n])!}\,. \tag{1.13}$$

Воспользуемся далее известной формулой Стирлинга:

$$n! = \sqrt{2\pi n} n^n e^{-n}(1+R(n)), \tag{1.14}$$

где $R(n) \to 0$ при $n \to \infty$. Тогда

$$\frac{n!}{[\alpha n]!(n-[\alpha n])!} \sim \tag{1.15}$$

$$\sim \frac{\sqrt{2\pi n}\, n^n e^{-n}}{\sqrt{2\pi[\alpha n]}[\alpha n]^{[\alpha n]} e^{-[\alpha n]} \sqrt{2\pi(n-[\alpha n])}(n-[\alpha n])^{n-[\alpha n]}\; e^{-(n-[\alpha n])}}\,.$$

Убрав в правой части (1.15) знаки целой части и произведя необходимые сокращения, получим, вернувшись к (1.13):

$$\mathsf{P}\{S_n = [\alpha n]\} \sim \frac{1}{\sqrt{2\pi\alpha(1-\alpha)n}} 2^{-n} \alpha^{-\alpha n}(1-\alpha)^{-(1-\alpha)n},$$

откуда и следует теорема 2.

Замечание 4. Из теоремы 2 в частности следует, что при $p = 1/2$ вероятность того, что число единиц *в точности* равно числу нулей, весьма мала (но не экспоненциально мала) при больших (четных) n. А именно, пусть $n = 2k$. Тогда

$$\mathsf{P}\{S_n = k\} \sim \frac{1}{\sqrt{\pi k}}.$$

С помощью формулы Стирлинга можно получать также более аккуратные оценки. Зададимся вопросом, какова вероятность того, что число успехов в серии длины n равно $np + x$.

Теорема 3 (Локальная предельная теорема). Пусть $0 < p < 1$. Тогда для любого x такого, что $pn + x$ есть целое неотрицательное число и $x = O(n^{1/2})$

$$\mathsf{P}\{S_n = pn + x\} \sim \frac{1}{\sqrt{2\pi npq}} \exp\left(-\frac{x^2}{2npq}\right) \tag{1.16}$$

при $n \to \infty$.

Доказательство. Для упрощения вычислений докажем теорему для случая $p = 1/2$. Воспользуемся формулой Стирлинга:

$$\mathsf{P}\{S_n = \frac{n}{2} + x\} = 2^{-n} C_n^{(n/2)+x} = 2^{-n} \frac{n!}{(n/2+x)!(n/2-x)!} \sim$$

$$\sim 2^{-n} \frac{n^n e^{-n}}{e^{-(n/2+x)} e^{-(n/2-x)} \left(\frac{n}{2}+x\right)^{(n/2+x)} \left(\frac{n}{2}-x\right)^{(n/2-x)}} \times$$

$$\times \frac{\sqrt{2\pi n}}{\sqrt{2\pi(n/2+x)}\sqrt{2\pi(n/2-x)}} =$$

$$= \frac{2^{-n} n^n}{\sqrt{2\pi(1/4)(n-(4x^2/n))}} \left(\frac{n}{2}+x\right)^{-(n/2+x)} \left(\frac{n}{2}-x\right)^{-(n/2-x)} \sim$$

$$\sim \sqrt{\frac{2}{\pi n}} \left(1 + \frac{2x}{n}\right)^{-(n/2+x)} \left(1 - \frac{2x}{n}\right)^{-(n/2-x)}.$$

Поскольку $x = O(n^{1/2})$, то при больших n

$$1 + \frac{2x}{n} \sim \exp\left(\frac{2x}{n}\right), \quad 1 - \frac{2x}{n} \sim \exp\left(-\frac{2x}{n}\right),$$

и поэтому

$$\left(1 + \frac{2x}{n}\right)^{-x} \left(1 - \frac{2x}{n}\right)^{x} \sim \exp\left(-\frac{4x^2}{n}\right).$$

Далее,

$$\left(1 + \frac{2x}{n}\right)^{-\frac{n}{2}} \left(1 - \frac{2x}{n}\right)^{-\frac{n}{2}} = \left(1 - \frac{4x^2}{n^2}\right)^{-\frac{n}{2}} \sim \exp\left(\frac{2x^2}{n}\right).$$

Окончательно получаем:

$$\mathsf{P}\{S_n = \frac{n}{2} + x\} \sim \sqrt{\frac{2}{\pi n}} \exp\left(-\frac{2x^2}{n}\right).$$

Литература

Феллер В. Введение в теорию вероятностей и ее приложения. М., Мир, 1984. Том 1, главы 1, 4, 7, 9, 10.

Лекция 2

Аксиоматика Колмогорова

В прошлой лекции мы занимались конечными или счетными вероятностными пространствами. Точки назывались элементарными событиями, любое подмножество было объявлено событием. В случае несчетного пространства невозможно определить естественным образом вероятность любого подмножества. Самый естественный и широко применяемый в настоящее время способ ввести вероятность на части подмножеств дает *аксиоматика Колмогорова*, состоящая из приводимых ниже определений.

Определение 1. Пусть имеется произвольное множество (пространство) Ω. Система Σ его подмножеств называется *алгеброй* (подмножеств), если выполнены следующие условия.

1. Если $A \in \Sigma$, то его дополнение $\bar{A} = \Omega \setminus A$ также принадлежит Σ.

2. Если A и B принадлежат Σ, то их объединение и пересечение также принадлежат Σ. (Кратко можно сказать, что алгебра замкнута относительно операций конечного объединения и пересечения.)

3. Пустое подмножество $\emptyset$ принадлежит Σ.

Упражнение 1. Доказать, что:

а) вместо условия 2 достаточно потребовать, чтобы только пересечение $A \cap B$ принадлежало Σ;

б) если $A_1, \dots, A_k$ принадлежат Σ, то $\cap_{i=1}^k A_i \in \Sigma$ и $\cup_{i=1}^n A_i \in \Sigma$.

Определение 2. Функция $\mathsf{P} = \mathsf{P}(A)$ на алгебре Σ, принимающая значения в $[0, 1]$, называется *конечно-аддитивной вероятностной мерой* (или конечно-аддитивной вероятностью), если для любых двух непересекающихся множеств A и B из Σ, $A \cap B = \emptyset$,

$$\mathsf{P}(A \cup B) = \mathsf{P}(A) + \mathsf{P}(B) \tag{2.1}$$

и

$$\mathsf{P}(\Omega) = 1.$$

Определение 3. Алгебра подмножеств Ω называется *σ-алгеброй*, если она замкнута относительно взятия счетного числа пересечений. Элементы σ-алгебры (т.е. подмножества, входящие в σ-алгебру) называются *событиями.* Пара (Ω, Σ), где Σ есть σ-алгебра подмножеств Ω, называется *измеримым пространством.*

Определение 4. Пусть каждому подмножеству A, принадлежащему некоторой σ-алгебре Σ подмножеств множества Ω, сопоставлено число $\mathsf{P}(A)$, $0 \le \mathsf{P}(A) \le 1$. Функция $\mathsf{P}(A)$ называется *вероятностной мерой* (или просто вероятностью), если выполнены следующие два условия.

1. *Счетная аддитивность.* Мера объединения счетного числа непересекающихся событий равна сумме их мер:

$$\mathsf{P}\big(\bigcup_{i=1}^{\infty} A_i\big) = \sum_{i=1}^{\infty} \mathsf{P}(A_i), \qquad \text{если } A_i \cap A_j = \emptyset, \ \ i \ne j.$$

2. *Нормировка*: $\mathsf{P}(\Omega) = 1$.

Тройка $(\Omega, \Sigma, \mathsf{P})$ называется *вероятностным пространством.*

Замечание 1. В первой лекции, рассматривая конечное пространство Ω, мы называли вероятностным пространством пару (Ω, P), подразумевая, что Σ есть алгебра всех подмножеств Ω.

Замечание 2. Непосредственно из определения следует, что
1) $P(\emptyset) = 0$;
2) Если $A \subset B$, то $P(A) \le P(B)$;
3) $P(A \cup B) \le P(A) + P(B)$.

Упражнение 2. Доказать, что свойство счетной аддитивности эквивалентно любому из следующих свойств непрерывности:

а) для любой бесконечной последовательности $\{A_i\}$ расширяющихся множеств из Σ (т.е. таких, что $A_i \subseteq A_{i+1}$ для всех i)

$$\lim_{i \to \infty} \mathsf{P}(A_i) = \mathsf{P}\big(\bigcup_{i=1}^{\infty} A_i\big);$$

б) для любой последовательности $\{A_i\}$ вложенных множеств из Σ (т.е. таких, что $A_i \supseteq A_{i+1}$ для всех i)

$$\lim_{i \to \infty} \mathsf{P}(A_i) = \mathsf{P}\big(\bigcap_{i=1}^{\infty} A_i\big).$$

Определение 5. Вещественная функция $\xi(\omega) : \Omega \to \mathbf{R}$ называется *случайной величиной*, если для любого числа x множество

$\{\omega : \xi(\omega) < x\}$ является событием, т.е. принадлежит σ-алгебре Σ. На языке теории функций такая функция называется измеримой относительно σ-алгебры Σ.

Пример. Обозначим через $\mathcal{A}$ систему множеств на прямой, каждое из которых есть конечная сумма непересекающихся интервалов вида $(a, b]$, $-\infty \le a < b \le \infty$:

$$A \in \mathcal{A}, \text{ если } A = \sum_{i=1}^{n} (a_i, b_i], \quad n < \infty.$$

Включим в $\mathcal{A}$ также пустое множество. Нетрудно видеть, что система $\mathcal{A}$ является алгеброй. Обозначим $\mathcal{B}(\mathbf{R})$ наименьшую σ-алгебру, содержащую алгебру $\mathcal{A}$. (Доказать, что такая σ-алгебра существует.) Эта σ-алгебра называется *борелевской*, а входящие в нее множества — борелевскими.

В некоторых случаях более удобным является следующее *эквивалентное* определение случайной величины.

Определение 5$'$. Функция $\xi = \xi(\omega)$ называется случайной величиной, если для любого $B \in \mathcal{B}(\mathbf{R})$

$$\xi^{-1}(B) \equiv \{\omega : \xi(\omega) \in B\} \in \Sigma.$$

Эквивалентность двух определений легко следует из следующих хорошо известных фактов теории меры.

1. Система полубесконечных интервалов $\mathcal{I} = \{ (-\infty, x), \ x \in \mathbf{R} \}$ порождает σ-алгебру $\mathcal{B}(\mathbf{R})$ борелевских множеств на прямой, т.е. наименьшая σ-алгебра $\sigma(\mathcal{I})$, содержащая систему множеств $\mathcal{I}$, совпадает с $\mathcal{B}(\mathbf{R})$.

2. Пусть $\mathcal{E}$ — некоторая система множеств такая, что $\sigma(\mathcal{E}) = \mathcal{B}(\mathbf{R})$. Для того, чтобы некоторая функция $\xi = \xi(\omega)$ была измеримой, необходимо и достаточно, чтобы

$$\{\omega : \xi(\omega) \in E\} \in \Sigma$$

для всех $E \in \mathcal{E}$.

(Доказательство можно найти, например, в книге Ширяева, глава II, §4.)

Функция $\varphi : \mathbf{R} \to \mathbf{R}$ называется *борелевской*, если она есть измеримое отображение измеримого пространства $(\mathbf{R}, \mathcal{B}(\mathbf{R}))$ в себя, т.е. для любого $B \in \mathcal{B}(\mathbf{R})$ имеет место $\varphi^{-1}(B) \in \mathcal{B}(\mathbf{R})$. Например, все непрерывные функции являются борелевскими.

Предложение 1. Пусть ξ — случайная величина, а φ — борелевская функция. Тогда $\eta = \varphi(\xi)$ — случайная величина.

Доказательство. Так как $\eta = \varphi \circ \xi$, то $\eta^{-1}(B) = \xi^{-1}(\varphi^{-1}(B))$ для всех B. Возьмем произвольное $B \in \mathcal{B}(\mathbf{R})$. Поскольку φ — борелевская, имеем $\varphi^{-1}(B) \in \mathcal{B}(\mathbf{R})$. Поскольку ξ — случайная величина, получаем $\eta^{-1}(B) = \xi^{-1}(\varphi^{-1}(B)) \in \mathcal{F}$.

Определение 6. *Математическим ожиданием* $\mathbf{E}\xi$ случайной величины $\xi(\omega)$ называется интеграл Лебега: $\int\limits_{\Omega} \xi(\omega) dP(\omega)$.

В дальнейшем мы приведем определение интеграла Лебега. Однако в ближайших лекциях нам понадобится только интеграл Лебега от кусочно-постоянных функций, т.е. таких, что существует разбиение $\Omega = \cup A_i$ на конечное число попарно непересекающихся множеств $A_i \in \Sigma$, так что $\xi(\omega)$ постоянна (и равна некоторому c_i) на каждом A_i. Такие функции называются иногда простыми. В этом случае по определению

$$\mathsf{E}\xi = \int\limits_{\Omega} \xi(\omega) d\mathsf{P}(\omega) = \sum_i c_i \mathsf{P}(A_i).$$

Схема Бернулли

Пусть Ω — множество всех бесконечных последовательностей ω нулей и единиц, т.е. $\omega = (x_1, \ldots, x_i, \ldots)$, $x_i = 0$ или 1, а индекс i пробегает все целые положительные числа. Каждой точке $\omega_n = (x_1, \ldots, x_n)$ пространства $\Omega^{(n)}$, построенного в предыдущей лекции (описывающего n испытаний Бернулли), соответствует подмножество множества Ω, которое мы будем обозначать той же буквой ω_n — множество всех последовательностей ω, первые n координат которых совпадают с последовательностью ω_n. Мы получаем разбиение Ω на 2^n непересекающихся подмножеств. Объединения всевозможных ω_n образуют алгебру Σ_n, изоморфную алгебре всех подмножеств пространства $\Omega^{(n)}$. Поэтому вероятность $\mathsf{P} = \mathsf{P}^{(n)}$, введенная ранее на $\Omega^{(n)}$, является в то же время конечно-аддитивной вероятностной мерой на Σ_n. При этом $\mathsf{P}^{(n)}(\omega_n) = p^k q^{n-k}$, где k — число единиц в последовательности ω_n.

Упражнение 3. Показать, что при $m < n$ алгебра Σ_m является подалгеброй Σ_n, и меры $\mathsf{P}^{(m)}$ и $\mathsf{P}^{(n)}$ совпадают на Σ_m.

Нам хотелось бы определить меру P на *всеобъемлющем* пространстве Ω такую, что на Σ_n она совпадала бы с $\mathsf{P}^{(n)}$. Зачем это нужно, мы увидим дальше, а делается это с помощью следующей теоремы, которую мы приводим без доказательства.

Теорема 1 (Теорема о продолжении меры). Пусть в Ω дана алгебра множеств Σ' и конечно-аддитивная вероятностная мера на ней, удовлетворяющая одному из двух следующих эквивалентных условий (см. лемму 1).

1. *Счетная аддитивность*: для любой бесконечной системы попарно непересекающихся множеств $A_1, A_2, \ldots$ из Σ' таких, что их объединение принадлежит Σ',

$$\mathsf{P}(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathsf{P}(A_i).$$

2. *Непрерывность*: если имеется последовательность убывающих множеств $A_1 \supseteq A_2 \supseteq \ldots$ из Σ' такая, что их пересечение пусто, $\cap A_i = \emptyset$, то $\mathsf{P}(A_i) \to 0$ при $i \to \infty$.

Тогда эта мера может быть единственным образом продолжена до вероятностной меры на минимальной σ-алгебре Σ, содержащей Σ'.

В первом условии этой теоремы есть тонкий пункт: если множества A_i, образующие счетное семейство, принадлежат Σ' и попарно не пересекаются, то вовсе не обязательно, что их объединение принадлежит алгебре Σ'. Но мы требуем выполнение следующего свойства: *если оно принадлежит*, то его мера равна сумме мер A_i.

Лемма 1. Пусть задана конечно-аддитивная вероятностная мера P на алгебре Σ'. Тогда свойства счетной аддитивности и непрерывности из теоремы 1 эквивалентны.

Доказательство.

1. Непрерывность + конечная аддитивность $\Rightarrow$ счетная аддитивность. Пусть последовательность множеств $\{B_n\}$ такова, что $B_i \cap B_j = \emptyset$ при $i \neq j$ и $B = \cup B_i \in \Sigma'$. Обозначим $A_n = B \setminus (\cup_{i=1}^{n} B_i)$. Тогда, очевидно, для любого n $\quad A_{n+1} \subseteq A_n$ и $\cap_{n=1}^{\infty} A_n = \emptyset$. В силу свойства непрерывности $\mathsf{P}(A_i) \to 0$ при $i \to \infty$, а это означает, что

$$\mathsf{P}(B) - \sum_{i=1}^{n} \mathsf{P}(B_i) \to 0, \quad n \to \infty,$$

то есть

$$\sum_{i=1}^{\infty} \mathsf{P}(B_i) = P(B).$$

2. Счетная аддитивность $\Rightarrow$ непрерывность. Пусть задана последовательность вложенных множеств $A_1 \supseteq A_2 \supseteq \ldots$ из Σ' с пустым пересечением $\cap A_i = \emptyset$. Рассмотрим последовательность множеств B_i, $B_i = A_i \setminus A_{i-1}$. Очевидно, $B_i \cap B_j = \emptyset$ при $i \neq j$. Кроме того, поскольку $\cap_{n=1}^{\infty} A_n = \emptyset$, то $\cup_{i=1}^{\infty} B_i = A_1$. В силу счетной аддитивности

$$\sum_{i=1}^{\infty} \mathsf{P}(B_i) = \mathsf{P}(A_1).$$

Но поскольку

$$\sum_{i=1}^{n} \mathsf{P}(B_i) = \mathsf{P}(\cup_{i=1}^{n} B_i) = \mathsf{P}(A_1 \setminus A_n) = \mathsf{P}(A_1) - \mathsf{P}(A_n),$$

то

$$\mathsf{P}(A_n) = \mathsf{P}(A_1) - \sum_{i=1}^{n} \mathsf{P}(B_i) \to 0$$

при $n \to \infty$.

Замечание 3. Свойство непрерывности можно эквивалентным образом сформулировать так. Для любой последовательности вложенных множеств $A_1 \supseteq A_2 \supseteq \ldots$ из Σ', такой что $\cap_{n=1}^{\infty} A_n = A \in \Sigma'$, $\mathsf{P}(A_n) \to \mathsf{P}(A)$ при $n \to \infty$. (Упражнение: доказать это.)

Определим Σ' как объединение всех Σ_n, т.е. как систему множеств, содержащую все элементы каждой из Σ_n. Очевидно, что Σ' алгебра. Определим функцию P на элементах Σ' так. Пусть $A \in \Sigma'$, тогда существует n такое, что $A \in \Sigma_n$. Положим $\mathsf{P}(A) = \mathsf{P}^{(n)}(A)$. Упражнение 3 показывает, что такое определение корректно, т.е. если A можно представить как элемент разных Σ_n и Σ_m, то $\mathsf{P}(A)$ одна и та же. Заметим, что $\mathsf{P}(\emptyset) = 0$, поэтому $\mathsf{P}(\Omega) = 1$. Из упражнения 3 следует также, что для любой конечной системы множеств из Σ' найдется n такое, что все они принадлежат Σ_n. Но вероятность $\mathsf{P}^{(n)}$ по построению конечно-аддитивна, поэтому и P является конечно-аддитивной вероятностной мерой на (Ω, Σ').

Теорема 2. Определенная выше бернуллиевская мера P на Σ' удовлетворяет условию теоремы о продолжении меры.

Доказательство. Пусть задана последовательность $\{A_n\}$ вложенных множеств, принадлежащих Σ'. Покажем, что пересечение этих множеств всегда не пусто. Рассмотрим первое множество A_1 этой последовательности. Оно принадлежит Σ', следовательно, найдется номер N такой, что $A_1 \in \Sigma_N$. Это означает, что $A_1 = \cup_{k=1}^{m} \omega_N^k$, где ω_N^k есть множество бесконечных последовательностей с фиксированными первыми N координатами. Поскольку каждое из последующих множеств A_i, $i > 1$, является подмножеством множества A_1, то в каждом из них найдется точка, принадлежащая какому-либо из множеств ω_N^k (то есть бесконечная последовательность, первые N координат которой совпадают с первыми N координатами какого-то из ω_N^k). Поскольку последовательность вложенных множеств бесконечна, а множеств ω_N^k конечное число, то найдутся бесконечная подпоследовательность $A_{n_1^{(1)}}, A_{n_2^{(1)}}, \ldots, A_{n_p^{(1)}}, \ldots$, последовательности $\{A_n\}$ и номер k_1, такие, что каждое из множеств этой подпоследовательности содержит точку, принадлежащую $\omega_N^{k_1}$. Рассмотрим первый элемент $A_{n_1^{(1)}}$ этой подпоследовательности и повторим предыдущие рассуждения. Найдется, очевидно, множество ω_M, где $M > N$, и бесконечная подпоследовательность $A_{n_1^{(2)}}, A_{n_2^{(2)}}, \ldots,$ $A_{n_p^{(2)}}, \ldots$ последовательности $A_{n_1^{(1)}}, A_{n_2^{(1)}}, \ldots, A_{n_p^{(1)}}, \ldots$, такие, что каждое из множеств $A_{n_p^{(2)}}$ содержит точку, принадлежащую ω_M. Будем повторять описанную схему и построим бесконечную последовательность ω, первые N координат которой совпадают с ω_N, следующие $M - N$ координат совпадают с ω_M, и т.д. Тогда, очевидно, $\omega \in \cap_{n=1}^{\infty} A_n$.

Далее, если $\{A_n\}$ — произвольная последовательность множеств из Σ', то $A = \cap A_n \notin \Sigma'$. В самом деле, если бы $A \in \Sigma'$, то $A_n \setminus A \in \Sigma'$ и последовательность $\{A_n \setminus A\}$ имела бы пустое пересечение, чего быть не может. Таким образом, условие теоремы о продолжении меры выполняется.

По теореме о продолжении меры мера P однозначно продолжается на минимальную σ-алгебру Σ, содержащую Σ'. Говорят, что Σ есть σ-алгебра, порожденная множествами ω_n.

Итак, мы построили вероятностное пространство $(\Omega, \Sigma, \mathsf{P})$, называемое схемой Бернулли (испытаниями Бернулли). Теперь мы ответим на вопрос, зачем нужна мера P на всем Ω. Дело в том, что многие события, которые естественно возникают в нашем воображении и в конкретных физических и других науках, не принадлежат алгебре Σ'.

Примеры. Суть примеров состоит в том, что мы приближаем события из σ-алгебры множествами из алгебры и используем при этом теорему о продолжении меры.

1. Чему равна вероятность, что при бесконечном числе бросаний монеты выпадут одни гербы, т.е. чему равна вероятность точки

$(1, 1, \ldots)$? Рассмотрим множества

$$A_n = \{\omega : x_1 = 1, \ldots, x_n = 1\}.$$

Очевидно, они образуют вложенную последовательность, а их пересечение их есть интересующее нас элементарное событие: $A_{n+1} \subset A_n$ и $\cap A_n = (1, 1, \ldots)$. По аксиоме непрерывности получаем, что

$$\mathsf{P}\{(1, 1, \ldots)\} = \lim P(A_n) = \lim p^n = 0.$$

Точно так же, если p не равно ни нулю, ни единице, вероятность любой точки ω равна нулю.

2. Чему равна вероятность события B, что выпадет конечное число единиц? Снова нулю. В самом деле, рассмотрим множества $B_n = \{\omega : x_i = 0 \ \forall i > n\}$. Событие B_n состоит из конечного числа точек, поэтому его вероятность равна нулю. Множества B_n образуют расширяющуюся последовательность и в объединении дают B, поэтому $\mathsf{P}(B) = \lim \mathsf{P}(B_n) = 0$.

3. Вероятность того, что ни разу не встретятся три нуля подряд, равна нулю при $q \neq 1$. Для доказательства нам понадобится понятие независимости.

Определение 7. События $A, B \in \Sigma$ называются *независимыми*, если

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

События $A_1, A_2, \ldots, A_n, \ldots$ называются *независимыми в совокупности,* если события, принадлежащие любой *конечной подсистеме* этой системы, независимы в совокупности, т.е. для любых $i_1, \ldots, i_m$

$$\mathsf{P}(A_{i_1} \cap \ldots \cap A_{i_m}) = \mathsf{P}(A_{i_1}) \ldots \mathsf{P}(A_{i_m}).$$

Случайные величины $\xi_1, \ldots, \xi_n, \ldots$ называются взаимно независимыми, или *независимыми в совокупности*, если принадлежащие любому *конечному набору* случайные величины взаимно независимы, то есть для любых $i_1, \ldots, i_m$ и любых $a_1, \ldots, a_m$ события

$$\{\xi_{i_1} < a_1\}, \ldots, \{\xi_{i_m} < a_m\}$$

независимы в совокупности.

Замечание 4. Если случайные величины принимают конечное или счетное число значений, в определении независимости достаточно потребовать, чтобы

$$\mathsf{P}(\xi_{i_1} = a_1, \ldots, \xi_{i_n} = a_n) = \mathsf{P}(\xi_{i_1} = a_1) \ldots \mathsf{P}(\xi_{i_n} = a_n).$$

Обозначим C_k событие, состоящее в том, что среди координат последовательности ω с номерами $k, k-1, k-2$ имеется хотя бы одна единица. Мы хотим вычислить вероятность события $C = \cap_{k=3}^{\infty} C_k$. Рассмотрим событие D, что три нуля подряд не встретятся на местах $3k, 3k-1, 3k-2$ ни для какого целого положительного k. Очевидно, $D = \cap_{k=1}^{\infty} C_{3k}$ и $D \supset C$. Подсчитаем вероятность события D. Обозначим $D_n = \cap_{k=1}^{n} C_{3k}$. Тогда $D = \cap_{n=1}^{\infty} D_n$. Прямым вычислением можно проверить, что события $C_{3k}, k = 1, \ldots, n$ независимы в совокупности. Поскольку $\mathsf{P}(C_k) = 1 - q^3$ для любого k, то $\mathsf{P}(D_n) = (1-q^3)^n$. По свойству непрерывности получаем, что $\mathsf{P}(D) = \lim \mathsf{P}(D_n) = 0$. Отсюда следует, что $\mathsf{P}(C) = 0$.

Литература

Колмогоров А. Н. Основные понятия теории вероятностей. М., Наука, 1976. Ширяев А. Н. Вероятность. М., Наука, 1980. Глава 1, §§1 – 4, глава 2, §§1 – 4.

Лекция 3

Закон больших чисел для схемы Бернулли

В прошлой лекции мы построили схему Бернулли с бесконечными сериями испытаний. В этой лекции мы изучим некоторые ее предельные свойства.

На пространстве Ω бесконечных серий зададим систему функций

$$\hat{\xi}_i(\omega) = 2x_i - 1,$$

где x_i есть значение i-й координаты последовательности ω. Таким образом, $\hat{\xi}_i(\omega) = 1$, если в i-м испытании серии ω наступил успех и -1, если в i-м испытании неудача.

Упражнение 1. Доказать, что $\hat{\xi}_i$ взаимно независимы.

Обозначим

$$\hat{S}_n = \hat{S}_n(\omega) = \sum_{i=1}^{n} \hat{\xi}_i(\omega).$$

Тогда, очевидно,

$$\hat{S}_n = 2S_n - n,$$

где S_n есть число успехов в n испытаниях Бернулли.

Определение 1 (Сходимость по вероятности). Последовательность случайных величин $\eta_n(\omega)$, заданных на одном вероятностном пространстве $(\Omega, \Sigma, \mathsf{P})$, сходится к случайной величине $\eta(\omega)$ *по вероятности*, если для любого $\delta > 0$

$$\mathsf{P}(\omega : \mid \eta_n(\omega) - \eta(\omega) \mid > \delta) \to 0 \quad \text{при } n \to \infty.$$

Теорема 1 (Закон больших чисел для схемы Бернулли). Последовательность $\hat{S}_n/n$ сходится по вероятности к константе $p - q$.

Доказательство было дано в лекции 1 в терминах случайных величин S_n.

Определение 2 (Сходимость с вероятностью 1). Последовательность случайных величин $\eta_n(\omega)$, заданных на одном вероятностном пространстве $(\Omega, \Sigma, \mathsf{P})$, сходится к случайной величине $\eta(\omega)$ *с вероятностью* 1, если $\eta_n(\omega) \to \eta(\omega)$ для всех ω, кроме множества, имеющего нулевую вероятность (множества меры нуль). Такая сходимость назывется также сходимостью *почти наверное* или сходимостью *почти всюду* (и сокращенно записывается п.н. или п.в.).

Теорема 2 (Усиленный закон больших чисел для схемы Бернулли). Последовательность $\hat{S}_n/n$ сходится с вероятностью 1 к константе $p-q$.

Доказательство опирается на лемму об экспоненциальных оценках.

Лемма 1 (экспоненциальные оценки). Для любого $\delta > 0$ существует $\varepsilon > 0$ такое, что

$$\mathsf{P}\{\mid \hat{S}_n - (p-q)n \mid > \delta n\} < 2\exp(-\varepsilon n).$$

Доказательство. Обозначим $m = p - q$. Представим интересующее нас событие в виде объединения двух событий:

$$\{\omega : \mid \hat{S}_n - mn \mid > \delta n\} = \{\omega :\ \hat{S}_n - mn > \delta n\} \cup \{\omega :\ \hat{S}_n - mn < -\delta n\}$$

и воспользуемся тем, что вероятность объединения событий не превосходит суммы вероятностей этих событий. Оценим вероятность первого события:

$$\mathsf{P}\{\hat{S}_n - mn > \delta n\} = \mathsf{P}\Big\{h\sum_{i=1}^{n}(\hat{\xi}_i - m) > h\delta n\Big\} =$$

$$= \mathsf{P}\Big\{\exp\Big(h\sum_{i=1}^{n}(\hat{\xi}_i - m)\Big) > \exp(h\delta n)\Big\} \le$$

$$\le \exp(-h\delta n)\mathsf{E}\exp\Big(h\sum_{i=1}^{n}(\hat{\xi}_i - m)\Big).$$

Здесь h — произвольная положительная константа, выбор которой мы осуществим позже, а последнее неравенство есть неравенство Чебышева. Случайные величины $\exp(h(\hat{\xi}_i - m))$ независимы в силу независимости случайных величин $\hat{\xi}_i$, поэтому (см. предложение 2 лекции 1)

$$\mathsf{E}\exp\Big(h\sum_{i=1}^{n}(\hat{\xi}_i - m)\Big) = \prod \mathsf{E}\exp(h(\hat{\xi}_i - m)).$$

Возьмем h столь малым, чтобы $\mid h(\hat{\xi}_i - m) \mid < 1$ и воспользуемся неравенством:

$$e^x < 1 + x + \frac{3}{2}x^2 \quad \text{при} \quad \mid x \mid < 1.$$

Получим:

$$\prod \mathsf{E} e^{h(\xi_i - m)} < \prod \Big[1 + h\mathsf{E}(\xi_i - m) + \frac{3}{2}h^2\mathsf{E}(\xi_i - m)^2\Big].$$

Второе слагаемое в правой части равно нулю, а третье не превосходит $c_1 h^2$ для некоторого c_1 в силу ограниченности ξ_i. Поэтому

$$\prod \mathsf{E} e^{h(\xi_i - m)} < \prod (1 + c_2 h^2)^n,$$

и окончательно получаем

$$\mathsf{P}\{\hat{S}_n - mn > \delta n\} < \exp(-n(\delta h - \ln(1 + c_2 h^2))) \leq \exp(-n\varepsilon),$$

где $\varepsilon = \delta h - \ln(1 + c_2 h^2) > 0$ для достаточно малых h. Аналогичным образом оценивается вероятность второго множества.

Доказательство усиленного закона больших чисел. Рассмотрим множество $B = \{\omega : \hat{S}_n(\omega)/n$ не стремится к $m\}$. Пусть $\omega \in B$. Это означает, что найдется $\delta > 0$ такое, что для любого k найдется $n > k$ такое, что $\mid \hat{S}_n/n \mid > \delta$. Мы хотим доказать, что $P(B) = 0$. Для этого достаточно показать, что для любого фиксированного δ вероятность события $B_\delta = \cap_{N=1}^{\infty} \cup_{n=N}^{\infty} A_{n,\delta}$ равна нулю, где

$$A_{n,\delta} = \left\{\omega : \left|\frac{\hat{S}_n(\omega)}{n}\right| > \delta\right\}.$$

Говорят, что событие B_δ состоит в том, что произошло бесконечное число событий $A_{n,\delta}$. Рассмотрим событие $D_N = \cup_{n=N}^{\infty} A_{n,\delta}$. Оценим его вероятность с помощью леммы об экспоненциальных оценках:

$$\mathsf{P}(D_N) \leq \sum_{n=N}^{\infty} P(A_{n,\delta}) \leq 2\sum_{n=N}^{\infty} \exp(-\varepsilon n) \leq C\exp(-\varepsilon N),$$

где константы C и ε не зависят от N. Поскольку $B_\delta = \cap_{N=1}^{\infty} D_N$, в силу свойства непрерывности

$$\mathsf{P}(B_\delta) = \lim_N \mathsf{P}(D_N) = 0.$$

Замечание 1. Разница между законом больших чисел и усиленным законом больших чисел состоит в следующем. Закон больших чисел утверждает, что вероятность каждого из множеств $A_{n,\delta}$ мала, начиная с некоторого n. Усиленный закон больших чисел гласит, что вероятность объединения $\cup_{n=N}^{\infty} A_{n,\delta}$ стремится к нулю при $N \to \infty$.

Закон нуля и единицы

Аналогично тому, как мы ввели схему Бернулли на пространстве бесконечных в одну сторону последовательностей из 0 и 1, можно ввести схему Бернулли на пространстве функций, принимающих значения 0 и 1 на любом счетном множестве (просто пронумеровав его натуральными числами). В частности, мы можем это сделать на множестве всех целых чисел Z. Получим бесконечные в обе стороны последовательности из нулей и единиц, где каждая координата независима от остальных и равна 0 с вероятностью q и 1 с вероятностью p. Обозначим соответствующее вероятностное пространство $(\Omega = \Omega(Z), \Sigma, \mathsf{P})$. Рассмотрим отображение сдвига $\theta : \Omega \to \Omega$, при котором последовательность просто сдвигается вправо: если $\omega = (x_k)_{k=-\infty}^{\infty}$ и $\theta(\omega) = (y_k)_{k=-\infty}^{\infty}$, то $y_k = x_{k-1}$ для всех k. Сдвиг множества определяется как сдвиг всех его элементов: $\theta A = \{\theta\omega : \omega \in A\}$. Сдвиг есть взаимно-однозначное преобразование в себя и "сохраняет меру", т.е. $\mathsf{P}(A) = \mathsf{P}(\theta A) = \mathsf{P}(\theta^{-1} A)$.

Определение 3. Событие A называется *инвариантным,* если $\theta A = A$.

Можно привести множество примеров инвариантных событий: точка $\omega = (\ldots, 1, 1, \ldots)$, т.е. последовательность из одних единиц; множество последовательностей, в которых на каждом интервале число единиц не превышает α процентов от длины интервала и т.д. Закон нуля и единицы утверждает следующее.

Теорема 3. В схеме Бернулли каждое инвариантное событие имеет вероятность 0 или 1.

Доказательство. Назовем событие *локальным*, если оно принадлежит алгебре Σ_m для некоторого m. Убедимся сначала, что утверждение теоремы верно для локальных событий. Пусть A локально и имеет меру $0 < \beta < 1$. Тогда для достаточно большого n события A и $\theta^n A$ независимы, так что $P(A \cap \theta^n A) = \beta^2$. Но, с другой стороны, $A = \theta A$ по предположению, так что $P(A \cap \theta^n A) =$

$P(A) = P(\theta^n A) = \beta$, что приводит к противоречию.

Для завершения доказательства достаточно показать, что для любого множества $B \in \Sigma$ и любого как угодно малого $\varepsilon > 0$ существует локальное событие A такое, что $P(A \triangle B) < \varepsilon$ и воспользоваться вышеприведенным рассуждением с точностью до ε. Здесь $A \triangle B$ есть симметрическая разность множеств A и B : $A \triangle B = (A \setminus B) \cup (B \setminus A)$. Нам понадобится следующая лемма, дающая представление о том, как можно построить σ-алгебру, содержащую заданную систему множеств.

Лемма 2. Пусть дана произвольная алгебра $\mathcal{F}$ подмножеств Ω. Необходимым и достаточным условием того, что $\mathcal{F}$ есть σ-алгебра, является ее замкнутость относительно взятия монотонного предела: для любой последовательности A_n вложенных множеств, принадлежащих $\mathcal{F}$ (т.е. $A_{n+1} \subset A_n$), следует, что $A = \cap_n A_n \in \mathcal{F}$; и для любой последовательности расширяющихся множеств $A_n \in \mathcal{F}$ (т.е. $A_{n+1} \supset A_n$) следует, что $A = \cup_n A_n \in \mathcal{F}$.

Доказательство. Каждая σ-алгебра, очевидно, обладает указанным свойством. Пусть теперь $\mathcal{F}$ замкнуто относительно взятия монотонного предела. Покажем, что $\mathcal{F}$ является σ-алгеброй. Пусть задана произвольная последовательность множеств B_n из $\mathcal{F}$. Так как $\mathcal{F}$ алгебра, то для любого n $A_n = \cup_{i=1}^n B_i \in \mathcal{F}$. Но $A_n \subseteq A_{n+1}$. Поэтому $\cup_{n=1}^\infty A_n = \cup_{i=1}^\infty B_i \in \mathcal{F}$. Аналогично можно показать, что $\cap_{i=1}^\infty B_i \in \mathcal{F}$.

Вернемся к доказательству теоремы 3. Обозначим $\mathcal{F}$ систему множеств такую, что для любого ε и для любого $B \in \mathcal{F}$ существует локальное событие A такое, что $P(A \triangle B) < \varepsilon$. Очевидно, алгебра Σ' локальных множеств принадлежит $\mathcal{F}$. Покажем, что $\mathcal{F}$ замкнуто относительно взятия монотонного предела. Пусть B_n есть последовательность расширяющихся множеств, принадлежащих $\mathcal{F}$. Обозначим $B = \cup B_n$. Последовательность $P(B_n)$ монотонно возрастает и ограничена сверху, поэтому для заданного ε существует n такое, что $P(B \setminus B_n) < \varepsilon/2$. Возьмем локальное множество A_n такое, что $P(A_n \triangle B_n) < \varepsilon/2$. Тогда

$$\mathsf{P}(B \triangle A_n) = \mathsf{P}(B \setminus A_n) + \mathsf{P}(A_n \setminus B) \leq$$

$$\leq \mathsf{P}((B_n \setminus A_n) \cup ((B \setminus B_n) \setminus A_n)) + \mathsf{P}(A_n \setminus B_n) \leq$$

$$\leq \mathsf{P}(B_n \setminus A_n) + \varepsilon/2 + \mathsf{P}(A_n \setminus B_n) \leq \mathsf{P}(A_n \triangle B_n) + \varepsilon/2 \leq \varepsilon.$$

Таким образом, $B \in \mathcal{F}$. Аналогично доказывается, что монотонный предел B последовательности убывающих множеств $B_n \downarrow B$ принадлежит $\mathcal{F}$. Согласно лемме 2 $\mathcal{F}$ есть σ-алгебра, но Σ является наименьшей σ-алгеброй, содержащей локальные множества, поэтому $\mathcal{F} = \Sigma$.

Итак, для любого $A \in \Sigma$ и любого ε найдется локальное B такое, что $P(A \triangle B) < \varepsilon$. Нам понадобятся три свойства симметрической разности, которые легко доказываются непосредственно.

Упражнение 2. Доказать следующие свойства.
1. Для любых C и D

$$| \mathsf{P}(C) - \mathsf{P}(D) | \leq \mathsf{P}(C \triangle D).$$

2. Сдвиг симметрической разности множеств есть симметрическая разность сдвигов этих множеств:

$$\theta(C \triangle D) = (\theta C) \triangle (\theta D).$$

3. Для любых A_1, A_2, B_1, B_2

$$(A_1 \cap A_2) \triangle (B_1 \cap B_2) \subseteq (A_1 \triangle B_1) \cup (A_2 \triangle B_2).$$

Рассмотрим произвольное инвариантное множество A. Для заданного ε выберем локальное B такое, что $\mathsf{P}(A \triangle B) < \varepsilon$. Так как преобразование θ сохраняет меру, в силу пункта 2 упражнения 2

$$\mathsf{P}((\theta^n A) \triangle (\theta^n B)) < \varepsilon. \tag{3.1}$$

Воспользуемся далее пунктом 3 упражнения 2:

$$\mathsf{P}((A \cap \theta^n A) \triangle (B \cap \theta^n B)) \leq \mathsf{P}(A \triangle B) + \mathsf{P}((\theta^n A) \cap (\theta^n B)) \leq 2\varepsilon. \tag{3.2}$$

Из (3.2) и пункта 1 упражнения 2 следует, что

$$| \mathsf{P}(A \cap \theta^n A) - \mathsf{P}(B \cap \theta^n B) | \leq 2\varepsilon. \tag{3.3}$$

Но B локальное множество, поэтому для достаточно больших n события B и $\theta^n B$ независимы, а так как θ сохраняет меру, то

$$\mathsf{P}(B \cap \theta^n B) = \mathsf{P}^2(B).$$

С другой стороны, в силу инвариантности A

$$\mathsf{P}(A \cap \theta^n A) = \mathsf{P}(A),$$

поэтому из (3.3) получаем

$$| \mathsf{P}(A) - \mathsf{P}^2(B) | \leq 2\varepsilon. \tag{3.4}$$

Заметим далее, что

$$| \mathsf{P}^2(B) - \mathsf{P}^2(A) | \leq 2 \mid \mathsf{P}(B) - \mathsf{P}(A) | \leq 2\varepsilon.$$

Подставляя последнее неравенство в (3.4), получим

$$| \mathsf{P}(A) - \mathsf{P}^2(A) | \leq 4\varepsilon,$$

откуда следует, что $\mathsf{P}(A) = 0$ либо 1.

Замечание 2. Изучением сохраняющих меру преобразований занимается эргодическая теория динамических систем.

Литература

Ширяев А. Н. Вероятность, глава 1, §§5, 6. Феллер В. Введение в теорию вероятностей и ее приложения, том 1, глава 6, §§1 − 4; глава 8, §§1 − 4.

Лекция 4

Интеграл Лебега. Математическое ожидание

Пусть имеется измеримое пространство (Ω, Σ, μ) со счетно-аддитивной конечной мерой. В лекции 2 мы определили математическое ожидание (интеграл Лебега) для кусочно-постоянных (простых) случайных величин. Определим теперь интеграл Лебега для произвольной неотрицательной измеримой функции.

Определение 1. Если $f(\omega) \geq 0$, то

$$\int\limits_{\Omega} f(\omega) d\mu(\omega) = \sup \int\limits_{\Omega} \varphi(\omega) d\mu(\omega), \tag{4.1}$$

где sup берется по всем кусочно-постоянным функциям φ таким, что $\varphi(\omega) \leq f(\omega)$. Далее, любую функцию $f(\omega)$ можно представить в виде

$$f(\omega) = f^{+}(\omega) - f^{-}(\omega),$$

где $f^{+}(\omega) = f(\omega)$, если $f(\omega) \geq 0$, и 0 в противном случае; $f^{-}(\omega) = -f(\omega)$, если $f(\omega) < 0$, и 0 в противном случае. Тогда определим

$$\int\limits_{\Omega} f(\omega) d\mu(\omega) = \int\limits_{\Omega} f^{+}(\omega) d\mu(\omega) - \int\limits_{\Omega} f^{-}(\omega) d\mu(\omega),$$

если хотя бы один из интегралов в правой части не равен ∞. В случае, если оба бесконечны, интеграл не определяется.

Замечание 1. Аналогично определяется интеграл по любому подмножеству $B \subset \Omega$, $B \in \Sigma$; нужно лишь взять ограничение кусочно-постоянных функций на множество B.

Определение 2. Функция f называется *интегрируемой по Лебегу*, если ее интеграл Лебега существует и конечен. Это означает, что

$$\int\limits_{\Omega} f^{+}(\omega) d\mu(\omega) < \infty \quad \text{и} \quad \int\limits_{\Omega} f^{-}(\omega) d\mu(\omega) < \infty,$$

то есть

$$\int\limits_{\Omega} |f| d\mu < \infty.$$

Упражнение 1. Показать, что непосредственно из определения вытекают следующие свойства интеграла Лебега.

1. Если $0 \le f(\omega) \le g(\omega)$ при всех ω, то

$$\int\limits_{\Omega} f d\mu \le \int\limits_{\Omega} g d\mu.$$

2. Для любой константы c

$$\int\limits_{\Omega} c f d\mu = c \int\limits_{\Omega} f d\mu.$$

3. Для любых интегрируемых $f_1,\ f_2$

$$\int\limits_{\Omega} (f_1 + f_2)\, d\mu = \int\limits_{\Omega} f_1\, dP + \int\limits_{\Omega} f_2\, d\mu.$$

4. $$\Big| \int\limits_{\Omega} f d\mu \Big| \le \int\limits_{\Omega} |f| d\mu.$$

5. Если $B \subset \Omega$ и $\mu(B) = 0$, то

$$\int\limits_{B} f d\mu = 0.$$

6. Если $f \ge 0$ и $\int\limits_{\Omega} f d\mu = 0$, то $f = 0$ почти всюду.

Определение 3. *Математическим ожиданием* $\mathsf{E}\xi$ случайной величины ξ, заданной на вероятностном пространстве $(\Omega, \Sigma, \mathsf{P})$, называется ее интеграл Лебега по мере P :

$$\mathsf{E}\xi = \int\limits_{\Omega} \xi(\omega)\, d\mathsf{P}.$$

Определение 4. Случайная величина, принимающая с положительной вероятностью значения $+\infty$ или $-\infty$, называется *расширенной.*

Определение интеграла Лебега очевидным образом распространяется и на расширенные случайные величины. Часто бывает полезна следующая лемма.

Лемма 1. Для любой (в том числе и расширенной) неотрицательной случайной величины ξ найдется последовательность простых (т.е. кусочно-постоянных) случайных величин $\xi_1, \xi_2, \ldots$ таких, что $\xi_n(\omega) \uparrow \xi(\omega)$ для каждого ω при $n \to \infty$.

Доказательство. Предъявим функции $\xi_n(\omega)$ в явном виде:

$$\xi_n(\omega) = \sum_{k=1}^{n2^n} \frac{k-1}{2^n} I_{A_{k,n}}(\omega) + nI_{\{\xi(\omega) \geq n\}}(\omega),$$

где

$$A_{k,n} = \left\{ \omega : \frac{k-1}{2^n} \leq \xi(\omega) < \frac{k}{2^n} \right\},$$

а I_A есть индикатор события A :

$$I_A(\omega) = \begin{cases} 1, & \text{если } \omega \in A; \\ 0 & \text{в противном случае.} \end{cases}$$

Лемма 1 позволяет дать эквивалентное определение математического ожидания случайной величины. Пусть $\xi = \xi(\omega)$ — неотрицательная случайная величина. Построим последовательность простых неотрицательных случайных величин $\{\xi_n\}$ таких, что $\xi_n(\omega) \uparrow \xi(\omega)$ при $n \to \infty$ для каждого $\omega \in \Omega$ (см. лемму 1). В силу определения математического ожидания простой случайной величины $\mathsf{E}\xi_n \leq \mathsf{E}\xi_{n+1}$, если $\xi_n \leq \xi_{n+1}$. Поэтому существует предел $\lim_n \mathsf{E}\xi_n$, возможно, равный $+\infty$.

Определение 5. Интегралом Лебега от неотрицательной случайной величины ξ, или ее математическим ожиданием, называется величина

$$\mathsf{E}\xi = \lim_n \mathsf{E}\xi_n,$$

где последовательность ξ_n такая, что $\xi_n(\omega) \uparrow \xi(\omega)$.

Для определения математического ожидания случайной величины общего вида нужно, как и раньше, представить ее в виде разности двух неотрицательных случайных величин: $\xi = \xi^+ - \xi^-$.

Определение 5 корректно, т.е. $\mathsf{E}\xi$ не зависит от выбора аппроксимирующей последовательности ξ_n. Кроме того, определения 1 и 5 эквивалентны. Это вытекает из следующей ниже теоремы 1, которая является вариантом теоремы о монотонной сходимости. Для формулировки теоремы 1 нам понадобятся некоторые сведения из теории меры.

Лемма 2 (из теории меры). Предел сходящейся при каждом ω последовательности случайных величин $\xi_n(\omega)$, заданных на одном вероятностном пространстве $(\Omega, \Sigma, \mathsf{P})$, является случайной величиной.

Доказательство. Покажем вначале, что $\sup \xi_n$ и $\inf \xi_n$ являются случайными величинами. Это следует из тождеств

$$\{\omega :\ \sup \xi_n > a\} = \cup_n \{\omega :\ \xi_n > a\} \in \Sigma,$$

$$\{\omega :\ \inf \xi_n < a\} = \cup_n \{\omega :\ \xi_n < a\} \in \Sigma.$$

Далее, поскольку

$$\limsup \xi_n = \inf_n \sup_{m \geq n} \xi_m, \quad \liminf \xi_n = \sup_n \inf_{m \geq n} \xi_m,$$

то $\limsup \xi_n$ и $\liminf \xi_n$ также являются случайными величинами. Обозначим $\xi(\omega) = \lim_{n \to \infty} \xi_n(\omega)$. Тогда

$$\begin{aligned}\{\omega :\ \xi(\omega) < a\} &= \{\omega :\ \lim \xi_n(\omega) < a\} = \\ = \{\omega :\ \limsup \xi_n(\omega) = \liminf \xi_n\} &\cap \{\omega :\ \limsup \xi_n(\omega) < a\} = \\ &= \Omega \cap \{\omega :\ \limsup \xi_n(\omega) < a\} \in \Sigma.\end{aligned}$$

Следствие 1. Пусть последовательность случайных величин $\xi_n(\omega)$ сходится к функции $\xi(\omega)$ почти всюду на Ω. Тогда $\xi(\omega)$ есть случайная величина.

Доказательство. Обозначим через A множество тех ω, для которых $\xi_n(\omega)$ не сходится к $\xi(\omega)$. По условию, $A \in \Sigma$ и $\mathsf{P}(A) = 0$. Нам удобно считать, что для любого события B такого, что $\mathsf{P}(B) = 0$, все его подмножества также являются событиями и, следовательно, также имеют нулевую вероятность. (В противном случае мы пополним Σ такими множествами.) Следовательно, для любого a

$$\{\omega :\ \xi(\omega) < a\} = (\{\omega :\ \xi(\omega) < a\} \cap (\Omega \setminus A)) \cup (\{\omega :\ \xi(\omega) < a\} \cap A) \in \Sigma.$$

Замечание 2. Из следствия 1 видно, что предел $\xi(\omega)$ сходящейся почти всюду последовательности случайных величин может быть определен с точностью до множества меры нуль: значения $\xi(\omega)$ на множестве A не играют роли. Таким образом, в вероятностных рассуждениях случайную величину можно заменить другой, отличающейся от исходной на множестве меры нуль. Вообще, если какое-либо свойство выполнено для всех ω за исключением, быть может, ω, принадлежащих множеству, имеющему нулевую вероятность, будем говорить, что это свойство выполнено *почти всюду* на Ω или *для почти всех* ω или *почти наверное* и кратко записывать это так: п.в. или п.н.

Теорема 1. Если функция $f(\omega)$ есть для почти всех ω сумма ряда из неотрицательных кусочно-постоянных функций $f_n(\omega)$, т.е. если

$$f(\omega) = \sum_{n=1}^{\infty} f_n(\omega) \quad \text{п.в.},$$

то

$$\int\limits_{\Omega} f(\omega) d\mathsf{P}(\omega) = \sum_{n=1}^{\infty} \int\limits_{\Omega} f_n(\omega) d\mathsf{P}(\omega).$$

Для доказательства теоремы 1 нам понадобится свойство счетной аддитивности интеграла Лебега.

Лемма 3 (из теории меры). Пусть $f \geq 0$ на Ω. Определим функцию Φ на элементах σ-алгебры Σ следующим образом: для любого $B \in \Sigma$ положим

$$\Phi(B) = \int\limits_{B} f d\mathsf{P}.$$

Тогда функция Φ счетно-аддитивна на Σ.

Доказательство. Пусть $B_n \in \Sigma$ — произвольная последовательность попарно непересекающихся множеств и $\cup_{n=1}^{\infty} B_n = B$. Нам нужно доказать, что

$$\Phi(B) = \sum_{n=1}^{\infty} \Phi(B_n). \tag{4.2}$$

Если f есть индикатор некоторого множества $A \in \Sigma$, т.е.

$$f(\omega) = \begin{cases} 1, & \text{если } \omega \in A, \\ 0, & \text{если } \omega \notin A, \end{cases}$$

то

$$\int\limits_{B} f d\mathsf{P} = \mathsf{P}(A \cap B),$$

и счетная аддитивность функции Φ есть не что иное, как счетная аддитивность вероятности P. Далее, если f кусочно-постоянна, то легко проверить, что утверждение леммы 3 также выполняется. Пусть теперь f произвольная неотрицательная измеримая функция. Для любой кусочно-постоянной φ такой, что $0 \leq \varphi \leq f$, имеем:

$$\int\limits_B \varphi d\mathsf{P} = \sum_{n=1}^{\infty} \int\limits_{B_n} \varphi d\mathsf{P} \leq \sum_{n=1}^{\infty} \Phi(B_n).$$

Поэтому, согласно определению интеграла,

$$\Phi(B) \leq \sum_{n=1}^{\infty} \Phi(B_n). \tag{4.3}$$

Заметим теперь, что если $\Phi(B_n) = \infty$ для какого-нибудь n, то поскольку $\Phi(B) \geq \Phi(B_n)$, (4.2) очевидным образом выполнено. Поэтому будем считать, что $\Phi(B_n) < \infty$ для всех n. Тогда для любого сколь угодно малого $\varepsilon > 0$ найдется кусочно-постоянная функция φ такая, что

$$\int\limits_{B_1} \varphi d\mathsf{P} \geq \int\limits_{B_1} f d\mathsf{P} - \varepsilon, \quad \int\limits_{B_2} \varphi d\mathsf{P} \geq \int\limits_{B_2} f d\mathsf{P} - \varepsilon. \tag{4.4}$$

Поэтому

$$\Phi(B_1 \cup B_2) = \int\limits_{B_1 \cup B_2} f d\mathsf{P} \geq \int\limits_{B_1 \cup B_2} \varphi d\mathsf{P} =$$

$$= \int\limits_{B_1} \varphi d\mathsf{P} + \int\limits_{B_2} \varphi d\mathsf{P} \geq \Phi(B_1) + \Phi(B_2) - 2\varepsilon,$$

а значит

$$\Phi(B_1 \cup B_2) \geq \Phi(B_1) + \Phi(B_2).$$

Точно так же последовательно при каждом n получаем

$$\Phi(B_1 \cup \ldots \cup B_n) \geq \Phi(B_1) + \ldots + \Phi(B_n). \tag{4.5}$$

Поскольку $B \supset B_1 \cup \ldots \cup B_n$, то из (4.5) следует, что

$$\Phi(B) \geq \sum_{n=1}^{\infty} \Phi(B_n). \tag{4.6}$$

Теперь (4.2) вытекает из (4.3) и (4.6).

Доказательство теоремы 1. Рассмотрим функции

$$F_N = \sum_{n=1}^{N} f_n(\omega).$$

Очевидно, F_N кусочно-постоянны, $F_N \geq 0$ и $F_N(\omega) \uparrow f(\omega)$ почти всюду. Интеграл Лебега от кусочно-постоянной функции не зависит, очевидно, от ее представления, т.е.

$$\int\limits_{\Omega} F_N(\omega) d\mathsf{P}(\omega) = \sum_{n=1}^{N} \int\limits_{\Omega} f_n(\omega) d\mathsf{P}(\omega).$$

В силу монотонности последовательности $\int F_N d\mathsf{P}$ существует предел

$$\lim_{N\to\infty} \int\limits_{\Omega} F_N d\mathsf{P} = \sum_{n=1}^{\infty} \int\limits_{\Omega} f_n(\omega) d\mathsf{P}(\omega). \tag{4.7}$$

Поэтому

$$\int\limits_{\Omega} f d\mathsf{P} \geq \sup_N \int\limits_{\Omega} F_N d\mathsf{P} = \sum_{n=1}^{\infty} \int\limits_{\Omega} f_n(\omega) d\mathsf{P}(\omega). \tag{4.8}$$

Возьмем произвольное $\varepsilon > 0$ и произвольную кусочно-постоянную функцию $\varphi(\omega) \geq 0$ такую, что $0 \leq \varphi \leq f$. Положим

$$B_N = \{\omega : \ F_N(\omega) > \varphi(\omega) - \varepsilon\}.$$

В силу монотонности последовательности $F_N(\omega)$ при каждом ω множества B_N образуют возрастающую последовательность:

$$B_1 \subset B_2 \subset \ldots \subset B_N \subset \ldots .$$

Для любого N

$$\int\limits_{\Omega} F_N \, d\mathsf{P} \geq \int\limits_{B_N} F_N \, d\mathsf{P} \geq \int\limits_{B_N} (\varphi - \varepsilon) d\mathsf{P} \geq$$

$$\geq \int\limits_{B_N} \varphi \, d\mathsf{P} - \varepsilon \mathsf{P}(B_N) \geq \int\limits_{B_N} \varphi \, d\mathsf{P} - \varepsilon. \tag{4.9}$$

Обозначим $B = \cup_{n=1}^{\infty} B_n$. По лемме 3 интеграл является счетно-аддитивной функцией множества. В силу эквивалентности свойств счетной аддитивности и непрерывности имеем:

$$\lim_{N\to\infty} \int\limits_{B_N} \varphi\, d\mathsf{P} = \int\limits_{B} \varphi\, d\mathsf{P}.$$

Так как

$$f(\omega) = \lim_{N\to\infty} F_N(\omega) \geq \varphi(\omega)$$

почти всюду, то $\mathsf{P}(B) = 1$. Поэтому по свойству 5 упражнения 1

$$\int\limits_{B} \varphi\, d\mathsf{P} = \int\limits_{\Omega} \varphi\, d\mathsf{P}. \tag{4.10}$$

Перейдем в (4.9) к пределу по N, учитывая (4.10):

$$\sum_{n=1}^{\infty} \int f_n\, d\mathsf{P} \geq \int\limits_{\Omega} \varphi\, d\mathsf{P} - \varepsilon. \tag{4.11}$$

В силу произвольности ε

$$\sum_{n=1}^{\infty} \int f_n\, d\mathsf{P} \geq \int\limits_{\Omega} \varphi\, d\mathsf{P},$$

и поскольку $\varphi \leq f$ — любая кусочно-постоянная функция, то

$$\sum_{n=1}^{\infty} \int f_n\, d\mathsf{P} \geq \int\limits_{\Omega} f\, d\mathsf{P}. \tag{4.12}$$

Теорема 1 следует теперь из (4.7), (4.8) и (4.12).

Замечание 3. Наряду с математическим ожиданием $\mathsf{E}\xi$ важными числовыми характеристиками случайной величины ξ являются величины $\mathsf{E}\xi^r$, называемые моментами r-го порядка.

Теорема 2. Пусть ξ и η — независимые случайные величины и пусть $\mathsf{E}|\xi| < \infty$ и $\mathsf{E}|\eta| < \infty$. Тогда $\mathsf{E}|\xi\eta| < \infty$ и

$$\mathsf{E}\xi\eta = \mathsf{E}\xi\mathsf{E}\eta.$$

Доказательство можно найти, например, в учебнике А.Н. Ширяева. Суть его состоит в том, что для простых функций утверждение проверяется непосредственно (мы проделали это в одной из предыдущих лекций), а затем случайная величина приближается простыми и совершается предельный переход под знаком интеграла Лебега.

Приведем теперь важнейшие неравенства для математических ожиданий, часто используемые в теории вероятностей. Первое из них — неравенство Чебышева — мы уже использовали для дискретных случайных величин.

Неравенство Чебышева. Пусть ξ — неотрицательная случайная величина. Тогда для любого $\varepsilon > 0$

$$\mathsf{P}(\xi \geq \varepsilon) \leq \frac{\mathsf{E}\xi}{\varepsilon}. \tag{4.13}$$

Доказательство. Обозначим $I_{\{\xi \geq \varepsilon\}}$ индикатор множества тех ω, для которых $\xi(\omega) \geq \varepsilon$. Тогда

$$\mathsf{E}\xi \geq \mathsf{E}[\xi I_{\{\xi \geq \varepsilon\}}] \geq \varepsilon \mathsf{P}(\xi \geq \varepsilon).$$

Из (4.13) получаем следующие разновидности неравенства Чебышева: если ξ — произвольная случайнаяя величина, то

$$\mathsf{P}(\xi \geq \varepsilon) \leq \frac{\mathsf{E}\xi^2}{\varepsilon^2}, \tag{4.14}$$

и

$$\mathsf{P}(|\xi - \mathsf{E}\xi| \geq \varepsilon) \leq \frac{\mathsf{D}\xi}{\varepsilon^2}, \tag{4.15}$$

где $\mathsf{D}\xi = \mathsf{E}(\xi - \mathsf{E}\xi)^2$ — дисперсия случайной величины ξ.

Неравенство Коши – Буняковского. Пусть случайные величины ξ и η таковы, что $\mathsf{E}\xi^2 < \infty$, $\mathsf{E}\eta^2 < \infty$. Тогда $\mathsf{E}|\xi\eta| < \infty$ и

$$(\mathsf{E}|\xi\eta|)^2 \leq \mathsf{E}\xi^2 \mathsf{E}\eta^2. \tag{4.16}$$

Доказательство. Предположим сначала, что $\mathsf{E}\xi^2 > 0$, $\mathsf{E}\eta^2 > 0$. Обозначим

$$\tilde{\xi} = \frac{\xi}{\sqrt{\mathsf{E}\xi^2}}, \quad \tilde{\eta} = \frac{\eta}{\sqrt{\mathsf{E}\eta^2}}.$$

Поскольку

$$2|\tilde{\xi}\tilde{\eta}| \leq \tilde{\xi}^2 + \tilde{\eta}^2,$$

то

$$2\mathsf{E}|\tilde{\xi}\tilde{\eta}| \leq \mathsf{E}\tilde{\xi}^2 + \mathsf{E}\tilde{\eta}^2 = 2,$$

т.е. $\mathsf{E}|\tilde{\xi}\tilde{\eta}| \le 1$, откуда и следует (4.16). В случае, если, например, $\mathsf{E}\xi^2 = 0$, то по свойству 6 интеграла Лебега $\xi = 0$ почти всюду. Следовательно, $\mathsf{E}\xi\eta = 0$, т.е. (4.16) также выполнено.

Неравенство Иенсена. Пусть $g = g(x)$ — выпуклая книзу борелевская функция и пусть ξ — случайная величина с $\mathsf{E}|\xi| < \infty$. Тогда

$$g(\mathsf{E}\xi) \le \mathsf{E}g(\xi). \tag{4.17}$$

Доказательство. Если функция $g = g(x)$ выпуклая книзу, то для каждого $x_0 \in \mathbf{R}$ найдется число $\lambda(x_0)$ такое, что для всех $x \in \mathbf{R}$

$$g(x) \ge g(x_0) + (x - x_0)\lambda(x_0). \tag{4.18}$$

Положим $x_0 = \mathsf{E}\xi$ и подставим $x = \xi(\omega)$ в (4.18). Получим

$$g(\xi) \ge g(\mathsf{E}\xi) + (\xi - \mathsf{E}\xi)\lambda(\mathsf{E}\xi)$$

и, следовательно, $\mathsf{E}g(\xi) \ge g(\mathsf{E}\xi)$.

Литература

Колмогоров А. Н., Фомин С. В. Элементы теории функций и функционального анализа. М., Наука, 1986, глава 5. Ширяев А. Н. Вероятность, глава 2, §6.

Лекция 5

Одномерное случайное блуждание

В лекции 3 мы рассматривали суммы $\hat{S}_n$ независимых случайных величин $\hat{\xi}_i$, принимающих значения $+1$ с вероятностью p и -1 с вероятностью $q = 1 - p$. Если считать индекс n временем, то можно установить взаимно однозначное соответствие между такими суммами (или схемой Бернулли) и одномерным случайным блужданием. А именно, рассмотрим частицу, находящуюся в начальный момент времени в нуле и совершающую в последующие (дискретные) моменты времени скачки по узлам одномерной целочисленной прямой. Если $\hat{\xi}_i = 1$, то в момент i совершается единичный скачок вправо, если же $\hat{\xi}_i = -1$, то в момент i совершается единичный скачок влево. Тогда в момент времени n координата блуждающей частицы равна $\hat{S}_n$.

Если на плоскости откладывать время n по горизонтальной оси, а по вертикальной оси откладывать положение частицы (т.е. $\hat{S}_n$), и соединить полученные точки отрезком, то получим ломаную линию (путь), которая называется траекторией случайного блуждания. Такое представление очень удобно для геометрического "видения" отдельных событий.

Точно так же можно определить блуждание, выходящее из точки $x \neq 0$. Положение частицы в момент времени n равно в этом случае $x + \hat{S}_n$.

В этой лекции мы познакомимся с некоторыми свойствами случайного блуждания.

Введем случайную величину η_k, принимающую значение 1, если в момент k случайное блуждание попадает в 0, и принимающую значение 0 в противном случае. В терминах схемы Бернулли событие $\eta_k = 1$ означает, что в серии длины k число успехов равно числу неудач. Очевидно, $\sum_{k=1}^{n} \eta_k$ есть (случайное) число попаданий в 0 за время n. Все случайные величины η_k определены на одном вероятностном пространстве $(\Omega, \Sigma, \mathsf{P})$, построенном в лекции 2. Для каждого $\omega \in \Omega$ определим функцию

$$\eta(\omega) = \sum_{k=1}^{\infty} \eta_k(\omega) \tag{5.1}$$

как предел последовательности $\sum_{k=1}^{n} \eta_k(\omega)$. Величина $\eta(\omega)$ равна, очевидно, числу попаданий блуждающей частицы в начало координат. Так как случайные величины η_k неотрицательны, то при каждом ω эта последовательность монотонно неубывает, поэтому существует ее предел, возможно, равный ∞. Тот факт, что η является (расширенной) случайной величиной, гарантируется леммой 2 из лекции 4.

Математическое ожидание $m_{00} = \mathsf{E}\eta$, то есть среднее число возвращений блуждающей частицы в начало координат, также определено по теореме 1 лекции 4.

Лемма 1. Если $p = q = \frac{1}{2}$, то $m_{00} = \infty$. Если $p \neq q$, то $m_{00} < \infty$.

Доказательство. По теореме 1 из лекции 4

$$m_{00} = \mathsf{E}\sum_{k=1}^{\infty} \eta_k = \sum_{k=1}^{\infty} \mathsf{E}\eta_k.$$

Так как случайная величина η_k принимает значения 0 и 1, то $\mathsf{E}\eta_k = \mathsf{P}(\eta_k = 1)$. Напомним, что $\mathsf{E}\hat{\xi}_i = p - q, \quad \mathsf{E}\hat{S}_n = n(p-q)$. При этом $\mathsf{P}(\eta_n = 1) = \mathsf{P}(\hat{S}_n = 0)$.

Пусть $p = q$. Воспользуемся локальной предельной теоремой для $\hat{S}_n$. При больших n

$$\mathsf{P}(\hat{S}_n = 0) \sim \sqrt{\frac{2}{\pi n}},$$

поэтому ряд для m_{00} расходится:

$$m_{00} = \sum_{n=1}^{\infty} \mathsf{E}\eta_n \geq \sum_{n=1}^{\infty} \frac{c}{\sqrt{n}} = \infty.$$

При $p \neq q$ ввиду экспоненциальных оценок (лемма 1 из лекции 3) для достаточно больших n

$$\mathsf{P}(\hat{S}_n = 0) \leq 2\exp(-\gamma n)$$

для некоторого $\gamma > 0$. Отсюда следует, что ряд для m_{00} сходится:

$$m_{00} = \sum_{n=1}^{\infty} \mathsf{E}\eta_n \leq \sum_{n=1}^{\infty} c\exp(-\gamma n) < \infty.$$

Но чему равна сама вероятность вернуться в нуль хотя бы один раз? Для ответа на этот вопрос нам понадобится понятие условной вероятности.

Определение 1. Предположим, что $\mathsf{P}(B) > 0$. Условная вероятность $\mathsf{P}(A \mid B)$ события A при условии, что совершилось событие B (кратко, условная вероятность A при условии B) определяется так:

$$\mathsf{P}(A \mid B) = \frac{\mathsf{P}(A \cap B)}{\mathsf{P}(B)} \tag{5.2}$$

Замечание 1. Если события A и B независимы, то $\mathsf{P}(A \mid B) = \mathsf{P}(A)$.

Утверждение 1 (Формула полной вероятности). Пусть задано разбиение пространства Ω на конечное или счетное число событий B_k, имеющих ненулевую вероятность, т.е. $\Omega = \cup_k B_k$, $B_i \cap B_j = \emptyset$, $\mathsf{P}(B_k) > 0$. Тогда вероятность произвольного события A может быть представлена в виде

$$\mathsf{P}(A) = \sum_k \mathsf{P}(A \mid B_k)\mathsf{P}(B_k). \tag{5.3}$$

Доказательство. Подставив (5.2) в правую часть (5.3) и воспользовавшись свойством аддитивности, получим тождество.

Пример. Вычислим условную вероятность того, что в момент времени n блуждание находится в точке x при условии, что в момент $n-1$ оно находилось в точке y. Поскольку $\mathsf{P}(\{\hat{S}_n = x\} \cap \{\hat{S}_{n-1} = y\}$ отлична от нуля лишь если $x = y + 1$ либо $x = y - 1$, то этим же свойством обладает и условная вероятность. Непосредственным вычислением легко получить, что

$$\mathsf{P}(\{\hat{S}_{n+1} = y + 1\} \cap \{\hat{S}_n = y\}) = p\mathsf{P}(\{\hat{S}_n = y\}).$$

Аналогично

$$\mathsf{P}(\{\hat{S}_{n+1} = y - 1\} \cap \{\hat{S}_n = y\}) = q\mathsf{P}(\{\hat{S}_n = y\}).$$

Поэтому окончательно имеем:

$$\mathsf{P}(\{\hat{S}_n = x\} \mid \{\hat{S}_{n-1} = y\}) = \begin{cases} p, & \text{если } x = y + 1; \\ q, & \text{если } x = y - 1; \\ 0 & \text{в остальных случаях.} \end{cases} \tag{5.4}$$

Обозначим $\hat{S}_n^x = x + \hat{S}_n$. Величина $\hat{S}_n^x$ равна положению блуждающей частицы в момент времени n, если в начальный момент времени частица находилась в точке x. Условные вероятности для $\hat{S}_n^x$ те же самые, т.е.

$$\mathsf{P}(\{\hat{S}_n^x = z\} \mid \{\hat{S}_{n-1}^x = y\}) = \begin{cases} p, & \text{если } z = y+1; \\ q, & \text{если } z = y-1; \\ 0 & \text{в остальных случаях.} \end{cases} \tag{5.5}$$

Обозначим через $p(x)$ вероятность, выйдя из точки $x > 0$, когда-либо вернуться в 0 :

$$p(x) = \mathsf{P}\{\omega : \exists n, \ \hat{S}_n^x(\omega) = 0\}.$$

Теорема 1.

$$p(x) = \begin{cases} 1 & \text{при } q \geq p, \\ \left(\frac{q}{p}\right)^x & \text{при } q < p. \end{cases} \tag{5.6}$$

Доказательство. Функция $p(x)$ удовлетворяет следующей системе уравнений при $x > 1$:

$$p(x) = pp(x+1) + qp(x-1). \tag{5.7}$$

Для доказательства (5.7) достаточно применить формулу полной вероятности, взяв в качестве разбиения пару множеств B_1, B_2, где B_1 соответствует тому, что на первом шаге блуждание попадает в точку $x+1$:

$$B_1 = \{\omega : \ \hat{S}_1^x(\omega) = x+1, \ \exists n, \ \hat{S}_n^x(\omega) = 0\};$$

а B_2 соответствует тому, что на первом шаге частица попадает в $x-1$:

$$B_2 = \{\omega : \ \hat{S}_1^x(\omega) = x-1, \ \exists n, \ \hat{S}_n^x(\omega) = 0\}.$$

При $x = 1$ первый же скачок может привести частицу в 0, поэтому

$$p(1) = q + pp(2). \tag{5.8}$$

Чтобы объединить уравнения (5.7) и (5.8), положим $p(0) = 1$, тогда $p(x)$ удовлетворяет (5.7) для всех $x > 0$.

Предложение 1. При $p \neq q$ произвольное решение системы (5.7) имеет вид

$$p(x) = c_1\lambda_1^x + c_2\lambda_2^x, \tag{5.9}$$

где $\lambda_1 = 1,\ \lambda_2 = q/p$ — два корня характеристического уравнения

$$1 = p\lambda + q\lambda^{-1}. \tag{5.10}$$

При $p = q$ характеристическое уравнение (5.10) имеет кратный корень 1. Решение системы (5.7) в этом случае имеет вид

$$p(x) = c_1 + c_2 x. \tag{5.11}$$

Доказательство. Будем искать решение системы (5.7) в виде геометрической прогрессии

$$p(x) = a\lambda^x.$$

Подстановка такого вида решения в уравнение приводит к квадратному уравнению (5.10). Если $p \neq q$, то корни этого уравнения различны: $\lambda_1 = 1,\ \lambda_2 = q/p$. Таким образом, мы нашли два линейно независимых решения (5.7): $p(x) = 1$ и $p(x) = (q/p)^x$.

Покажем, что все решения исчерпываются линейными комбинациями $c_1 + c_2(q/p)^x$ двух найденных. В самом деле, всякое решение $p(x)$ системы (5.7) однозначно определяется своими значениями $p(0)$ и $p(1)$. Обозначим $Y(x)$ и $Z(x)$ решения, удовлетворяющие условиям

$$\begin{aligned} Y(0) = 1,\ \ Y(1) = 0;\\ Z(0) = 0,\ \ Z(1) = 1. \end{aligned} \tag{5.12}$$

Произвольное решение $p(x)$, очевидно, является линейной комбинацией этих решений:

$$p(x) = p(0)Y(x) + p(1)Z(x).$$

Легко убедиться, что функции $Y(x)$ и $Z(x)$ в свою очередь представимы в виде (5.9):

$$\begin{aligned} Y(x) &= \frac{q}{q-p} - \frac{p}{q-p}\left(\frac{q}{p}\right)^x,\\ Z(x) &= -\frac{p}{q-p} + \frac{p}{q-p}\left(\frac{q}{p}\right)^x. \end{aligned} \tag{5.13}$$

В случае $p = q$ разностное уравнение принимает вид

$$p(x-1) - 2p(x) + p(x+1) = 0,$$

или

$$(p(x-1) - p(x)) - (p(x) - p(x+1)) = 0.$$

Мы видим, что разность $p(x-1)-p(x)$ постоянна, т.е. $p(x)$ — линейная функция, $p(x)=c_1+c_2x$. Заметим, что решения $Y(x)$ и $Z(x)$, удовлетворяющие (5.12), могут быть при $p=q$ представлены в виде линейных функций:

$$Y(x)=-x+1,\quad Z(x)=x,$$

откуда следует, что все решения исчерпываются линейными функциями.

Вернемся к доказательству теоремы.

I. Рассмотрим сначала случай $q>p$. Тогда в (5.9) $c_2=0$, так как $p(x)\le 1$. Но тогда $c_1=1$, так как $p(0)=1$ и поэтому $p(x)=1$ для всех $x\ge 0$.

II. При $p=q$ в силу $p(x)\le 1$ из (5.11) опять следует $c_2=0$, и $p(x)=1$ для всех $x>0$.

III. Если $q<p$, то из (5.9) получаем, что $c_1+c_2=1$, так как $p(0)=1$, и мы имеем континуум решений:

$$p(x)=c_1+(1-c_1)\Big(\frac{q}{p}\Big)^x. \tag{5.14}$$

Как выбрать истинное решение? Только одно из них стремится к нулю при $x\to\infty$. Ему соответствует $c_1=0$. Покажем, что оно и есть искомое. Для этого нам понадобятся следующие утверждения.

Предложение 2. Для любого $x\ge 1$

$$p(2x)=p^2(x). \tag{5.15}$$

Доказательство. В самом деле, все траектории блуждания, выходящие из $2x$ и попадающие в 0, проходят через точку x. Обозначим

$$A_m=\{\omega:\ \hat S_i^{2x}(\omega)\ne x \text{ для } i<m;\ \hat S_m^{2x}(\omega)=x,\ \ \exists n>m:\ \hat S_n^x(\omega)=0\}$$

событие, состоящее в том, что блуждающая частица, выйдя в начальный момент времени из $2x$, впервые попадет в x на шаге m, а затем когда-нибудь попадет в 0. Тогда, поскольку события A_m не пересекаются,

$$p(2x)=\mathsf{P}(\bigcup_{m\ge x}A_m)=\sum_{m\ge x}\mathsf{P}(A_m).$$

В силу независимости испытаний Бернулли, после попадания частицы в x последующая часть ее траектории не зависит от предшествующей части. Обозначим через $f^{(m)}(y,z)$ вероятность, выйдя из y, впервые попасть в z на шаге m. Тогда

$$\mathsf{P}(A_m) = f^{(m)}(2x,x)p(x).$$

Поэтому

$$p(2x) = p(x)\sum_m f^{(m)}(2x,x) = p(x)\sum_m f^{(m)}(x,0).$$

В последнем равенстве мы воспользовались тем, что вероятности траекторий блуждания инвариантны относительно сдвига по пространству (см. (5.4) и (5.5)). Но

$$\sum_m f^{(m)}(x,0) = p(x), \tag{5.16}$$

откуда и получаем утверждение (5.15).

Предложение 3. Либо $p(x) = 1$ при всех x, либо $p(x) \to 0$ при $x \to \infty$.

Доказательство. Функция $p(x)$ является монотонно невозрастающей. Действительно, если $x' > x$, то

$$p(x') = f(x',x)p(x) \le p(x),$$

где $f(x',x)$ есть вероятность когда-нибудь попасть из точки x' в x. Если найдется x такое, что $p(x) < 1$, то в силу предложения 2 последовательность $p(x2^n)$ стремится к нулю при $n \to \infty$. Но функция $p(x)$ монотонна и ограничена, следовательно ее предел равен нулю.

Предложение 4. При $q < p$ существует x, такое, что $p(x) < 1$.

Доказательство. Заметим, что вероятность $f^{(m)}(x,0)$ впервые попасть из x в 0 за m, $m \ge x$ шагов не превосходит

$$f^{(m)}(x,0) \le C_m^{(m-x)/2} p^{(m-x)/2} q^{(m+x)/2}. \tag{5.17}$$

При этом ясно, что число m должно иметь ту же четность, что и x, поэтому $(m-x)/2$ и $(m+x)/2$ целые числа. В самом деле, $f^{(m)}(x,0)$ не превосходит вероятности того, что частица, вышедшая из x, в момент времени m находится в нуле (возможно, не впервые). Для этого частице необходимо ровно $(m-x)/2$ раз прыгнуть вверх и $(m+x)/2$ раз прыгнуть вниз. Таких траекторий существует ровно

$C_m^{(m-x)/2}$. Далее, поскольку $pq \le 1/4$ и равенство $pq = 1/4$ достигается при $p = q$, то при $q < p$ существует $\alpha < 1$ такое, что

$$2\sqrt{pq} \le \alpha,$$

и поэтому правая часть (5.17) не превосходит

$$C_m^{(m-x)/2} p^{(m-x)/2} q^{(m+x)/2} \le 2^m (pq)^{m/2} \left(\frac{q}{p}\right)^{x/2} \le \left(\frac{q}{p}\right)^{x/2} \alpha^m. \quad (5.18)$$

Из (5.16) и (5.18) следует, что

$$p(x) = \sum_{m \ge x} f^{(m)}(x, 0) \le \left(\frac{q}{p}\right)^{x/2} \sum_{m \ge x} \alpha^m < 1$$

для достаточно больших x.

Таким образом, из предложений 2, 3 и 4 следует, что $p(x) \to 0$ при $x \to \infty$, и из (5.14) следует, что $p(x) = (q/p)^x$. Теорема 1 доказана.

Теорема 2. Вероятность достижения нуля из точки $x < 0$ равна

$$p(x) = \begin{cases} 1 & \text{при } p \ge q, \\ \left(\frac{p}{q}\right)^x & \text{при } p < q. \end{cases} \quad (5.19)$$

Доказательство аналогично доказательству теоремы 1.

Следствие. Из (5.6) и (5.19) получаем, что вероятность f_{00} того, что частица, выйдя из нуля, когда-либо туда вернется, равна

$$f_{00} = pp(1) + qp(-1) = \begin{cases} 2q, & \text{если } q \le p, \\ 2p, & \text{если } q > p. \end{cases} \quad (5.20)$$

Таким образом, частица, выйдя из нуля, вернется в 0 с вероятностью 1 в том и только в том случае, когда блуждание симметрично, т.е. $p = q = 1/2$. Интуитивно это ясно, так как при $p \ne q$ частица, выйдя из начала координат, будет смещаться к $+\infty$, если $p > q$, и к $-\infty$, если $p < q$.

Зададимся вопросом: какова вероятность того, что частица вернется в 0 бесконечное число раз?

Теорема 3. Обозначим A_0 событие, состоящее в том, что частица, выйдя из нуля, вернется туда бесконечное число раз. Тогда

$$\mathsf{P}(A_0) = \begin{cases} 1 & \text{при } p = q, \\ 0 & \text{при } p \ne q. \end{cases} \quad (5.21)$$

Доказательство. Обозначим $A_{0,N}$ событие, состоящее в том, что частица, выйдя из нуля, вернется туда по крайней мере N раз. Обозначим $f_{00}^{(k)}$ вероятность того, что в первый раз частица вернется в начало координат в момент времени k. Представив $A_{0,N}$ как объединение непересекающихся событий с фиксированным моментом первого возвращения, получим:

$$\mathsf{P}(A_{0,N}) = \sum_{k=1}^{\infty} f_{00}^{(k)} \mathsf{P}(A_{0,N-1}) = \mathsf{P}(A_{0,N-1}) f_{00}, \tag{5.22}$$

где

$$f_{00} = \sum_{k=1}^{\infty} f_{00}^{(k)}$$

есть вероятность когда-либо вернуться в 0. Последовательно применяя (5.22), получим:

$$\begin{aligned} \mathsf{P}(A_{0,N}) &= \mathsf{P}(A_{0,N-1}) f_{00} = \mathsf{P}(A_{0,N-2})(f_{00})^2 = \ldots \\ &= (f_{00})^{N-1} \mathsf{P}(A_{0,1}) = (f_{00})^N, \end{aligned} \tag{5.23}$$

так как $\mathsf{P}(A_{0,1}) = f_{00}$. Поскольку в силу свойства непрерывности

$$\mathsf{P}(A_0) = \lim_{N \to \infty} \mathsf{P}(A_{0,N}),$$

то из (5.23) следует, что $\mathsf{P}(A_0) = 1$ при $f_{00} = 1$ и $\mathsf{P}(A_0) = 0$ при $f_{00} < 1$. Отсюда с учетом (5.20) получаем утверждение теоремы 3.

Сколько же времени понадобится частице, начавшей блуждание в точке $x > 0$, чтобы достичь начала координат? Ответ дает следующая теорема. Обозначим $\tau_{x,0}$ случайное время первого попадания в начало координат частицы, начавшей блуждание в точке $x > 0$.

Теорема 4. При $p \geq q$ $\mathsf{E}\tau_{x,0} = \infty$. При $p < q$ $\mathsf{E}\tau_{x,0} = x/(q-p)$.

Для доказательства теоремы 4 нам понадобится понятие *условного математического ожидания.* Пусть ξ — дискретная неотрицательная случайная величина, то есть принимающая конечное или счетное число значений $x_1, x_2, \ldots;$ $x_i \geq 0$ для всех i. Пусть событие $D \subset \Omega$ имеет положительную вероятность.

Определение 2. *Условным математическим ожиданием случайной величины ξ при условии события D* называется число

$$\mathsf{E}(\xi | D) = \sum_{j=1}^{\infty} x_j \; P(\xi = x_j | D). \tag{5.24}$$

Заметим, что

$$\mathsf{P}(D)\mathsf{E}(\xi|D) = \sum_{j=1}^{\infty} x_j \mathsf{P}(\xi = x_j \cap D) \le \sum_{j=1}^{\infty} x_j \mathsf{P}(\xi = x_j) = \mathsf{E}\xi, \quad (5.25)$$

поэтому $\mathsf{E}(\xi|D) < \infty$, если $\mathsf{E}\xi < \infty$.

Пусть η — случайная величина, принимающая конечное число значений $y_1, \ldots, y_r$. Обозначим $D_i = \{\omega : \eta(\omega) = y_i\}$.

Определение 3. *Условным математическим ожиданием* $\mathsf{E}(\xi|\eta)$ *случайной величины* ξ *относительно случайной величины* η называется случайная величина, постоянная на каждом из множеств D_i и равная на множестве D_i величине $\mathsf{E}(\xi|D_i)$.

Нам понадобится следующее утверждение — аналог формулы полной вероятности.

Утверждение 2.

$$\mathsf{E}\xi = \mathsf{E}\mathsf{E}(\xi|\eta) = \sum_{i=1}^{r} \mathsf{P}(D_i)\mathsf{E}(\xi|D_i). \quad (5.26)$$

Доказательство. Пусть $\mathsf{E}\xi < \infty$. Тогда в силу (5.25) $\mathsf{E}(\xi|D_i) < \infty$ для каждого i. Воспользовавшись формулой полной вероятности, получим:

$$\mathsf{E}\xi = \sum_{j=1}^{\infty} x_j\ \mathsf{P}(\xi = x_j) = \sum_{j=1}^{\infty} x_j \sum_{i=1}^{r}\ \mathsf{P}(\xi = x_j|D_i)\mathsf{P}(D_i) =$$

$$= \sum_{i=1}^{r} \mathsf{P}(D_i) \sum_{j=1}^{\infty} x_j\ \mathsf{P}(\xi = x_j|D_i) = \sum_{i=1}^{r} \mathsf{P}(D_i)\mathsf{E}(\xi|D_i).$$

Если же $\mathsf{E}\xi = \infty$, то $\mathsf{E}(\xi|D_i) = \infty$ хотя бы для одного i. В самом деле, предположив, что $\mathsf{E}(\xi|D_i) < \infty$ для всех i и рассмотрев предыдущую цепочку равенств в обратном направлении, получим, что $\mathsf{E}\xi < \infty$, что является противоречием. Таким образом, утверждение 2 снова выполнено.

Перейдем к доказательству теоремы 4.

Доказательство теоремы 4. Воспользуемся утверждением 2, положив $\xi = \tau_{x,0}$, а η — бернуллиевская случайная величина, равная величине первого скачка частицы, находящейся в точке x. Если свой первый скачок из точки x частица совершает вправо, то условное математическое ожидание времени первого попадания в 0 равно

$\mathsf{E}\tau_{x+1,0}+1$. Аналогично, условное математическое ожидание времени первого попадания в 0 при условии, что первый скачок был сделан влево, равно $\mathsf{E}\tau_{x-1,0}+1$. (Мы полагаем $\mathsf{E}\tau_{0,0}=0$.) Таким образом, величины $T_x=\mathsf{E}\tau_{x,0}$ удовлетворяют уравнению

$$T_x = pT_{x+1} + qT_{x-1} + 1, \quad T_0 = 0. \tag{5.27}$$

Из (5.6) следует, что при $p>q$ $\tau_{x,0}=\infty$ с вероятностью $1-(q/p)^x>0$, и поэтому $\mathsf{E}\tau_{x,0}=\infty$.

Пусть теперь $q>p$. Уравнение (5.27) неоднородное. Соответствующее ему однородное уравнение мы изучали ранее; при $p\neq q$ оно имеет общее решение $A+B(q/p)^x$. Частное решение неоднородного уравнения при $p\neq q$ есть $x/(q-p)$. Таким образом, общее решение уравнения (5.27)

$$T_x = \frac{x}{(q-p)} + A + B\Big(\frac{q}{p}\Big)^x. \tag{5.28}$$

Граничное условие $T_0=0$ влечет $A+B=0$, т.е.

$$T_x = \frac{x}{(q-p)} + B\Big(\frac{q}{p}\Big)^x - B. \tag{5.29}$$

При $q>p$ если $B\geq 0$, то решение (5.29) положительно при всех $x>0$.

Итерируя уравнение (5.27), мы можем представить его в следующем виде:

$$T_x = \sum_{n=0}^{N-1}\sum_{\theta:n} \mathsf{P}(\theta) \; + \; \sum_{y>0}\sum_{\theta:y} \mathsf{P}(\theta) T_y. \tag{5.30}$$

Здесь $\sum_{\theta:n}$ берется по всем путям длины n, выходящим из точки x и не проходящим через начало координат, $\sum_{\theta:y}$ берется по всем путям длины N, выходящим из x и оканчивающимся в y и не проходящим через 0, а $\mathsf{P}(\theta)$ есть вероятность пути θ :

$$\mathsf{P}(\theta) = p^{r(\theta)} q^{l(\theta)},$$

где $r(\theta)$ — число скачков вправо, а $l(\theta)$ — число скачков влево у пути θ. Но $\sum_{\theta:n} \mathsf{P}(\theta)$ в правой части (5.30) есть не что иное, как вероятность того, что до момента n включительно блуждающая частица не посетила 0, т.е. $P(\tau_{x,0}>n)$. Таким образом, (5.30) записывается в виде:

$$T_x = \sum_{n=0}^{N-1} \mathsf{P}(\tau_{x,0}>n) \; + \; \sum_{y>0}\sum_{\theta:y} P(\theta) T_y. \tag{5.31}$$

Заметим, что при $B < 0 \quad T_x < 0$ для достаточно больших x, так что решение (5.29) не имеет вероятностного смысла. Поскольку при $B > 0$ и $q > p \quad T_x > 0$ при всех $x > 0$, то второе слагаемое в (5.31) положительно и, значит, для любого N и любого $x > 0$

$$\sum_{n=0}^{N-1} \mathsf{P}(\tau_{x,0} > n) \le T_x. \tag{5.32}$$

Отсюда следует, что для любого $x > 0$ ряд

$$\sum_{n=0}^{\infty} \mathsf{P}(\tau_{x,0} > n) < \infty.$$

Но

$$\sum_{n=0}^{\infty} \mathsf{P}(\tau_{x,0} > n) = \sum_{n=0}^{\infty} n\mathsf{P}(\tau_{x,0} = n) = \mathsf{E}\tau_{x,0}.$$

Итак, мы показали, что при $q > p \quad \mathsf{E}\tau_{x,0} < \infty$, и, следовательно, оно совпадает с решением (5.29) уравнения (5.27).

Утверждение 3. Если среднее время попадания из x в 0 конечно, то оно растет с ростом x линейным образом, т.е. если $\mathsf{E}\tau_{x,0} < \infty$, то существует константа $c > 0$ такая, что $\mathsf{E}\tau_{x,0} = cx$.

Доказательство. Положим $c = \mathsf{E}\tau_{x,x-1}$ — среднее время первого попадания из x в $x-1$. Оно конечно, поскольку $\mathsf{E}\tau_{x,0} < \infty$. Но так как время $\tau_{x,0}$ первого попадания из x в 0 складывается из времени $\tau_{x,x-1}$ первого попадания из x в $x-1$, времени $\tau_{x-1,x-2}$ первого попадания из $x-1$ в $x-2$ и т.д., и все эти величины равны между собой, то

$$\mathsf{E}\tau_{x,0} = \mathsf{E}\tau_{x,x-1} + \ldots + \mathsf{E}\tau_{1,0} = cx, \tag{5.33}$$

что и требовалось доказать.

Из утверждения 3 вытекает, что решение (5.29) имеет вероятностный смысл лишь при $B = 0$, так как в противном случае $\mathsf{E}\tau_{x,0}$ росло бы экспоненциально. Итак, при $q > p \quad \mathsf{E}\tau_{x,0} = x/(q-p)$.

Нам осталось разобрать случай $p = q$. Уравнение (5.27) в этом случае имеет вид

$$T_x = \frac{1}{2}T_{x+1} + \frac{1}{2}T_{x-1} + 1, \quad T_0 = 0. \tag{5.34}$$

Общее решение соответствующего однородного уравнения имеет вид $c_1 + c_2 x$. Частное решение уравнения (5.34) $T_x = -x^2$. Таким образом, общее решение (5.34) равно

$$T_x = -x^2 + c_1 + c_2 x. \tag{5.35}$$

Решение (5.35) не является положительным на всей полуоси $x > 0$ ни при каких значениях c_1, c_2, поэтому (5.35) не имеет вероятностного смысла. Поэтому единственным неотрицательным решением является $T_x \equiv +\infty$. Теорема 4 доказана.

Литература

Феллер В. Введение в теорию вероятностей и ее приложения, том 1, глава 14, §§1 – 3. Спицер Ф. Принципы случайного блуждания. М., Мир, 1969, глава 1, §1.

Лекция 6
Цепи Маркова

Пусть заданы произвольное множество $S = \{1, 2, \ldots, n\}$ из n элементов и матрица $P = (p_{ij})$ размера $n \times n$ такая, что

$$0 \leq p_{ij} \leq 1, \ \sum_{j=1}^{n} p_{ij} = 1 \quad \text{для всех } i.$$

Такая матрица называется *стохастической*. С ней удобно связывать направленный граф, вершины которого соответствуют элементам множества S. Проведем направленное ребро из i в j тогда и только тогда, когда $p_{ij} \neq 0$.

Рассмотрим пространство $\Omega^{(T)}$ всевозможных последовательностей $\omega = (\omega_0, \omega_1, \ldots, \omega_T)$ длины $T+1$, элементы которых принадлежат S. Такие последовательности будем называть *траекториями* и представлять себе как путь частицы, блуждающей по графу.

Зададим вероятностную меру $\mathsf{P}^{(T)}$ на траекториях следующим образом:

$$\mathsf{P}^{(T)}(\omega) = p^{(0)}_{\omega_0} p_{\omega_0\omega_1} p_{\omega_1\omega_2} \ldots p_{\omega_{T-1}\omega_T}, \tag{6.1}$$

где $p^{(0)} = (p^{(0)}_i, \ i = 1, \ldots, n)$ — вероятностная мера на пространстве S. Мы интерпретируем $p^{(0)}_i$ как вероятность начать путь из состояния i. Если осуществляется событие $\omega_t = j$, говорят, что в момент времени t цепь находится в состоянии j. Говорят также, что p_{ij} есть вероятность перехода из вершины i в вершину j.

Предложение 1. Мера $\mathsf{P}^{(T)}$ задана с помощью (6.1) корректно, то есть

$$\sum_{\omega \in \Omega^{(T)}} \mathsf{P}^{(T)}(\omega) = 1. \tag{6.2}$$

Доказательство. Зафиксируем $\omega_0, \omega_1, \ldots, \omega_{T-1}$ и просуммируем в (6.2) по ω_T. Поскольку матрица P стохастическая, получим

$$\sum_{\omega_0, \ldots, \omega_T} p^{(0)}_{\omega_0} p_{\omega_0\omega_1} \ldots p_{\omega_{T-1}\omega_T} = \sum_{\omega_0, \ldots, \omega_{T-1}} p^{(0)}_{\omega_0} p_{\omega_0\omega_1} \ldots p_{\omega_{T-2}\omega_{T-1}}.$$

Суммируя последовательно по $\omega_{T-1}, \ldots, \omega_0$ и пользуясь на каждом шаге стохастичностью матрицы P, получим (6.2).

Определение 1. *Конечной цепью Маркова* с дискретным временем $t = 0, 1, \ldots, T$ с пространством состояний S, матрицей переходных вероятностей P и начальным распределением (или вектором начальных вероятностей) $p^{(0)}$ называется мера $\mathsf{P}^{(T)}$ на пространстве $\Omega^{(T)}$, заданная с помощью (6.1).

В случае, если все координаты вектора начальных вероятностей $p^{(0)}$, за исключением i-й, равны нулю, говорят, что марковская цепь имеет начальное состояние i.

Аналогично предложению 1 доказывается, что для любого $t' < t$ и любого $\omega \in \Omega^{(t)}$

$$\sum_{\omega_{t'+1}, \ldots, \omega_t} \mathsf{P}^{(t)}(\omega) = p^{(0)}_{\omega_0} p_{\omega_0 \omega_1} \ldots p_{\omega_{t'-1}\omega_{t'}} = \mathsf{P}^{(t')}(\omega^{(t')}),$$

где суммирование ведется по всем возможным значениеям $\omega_{t'+1} \in S, \ldots, \omega_t \in S$, а $\omega^{(t')}$ обозначает ограничение вектора $\omega \in \Omega^{(t)}$ на пространство $\Omega^{(t')}$. Таким образом, вероятность, индуцированная на пространстве $\Omega^{(t')}$ мерой $\mathsf{P}^{(t)}$, совпадает с $\mathsf{P}^{(t')}$. Подобно тому, как мы построили меру на бесконечных последовательностях для схемы Бернулли, мы можем построить меру на бесконечных последовательностях ω с элементами из S. Поэтому в дальнейшем мы иногда опускаем индекс у меры $\mathsf{P}^{(t)}$.

Пример. Рассмотрим отрезок $[-N, N]$ целочисленной оси и рассмотрим случайное блуждание, которое мы изучали в лекции 5, сузив его на этот отрезок. А именно, для любой внутренней точки y отрезка вероятности переходов задаются формулой (5.5). Для $y = -N$ положим

$$P(\hat{S}_t = x \mid \hat{S}_{t-1} = -N) = \begin{cases} p, & \text{если } x = -N + 1; \\ q, & \text{если } x = -N; \\ 0 & \text{в остальных случаях.} \end{cases}$$

Аналогично для $y = N$ положим

$$P(\hat{S}_t = x \mid \hat{S}_{t-1} = N) = \begin{cases} p, & \text{если } x = N; \\ q, & \text{если } x = N - 1; \\ 0 & \text{в остальных случаях.} \end{cases}$$

Последовательность случайных величин $\hat{S}_t$ образует конечную цепь Маркова, переходы которой графически изображены на рисунке 6.1.

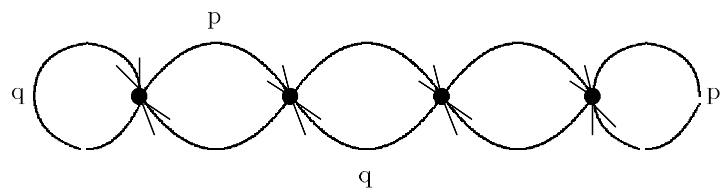


Рис. 6.1. Случайное блуждание на отрезке.

Пусть марковская цепь имеет начальное состояние i, т.е. $p_i^{(0)} = 1$. Обозначим $p_{ij}^{(t)}$ вероятность того, что в момент времени t цепь находится в состоянии j. В частности, $p_{ij}^{(1)} = p_{ij}$ и

$$p_{ij}^{(2)} = \sum_{k \in S} p_{ik} p_{kj}. \tag{6.3}$$

На языке матриц равенство (6.3) означает, что $p_{ij}^{(2)}$ является элементом матрицы P^2. По индукции мы получаем рекуррентную формулу

$$p_{ij}^{(t)} = \sum_{k \in S} p_{ik} p_{kj}^{(t-1)}, \tag{6.4}$$

которая означает, что интересующая нас вероятность $p_{ij}^{(t)}$ задается (ij)-м элементом t-й степени матрицы переходов P.

Пусть теперь вектор $p^{(0)}$ начальных вероятностей произвольный. Обозначим $p^{(t)}$ набор вероятностей $p_j^{(t)} = \mathsf{P}^{(t)}(\omega_t = j)$, $j \in S$. Он называется распределением марковской цепи в момент времени t. Тогда

$$p_j^{(t)} = \sum_{i \in S} p_i^{(0)} p_{ij}^{(t)},$$

то есть чтобы найти распределение вероятностей в момент t, нужно к вектору-строке $p^{(0)}$ начального распределения применить матрицу P^t :

$$p^{(t)} = p^{(0)} P^t.$$

Предложение 2. Пусть последовательность состояний i_0, i_1, $\ldots$, i_{t-1}, i такова, что

$$\mathsf{P}\{\omega_0 = i_0, \ldots, \omega_{t-1} = i_{t-1}, \omega_t = i\} > 0.$$

Тогда
$$\mathsf{P}\{\omega_{t+1}=j \mid \omega_0=i_0,\dots,\omega_t=i\}=p_{ij}.$$

Доказательство.
$$\mathsf{P}\{\omega_{t+1}=j \mid \omega_0=i_0,\dots,\omega_t=i\}=$$
$$=\frac{\mathsf{P}\{\omega_{t+1}=j,\omega_0=i_0,\dots,\omega_t=i\}}{\mathsf{P}\{\omega_0=i_0,\dots,\omega_{t-1}=i_{t-1},\omega_t=i\}}=$$
$$=\frac{p_{i_0}^{(0)}p_{i_0i_1}\dots p_{i_{t-1}i}p_{ij}}{p_{i_0}^{(0)}p_{i_0i_1}\dots p_{i_{t-1}i}}=p_{ij}.$$

Таким образом, условная вероятность оказаться в момент времени t в некотором состоянии при условии, что заданы состояния цепи во все предыдущие моменты времени, зависит лишь от состояния в момент $t-1$ и не зависит от всех предыдущих состояний. Это свойство иногда берется в качестве определения цепи Маркова. Его можно понимать так, что при фиксированном "настоящем" "будущее" не зависит от "прошлого" (см. также упражнение 1 ниже).

Предложение 3. Для любого $t>0$
$$\mathsf{P}\{\omega_{t+k}=j \mid \omega_t=i\}=p_{ij}^{(k)},$$
где $p_{ij}^{(k)}$ — элемент матрицы P^k.

Доказательство по индукции. При $k=1$
$$\mathsf{P}\{\omega_{t+1}=j \mid \omega_t=i\}=\frac{\mathsf{P}\{\omega_{t+1}=j,\omega_t=i\}}{\mathsf{P}\{\omega_t=i\}}=$$
$$=\frac{\sum_{i_0\dots i_{t-1}}p_{i_0}^{(0)}p_{i_0i_1}\dots p_{i_{t-1}i}p_{ij}}{\sum_{i_0\dots i_{t-1}}p_{i_0}^{(0)}p_{i_0i_1}\dots p_{i_{t-1}i}}=p_{ij}.$$
Далее, предположим, что для всех $k\le m$ утверждение доказано. Имеем
$$\mathsf{P}\{\omega_{t+m+1}=j \mid \omega_t=i\}=\frac{\mathsf{P}\{\omega_{t+m+1}=j,\omega_t=i\}}{\mathsf{P}\{\omega_t=i\}}=$$
$$=\frac{\sum_s\mathsf{P}\{\omega_{t+m+1}=j,\omega_{t+m}=s,\omega_t=i\}}{\mathsf{P}\{\omega_t=i\}}=$$
$$=\frac{\sum_s\mathsf{P}\{\omega_{t+m+1}=j \mid \omega_{t+m}=s,\omega_t=i\}\mathsf{P}(\omega_{t+m}=s,\omega_t=i)}{\mathsf{P}\{\omega_t=i\}}=$$

$$= \sum_s \mathsf{P}\{\omega_{t+m+1} = j \mid \omega_{t+m} = s, \omega_t = i\}\mathsf{P}\{\omega_{t+m} = s \mid \omega_t = i\} =$$

$$= \sum_s p_{is}^{(m)} p_{sj} = p_{ij}^{(m+1)}.$$

Здесь мы воспользовались тем, что по предположению индукции

$$\mathsf{P}\{\omega_{t+m} = s \mid \omega_t = i\} = p_{is}^{(m)},$$

а из предложения 2 следует, что

$$\mathsf{P}\{\omega_{t+m+1} = j \mid \omega_{t+m} = s, \omega_t = i\} = p_{sj}.$$

Упражнение 1. Доказать, что для любых $s, r > 0$ события $\{\omega_{t+s} = i\}$ и $\{\omega_{t-r} = j\}$ условно независимы при условии $\{\omega_t = k\}$, т.е.

$$\mathsf{P}\{\omega_{t+s} = i, \omega_{t-r} = j \,|\, \omega_t = k\} = \qquad (6.5)$$
$$= \mathsf{P}\{\omega_{t+s} = i \,|\, \omega_t = k\}\mathsf{P}\{\omega_{t-r} = j \,|\, \omega_t = k\}.$$

Условие положительности. Пусть для марковской цепи выполнено следующее условие положительности: существуют t_0 и $\gamma > 0$ такие, что

$$p_{ij}^{(t_0)} > \gamma \qquad (6.6)$$

для всех i, j. В этом случае будем говорить, что цепь Маркова *неприводима и апериодична.*

Примеры цепей Маркова, для которых условие положительности *не выполнено.*

1. Пространство состояний S разбито на два подмножества A и B такие, что $p_{ij} = 0$ для всех $i \in A, j \in B$ и всех $j \in A, i \in B$ (см. рис. 6. 2). Такая цепь называется приводимой.

2. Периодическая цепь: пример дает граф цепи из двух состояний на рисунке 6.3, где каждый переход имеет вероятность 1. Матрица переходов этой цепи равна

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

поэтому $P^k = P$, если k четно, и $P^k = E$, если k нечетно, где E — единичная матрица.

Определение 2. Состояние i называется *несущественным*, если с положительной вероятностью можно оказаться вне этого состояния за конечный промежуток времени, но нельзя в него вернуться,

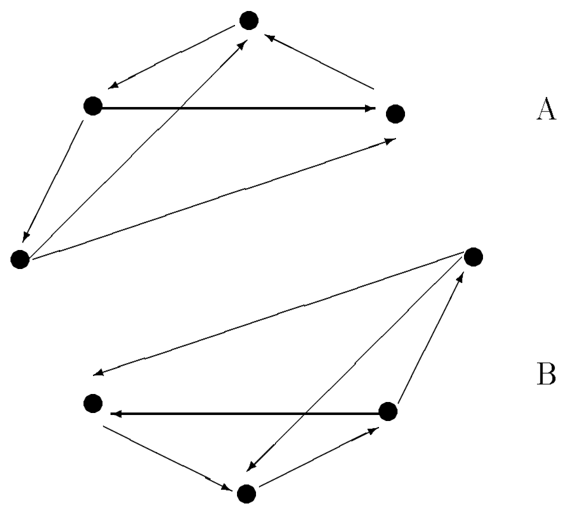


Рис. 6.2. Приводимая цепь.

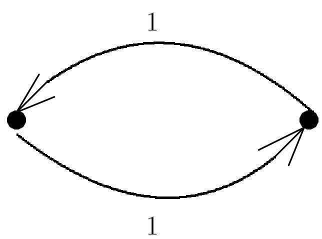


Рис. 6.3. Периодическая цепь.

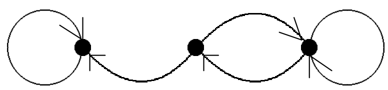


Рис. 6.4. Цепь с несущественными состояниями.

т.е. если существуют такие m и j, что $p_{ij}^{(m)} > 0$, но для всех t $p_{ji}^{(t)} = 0$. В противном случае состояние называется *существенным*.

На рисунке 6.4 даны примеры несущественных состояний. Здесь все состояния, кроме а, несущественны. Множество существенных состояний обладает, очевидно, тем свойством, что, попав в него, марковская цепь никогда уже его не покинет. Можно показать, что в случае, когда все состояния марковской цепи существенны, примеры невыполнимости условия положительности исчерпываются периодическими и приводимыми цепями.

Определение 3. Вероятностная мера $\pi = (\pi_1, \ldots, \pi_n)$ на множестве S называется *инвариантной мерой* марковской цепи, если $\pi = \pi P$.

Теорема 1 (Эргодическая теорема). Пусть для марковской цепи выполнено условие положительности (6.6). Тогда существует единственная вероятностная инвариантная мера π для этой марковской цепи, и для всех i, j

$$\lim_{t \to \infty} p_{ij}^{(t)} = \pi_j. \tag{6.7}$$

Более того, сходимость к инвариантному распределению экспоненциальна, т.е. существуют положительные константы c и α такие, что для любых i и j для достаточно больших t

$$| p_{ij}^{(t)} - \pi_j | < c \exp(-t\alpha). \tag{6.8}$$

Замечание. Из (6.7) следует, что для любого начального распределения $p^{(0)}$

$$\lim_{t \to \infty} p^{(0)} P^t = \pi.$$

Доказательство. Сначала приведем идею доказательства. Рассмотрим две вероятностные меры на пространстве $S : \mu'$ и μ''. Поскольку $\sum_i (\mu_i' - \mu_i'') = 0$, вектор $\mu' - \mu''$ можно представить как разность двух неотрицательных мер на S с непересекающимися носителями и с одинаковой массой, равной $\sum^+(\mu'_i - \mu''_i)$, где $\sum^+$ означает, что в сумму включаются лишь неотрицательные слагаемые. Под действием оператора, заданного стохастической матрицей P^{t_0}, положительная и отрицательная масса перераспределятся некоторым образом по состояниям. При этом поскольку $p_{ij}^{t_0} > \gamma$, то обязательно произойдут сокращения массы. Результирующую меру можно вновь представить как разность двух неотрицательных мер с непересекающимися носителями и с одинаковой массой, однако масса новой меры окажется строго меньше массы исходной меры. Взяв $\mu' = p^{(0)}P^{t'}$, $\mu'' = p^{(0)}P^{t''}$ и продолжая аналогично, получим, что масса разности $p^{(0)}P^{t'} - p^{(0)}P^{t''}$ мала для достаточно больших t', t'', а это и означает сходимость.

Приведем теперь подробное доказательство. Определим расстояние $d(\mu', \mu'')$ между вероятностными мерами следующим образом:

$$d(\mu', \mu'') = \frac{1}{2}\sum_{i=1}^{n} \mid \mu'_i - \mu''_i \mid . \qquad (6.9)$$

Пространство вероятностных мер на S с метрикой d является полным метрическим пространством. Заметим, что

$$d(\mu', \mu'') \leq 1.$$

Кроме того,

$$d(\mu', \mu'') = \frac{1}{2}\sum_{i=1}^{n} \mid \mu'_i - \mu''_i \mid = \frac{1}{2}\sum{}^{+}(\mu'_i - \mu''_i) +$$

$$+\frac{1}{2}\sum{}^{+}(\mu''_i - \mu'_i) = \sum{}^{+}(\mu'_i - \mu''_i).$$

Лемма 1.

а) Если Q — стохастическая матрица, а p — вероятностная мера, то pQ тоже вероятностная мера.

б) Если стохастическая матрица Q такова, что $q_{ij} \geq \gamma > 0$ для некоторого γ и для всех i, j, то для любых вероятностных мер μ' и μ''

$$d(\mu' Q, \mu'' Q) \leq (1-\gamma) d(\mu', \mu''). \qquad (6.10)$$

Доказательство.

а) Пусть $p' = pQ$. Тогда, очевидно, $p'_j \geq 0$ для всех j. Далее,

$$\sum_j p'_j = \sum_j \sum_i p_i q_{ij} = \sum_i p_i \sum_j q_{ij} = 1.$$

б)

$$d(\mu' Q, \mu'' Q) = \sum\nolimits_j^+ \sum_i (\mu'_i - \mu''_i) q_{ij} \le \sum\nolimits_j^+ \sum\nolimits_i^+ (\mu'_i - \mu''_i) q_{ij} =$$

$$= \sum\nolimits_i^+ (\mu'_i - \mu''_i) \sum\nolimits_j^+ q_{ij} \le d(\mu', \mu''),$$

так как $\sum_j^+ q_{ij} \le \sum_j q_{ij} = 1$. Заметим, что сумма $\sum_j^+ q_{ij}$ не может содержать суммирования по всем индексам j. Действительно, если бы для каждого j

$$\sum_{i=1}^n \mu'_i q_{ij} > \sum_{i=1}^n \mu''_i q_{ij},$$

то

$$\sum_{j=1}^n \sum_{i=1}^n \mu'_i q_{ij} > \sum_{j=1}^n \sum_{i=1}^n \mu''_i q_{ij},$$

и, значит, сумма координат вектора $(\mu' - \mu'')Q$ положительна, что невозможно, так как в силу пункта а) векторы $\mu' Q$ и $\mu'' Q$ вероятностные. Итак, в сумме $\sum_j^+ q_{ij}$ хотя бы один индекс j отсутствует, поэтому если $q_{ij} > \gamma$ для всех i, j, то

$$\sum\nolimits_j^+ q_{ij} < 1 - \gamma.$$

Отсюда следует, что

$$d(\mu' Q, \mu'' Q) \le (1 - \gamma) \sum\nolimits_i^+ (\mu'_i - \mu''_i) = (1 - \gamma)\, d(\mu', \mu'').$$

Лемма 1 доказана.

Продолжим доказательство эргодической теоремы. Пусть $p^{(0)}$ есть произвольное распределение вероятностей на S. Покажем, что последовательность

$$p^{(t)} = p^{(0)} P^t$$

является фундаментальной, т.е. для любого $\varepsilon > 0$ найдется N такое, что для любого $t > N$ и любого $m > 0$

$$d\,(p^{(t)}, p^{(t+m)}) < \varepsilon. \tag{6.11}$$

Положим

$$\mu' = p^{(t-t_0)}, \quad \mu'' = p^{(t+m-t_0)}, \quad Q = P^{t_0}$$

и воспользуемся леммой 1. Получим:

$$d(p^{(t)}, p^{(t+m)}) = d(p^{(0)} P^t, p^{(0)} P^{t+m}) \le$$

$$\leq (1-\gamma)\, d(p^{(0)}P^{t-t_0}, p^{(0)}P^{t+m-t_0}) \leq \ldots$$
$$\ldots \leq (1-\gamma)^k d(p^{(0)}P^{t-kt_0}, p^{(0)}P^{t+m-kt_0}) \leq (1-\gamma)^k, \qquad (6.12)$$

где k есть наибольшее целое, такое, что $kt_0 \leq t$. Для достаточно большого t отсюда следует (6.11). Обозначим

$$\pi = \lim_{t\to\infty} p^{(t)}.$$

Тогда

$$\pi P = (\lim_{t\to\infty} p^{(t)})P = \lim_{t\to\infty} (p^{(0)}P^t)P = \lim_{t\to\infty} p^{(0)}P^{t+1} = \pi,$$

то есть π — инвариантная мера:

$$\pi P = \pi.$$

Докажем теперь единственность инвариантной меры. Допустим, имеются две инвариантные меры, $\pi^{(1)}$ и $\pi^{(2)}$. Тогда

$$\pi^{(1)} = \pi^{(1)}P^{t_0}, \quad \pi^{(2)} = \pi^{(2)}P^{t_0}.$$

Поэтому

$$d(\pi^{(1)}, \pi^{(2)}) = d(\pi^{(1)}P^{t_0}, \pi^{(2)}P^{t_0}) \leq (1-\gamma)\, d(\pi^{(1)}, \pi^{(2)}),$$

откуда следует, что $d(\pi^{(1)}, \pi^{(2)}) = 0$.

Итак,

$$\lim_{t\to\infty} p^{(0)}P^t = \pi$$

не зависит от начального вектора $p^{(0)}$.

Возьмем в качестве начальной меры меру $\delta^{(i)}$, сосредоточенную в i-м состоянии, т.е. $\delta_j^{(i)} = 0$ для $i \neq j$ и $\delta_i^{(i)} = 1$. Тогда

$$p_{ij}^{(t)} = \delta^{(i)}P^t \to \pi_j, \quad j = 1, \ldots, n,$$

что и доказывает (6.7). Экспоненциальная сходимость (6.8) следует из (6.12).

Рассмотрим марковскую цепь, удовлетворяющую условию положительности. Пусть начальное состояние этой цепи есть i. Обозначим через τ_{ii} случайное время первого возвращения марковской цепи в состояние i.

Предложение 4. Существуют константы $C > 0, \alpha > 0$ такие, что для любого достаточно большого t выполнены экспоненциальные оценки

$$\mathsf{P}(\tau_{ii} > t) \leq C \exp(-\alpha t).$$

Доказательство.

$$\mathsf{P}(\tau_{ii} > t) \le \mathsf{P}(\omega_{t_0} \ne i, \omega_{2t_0} \ne i, \ldots, \omega_{kt_0} \ne i), \tag{6.13}$$

где $t = kt_0 + t'$, $t' < t_0$, а t_0 фигурирует в условии положительности. Из условия положительности следует, что правая часть (6.13) не превосходит $(1-\gamma)^k$, что и доказывает предложение 4.

Упражнение 2. Доказать, что для любой случайной величины ξ, принимающей целые неотрицательные значения,

$$\mathsf{E}\xi = \sum_{m=0}^{\infty} \mathsf{P}(\xi \ge m).$$

Предложение 5. Дисперсия случайной величины τ_{ii} конечна.

Доказательство. Из предложения 4 и упражнения 2 следует, что $\mathsf{E}\tau_{ii} < \infty$ и $\mathsf{E}\tau_{ii}^2 < \infty$:

$$\mathsf{E}\tau_{ii}^2 = \sum_{m=0}^{\infty} P(\tau_{ii}^2 \ge m) \le C \sum_{m=0}^{\infty} \exp(-\alpha\sqrt{m}) < \infty. \tag{6.14}$$

Рассмотрим последовательность случайных величин $\tau_{ii}^{(1)}, \tau_{ii}^{(2)} \ldots$ где $\tau_{ii}^{(k)}$ есть время между $(k-1)$-м и k-м посещением цепью состояния i. В силу марковского свойства эти случайные величины независимы и одинаково распределены. В частности, $\mathsf{E}\tau_{ii}^{(l)} = \mathsf{E}\tau_{ii}$ не зависит от l. В лекции 1 мы доказали закон больших чисел для схемы Бернулли, однако этот закон верен в гораздо более общей ситуации. В частности, мы увидим, что он выполнен и для последовательности случайных величин $\tau_{ii}^{(k)}$.

Теорема (закон больших чисел для независимых случайных величин). Пусть $\xi_1, \xi_2, \ldots$ — последовательность независимых одинаково распределенных случайных величин с $\mathsf{E}|\xi_j| < \infty$ и конечной дисперсией. Пусть $S_n = \sum_{i=1}^{n} \xi_i$. Тогда $S_n/n \to \mathsf{E}\xi_j$ по вероятности при $n \to \infty$, т.е. для всякого $\varepsilon > 0$

$$\mathsf{P}\left(\left|\frac{S_n}{n} - \mathsf{E}\xi_j\right| \ge \varepsilon\right) \to 0, \quad n \to \infty.$$

Доказательство аналогично приведенному в лекции 1 доказательству закона больших чисел для схемы Бернулли.

В силу предложения 5 последовательность случайных величин $\tau_{ii}^{(k)}$ удовлетворяет условиям закона больших чисел.

Теорема 2. Пусть π — инвариантная мера марковской цепи, для которой выполнено условие положительности. Тогда для любого $i \in S$

$$\pi_i = \frac{1}{\mathsf{E}\tau_{ii}}. \tag{6.15}$$

Доказательство. Пусть начальное состояние марковской цепи есть i. Обозначим через $r(N)$ случайное число попаданий в состояние i за промежуток времени длины N.

Предложение 6. Для достаточно больших N

$$\mathsf{E}r(N) \sim N\pi_i. \tag{6.16}$$

Доказательство предложения 6. Рассмотрим случайную величину χ_k, являющуюся индикатором события, состоящего в том, что в момент времени k марковская цепь находится в состоянии i. Тогда

$$\mathsf{E}\chi_k = p_{ii}^{(k)}.$$

Очевидно,

$$r(N) = \sum_{k=1}^{N} \chi_k,$$

поэтому, пользуясь теоремой 1, получаем:

$$\frac{\mathsf{E}r(N)}{N} = \frac{1}{N}\sum_{k=1}^{N} p_{ii}^{(k)} = \pi_i + o(1), \quad N \to \infty. \tag{6.17}$$

Предложение 6 доказано.

Выберем достаточно большое (мы уточним это ниже) целое положительное число M и рассмотрим поведение цепи Маркова за промежуток времени $N = [M\mathsf{E}\tau_{ii}]$. Согласно предложению 6,

$$\mathsf{E}r([M\mathsf{E}\tau_{ii}]) \sim [M\mathsf{E}\tau_{ii}]\pi_i, \quad M \to \infty. \tag{6.18}$$

С другой стороны, интуитивно ясно, что в среднем число попаданий цепи в состояние i за промежуток времени $M\mathsf{E}\tau_{ii}$ близко к M. Если мы докажем, что, наряду с (6.18), выполнено асимптотическое равенство

$$\mathsf{E}r([M\mathsf{E}\tau_{ii}]) \sim M, \quad M \to \infty, \tag{6.19}$$

то получим, что

$$[M\mathsf{E}\tau_{ii}]\pi_i \sim M, \tag{6.20}$$

откуда, очевидно, вытекает утверждение теоремы 2. Итак, нам осталось доказать следующую лемму.

Лемма 2.

$$\frac{\mathsf{E}r([M\mathsf{E}\tau_{ii}])}{M} = 1 + o(1), \quad M \to \infty. \tag{6.21}$$

Доказательство леммы 2. Рассмотрим случайный момент времени $t(M)$, в который произошло M-е возвращение цепи в состояние i :

$$t(M) = \sum_{k=1}^{M} \tau_{ii}^{(k)}.$$

Согласно закону больших чисел для любых ε, δ существует $M_0 = M_0(\varepsilon, \delta)$ такое, что для любого $M > M_0$

$$\left| \frac{t(M) - M\mathsf{E}\tau_{ii}}{M} \right| < \varepsilon \tag{6.22}$$

с вероятностью большей, чем $1-\delta$. Иными словами, с вероятностью, большей, чем $1-\delta$, случайный момент времени M-го возвращения в состояние i отличается от интересующего нас промежутка времени $M\mathsf{E}\tau_{ii}$ не более, чем на εM :

$$|t(M) - M\mathsf{E}\tau_{ii}| \leq M\varepsilon. \tag{6.23}$$

Отсюда можно сделать следующий вывод. Поскольку промежутки времени отличаются не больше, чем на εM, то и числа возвращений в i за данные промежутки отличаются не более, чем на εM. Точнее, верно следующее предложение.

Предложение 7. Для любых ε, δ существует M_0 такое, что для любого $M > M_0$

$$\mathsf{P}\{|r(M\mathsf{E}\tau_{ii}) - M| > M\varepsilon\} < \delta.$$

Доказательство. Рассмотрим любую траекторию марковской цепи, для которой выполнено (6.23). Поскольку промежутки времени $t(M)$ и $[M\mathsf{E}\tau_{ii}]$ отличаются друг от друга не больше чем на $M\varepsilon$, то и число возвращений в состояние i за время $[M\mathsf{E}\tau_{ii}]$ для данной траектории отличается от M не больше, чем на $M\varepsilon$. (Напомним, что по определению $r(t(M)) = M$.) Таким образом, для множества траекторий, вероятность которого превосходит $1-\delta$, получаем:

$$|r(M\mathsf{E}\tau_{ii}) - M| < M\varepsilon.$$

Предложение 7 доказано. Запишем его в таком виде:

$$\mathsf{P}\{M(1-\varepsilon) \le r(M\mathsf{E}\tau_{ii}) \le M(1+\varepsilon)\} > 1-\delta. \qquad (6.24)$$

Неравенство (6.24) означает, что случайная величина $r(M\mathsf{E}\tau_{ii})/M$ сходится по вероятности к 1 при $M \to \infty$. Но тогда математическое ожидание этой случайной величины также стремится к 1.

Упражнение 3. Доказать, что если последовательность равномерно ограниченных случайных величин сходится по вероятности к константе, то последовательность их математических ожиданий стремится к этой же константе.

Лемма 2, а с ней и теорема 2 доказаны.

Литература

Ширяев А. Н. Вероятность, глава 8. Феллер В. Введение в теорию вероятностей и ее приложения, том 1, глава 15.

Лекция 7

Распределения случайных величин. Функции распределения. Плотность. Свертка распределений

Пусть $(\Omega, \Sigma, \mathsf{P})$ — вероятностное пространство. Напомним, что случайной величиной называется измеримая функция $\xi(\omega)$ на этом пространстве. В практических применениях, однако, не всегда удобно пользоваться только таким языком, так как нас будут интересовать, как правило, только вероятности $\mathsf{P}(\xi \in A)$ того, что значение случайной величины принадлежит некоторому (открытому, замкнутому или, более общо, борелевскому) множеству на прямой. Напомним, что борелевской σ-алгеброй $\mathcal{B}(\mathbf{R})$ называется σ-алгебра на вещественной прямой, порожденная интервалами, т.е. наименьшая σ-алгебра, содержащая всевозможные интервалы.

Определение 1. Вероятностная мера P_ξ на измеримом пространстве $(\mathbf{R}, \mathcal{B}(\mathbf{R}))$, определяемая для любого $B \in \mathcal{B}(\mathbf{R})$ как

$$P_\xi(B) \stackrel{\mathrm{def}}{=} \mathsf{P}\{\xi^{-1}(B)\} \equiv \mathsf{P}\{\omega : \xi(\omega) \in B\},$$

называется *распределением случайной величины* ξ.

Теорема 1. Пусть $\xi = \xi(\omega)$ — случайная величина, P_ξ — ее распределение, $\varphi = \varphi(x)$ — некоторая борелевская функция. Тогда

$$\int\limits_{\mathrm{R}} \varphi(x)\, dP_\xi(x) = \int\limits_{\Omega} \varphi(\xi(\omega))\, d\mathsf{P}(\omega). \tag{7.1}$$

(Если один из интегралов существует в смысле Лебега, то и другой существует и они равны.)

Следствие. Полезные формулы для математических ожиданий.

$$\mathsf{E}\varphi(\xi) = \int\limits_{\mathrm{R}} \varphi(x)\, dP_\xi(x).$$

В частности, полагая $\varphi(x) \equiv x$, имеем

$$\mathsf{E}\xi = \int\limits_{\mathrm{R}} x\, dP_\xi(x).$$

Аналогично, $\mathsf{E}\xi^2 = \int\limits_{\mathrm{R}} x^2\, dP_\xi(x)$, $\quad \mathsf{E}\xi^k = \int\limits_{\mathrm{R}} x^k\, dP_\xi(x)$ и т.п.

Доказательство теоремы 1. Докажем вначале формулу (7.1) для индикатора борелевского множества B:

$$\varphi(x) = I_B(x) = \begin{cases} 1, & x \in B, \\ 0, & x \notin B. \end{cases}$$

Из определения интеграла Лебега легко следует, что в этом случае формула (7.1) принимает вид

$$P_\xi(B) = \mathsf{P}\{\omega : \xi(\omega) \in B\},$$

а это есть не что иное, как распределение P_ξ.

В силу линейности интеграла Лебега формула (7.1) верна для неотрицательных простых функций

$$\varphi(x) = \sum_i c_i I_{B_i}(x), \quad c_i > 0,\ B_i \cap B_j \neq \emptyset,\ B_i \in \mathcal{B}(\mathbf{R}).$$

Применяя теорему о монотонной сходимости для интеграла Лебега, убеждаемся, что формула (7.1) верна для произвольных неотрицательных борелевских функций φ.

Чтобы доказать (7.1) для произвольной измеримой функции φ, достаточно представить ее в виде $\varphi(x) = \varphi^+(x) - \varphi^-(x)$, где $\varphi^+(x) \geq 0$, $\varphi^-(x) \geq 0$, и заметить, что для φ^+ и φ^- формула (7.1) верна в силу доказанного выше.

Функция распределения случайной величины

Распределение P_ξ случайной величины ξ есть вероятностная мера на прямой. Вероятностные меры на прямой можно описывать и изучать в рамках двух эквивалентных подходов: функции распределения и характеристические функции. Ниже мы даем определение первого из этих понятий.

Определение 2. *Функцией распределения* случайной величины ξ называется функция $F_\xi : \mathbf{R} \to \mathbf{R}$, определяемая следующим образом

$$F_\xi(x) = P_\xi((-\infty, x]) = \mathsf{P}\{\omega : \xi(\omega) \le x\}.$$

Предложение 1 (Общие свойства функций распределения).

1. $0 \le F_\xi(x) \le 1$ для любого x.
2. $P_\xi(a, b] = F_\xi(b) - F_\xi(a)$.
3. $F_\xi(x)$ — неубывающая функция:

$$F_\xi(x_1) \le F_\xi(x_2), \quad \text{если } x_1 < x_2.$$

4. Пределы на бесконечности:

$$F_\xi(-\infty) \equiv \lim_{x\to-\infty} F_\xi(x) = 0,$$

$$F_\xi(+\infty) \equiv \lim_{x\to+\infty} F_\xi(x) = 1.$$

5. Функция $F_\xi(x)$ непрерывна справа в каждой точке:

$$F_\xi(x+0) \equiv \lim_{y\to x+0} F_\xi(y) = F_\xi(x).$$

Доказательство. 1, 2 — очевидно.

$$3.\quad \begin{aligned} F_\xi(x_2) - F_\xi(x_1) &= \mathsf{P}\{\omega : \xi(\omega) \le x_2\} - \mathsf{P}\{\omega : \xi(\omega) \le x_1\} = \\ &= \mathsf{P}\{\omega : \xi(\omega) \in (x_1, x_2]\} \ge 0. \end{aligned}$$

4. Докажем, что $F_\xi(+\infty) = 1$. Рассмотрим последовательность событий

$$A_n = \{\omega : \xi(\omega) \le n\}, \quad A_n \subseteq A_{n+1}, \quad \bigcup_n A_n = \Omega.$$

По свойству непрерывности вероятностной меры (см. лекцию 2) получим

$$\lim_{n\to\infty} F_\xi(n) = \lim_{n\to\infty} \mathsf{P}(A_n) = \mathsf{P}\big(\bigcup_n A_n\big) = \mathsf{P}(\Omega) = 1.$$

Аналогично доказывается свойство $F_\xi(-\infty) = 0$.

5. Пусть $x_n \downarrow x,\ n \to \infty$. Тогда

$$\{\omega : \xi(\omega) \le x_n\} \supseteq \{\omega : \xi(\omega) \le x_{n+1}\}$$

и

$$\{\omega:\ \xi(\omega)\le x\}=\bigcap_n\{\omega:\ \xi(\omega)\le x_n\}.$$

По непрерывности вероятностной меры

$$\lim_{n\to\infty}F_\xi(x_n)=\lim_{n\to\infty}\mathsf{P}\{\omega:\ \xi(\omega)\le x_n\}=\mathsf{P}\{\omega:\ \xi(\omega)\le x\}=F_\xi(x).$$

Предложение доказано.

Замечание. Любая функция, удовлетворяющая перечисленным в предложении 1 свойствам 1, 3, 4, 5, определяет свойством 2 (по теореме о продолжении меры) вероятностную меру на прямой. Поскольку мера P_ξ однозначно восстанавливается по функции распределения $F_\xi(x)$, интеграл Лебега $\int g(x)\,dP_\xi$ мы часто будем обозначать $\int g(x)\,dF_\xi(x)$.

Упражнения.

1. Показать, что у функции распределения $F_\xi(x)$ в каждой точке существует предел слева, т.е. существует

$$F_\xi(x-0)=\lim_{y\to x-0}F_\xi(x).$$

2. Показать, что множество точек разрыва функции $F_\xi(x)$ не более, чем счетно (x — точка разрыва, если $F_\xi(x)-F_\xi(x-0)>0$).

Замечание. Каков *вероятностный смысл* скачка функции распределения? Величина скачка в точке x равна вероятности того, что случайная величина примет значение x :

$$\mathsf{P}\{\omega:\ \xi(\omega)=x\}=F_\xi(x)-F_\xi(x-0).$$

Пример. Рассмотрим функцию распределения кусочно-постоянной (простой) случайной величины. Пусть

$$\Omega=A_1\cup A_2\cup\ldots\cup A_n,\quad A_i\cap A_j=\emptyset,\quad \mathsf{P}(A_i)=p_i>0,\quad \sum_{i=1}^n p_i=1.$$

Пусть $\xi(\omega)=x_i$ для $\omega\in A_i$ и $-\infty<x_1<x_2<\ldots<x_n<\infty$. Тогда

$$F_\xi(x)=\sum_{i:\,x_i\le x}p_i.$$

В частности, если $\xi=c$ константа, то

$$F_\xi(x)=\begin{cases}0, & \text{если } x<c,\\ 1, & \text{если } x\ge c.\end{cases}$$

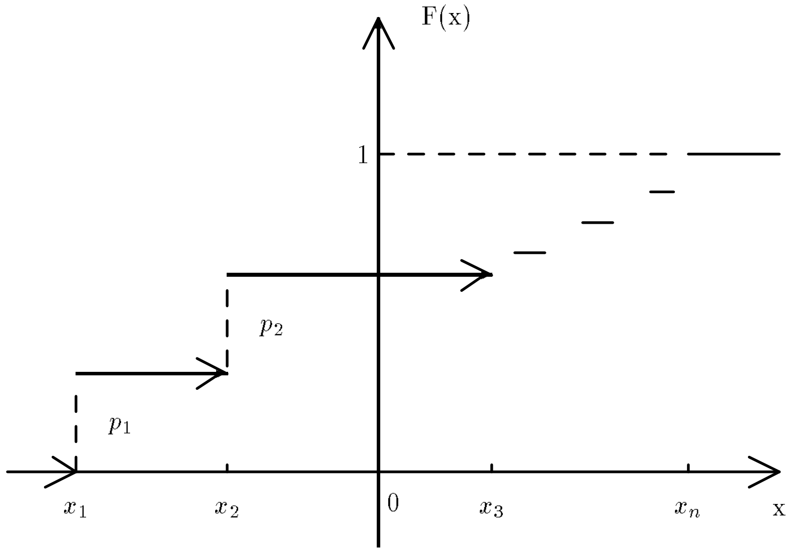


Рис. 7.1. Функция распределения дискретной случайной величины.

На рисунке 7.1 приведена функция распределения случайной величины, принимающей конечное число значений $x_1, \ldots, x_n$ с вероятностями $p_1, \ldots, p_n$ соответственно.

Замечание. Случайная величина ξ называется *дискретной*, если $F_\xi(x)$ кусочно-постоянна; *непрерывной*, если $F_\xi(x)$ непрерывна; *абсолютно непрерывной*, если у распределения случайной величины ξ существует плотность (см. ниже).

Плотность распределения

Определение 3. Мы будем называть распределение P_ξ *абсолютно непрерывным* (относительно меры Лебега), если существует неотрицательная измеримая функция $p_\xi(x)$, называемая *плотностью распределения*, такая, что для любого $B \in \mathcal{B}(\mathbf{R})$

$$P_\xi(B) = \int_B p_\xi(x)\, dx.$$

Ясно, что плотность определяется с точностью до значений на множествах лебеговой меры 0.

Приведем некоторые простейшие свойства плотности.

1. Полагая $B = (-\infty, \infty)$, получим

$$\int\limits_{\mathrm{R}} p_\xi(x)\,dx = 1$$

для любой плотности $p_\xi(x)$.

2. Полагая $B = (-\infty, x]$, получим связь плотности и функции распределения:

$$F_\xi(x) = \int\limits_{-\infty}^{x} p_\xi(x)\,dx \tag{7.2}$$

для каждого x.

3. Аналогично,

$$\mathsf{P}\{a \le \xi \le b\} = \int\limits_{a}^{b} p_\xi(x)\,dx. \tag{7.3}$$

4. Почти всюду по мере Лебега

$$\frac{d}{dx}F_\xi(x) = p_\xi(x).$$

Если $p_\xi(x)$ непрерывна, равенство имеет место всюду. (Доказательство см. в книге Колмогорова и Фомина, глава 6.)

Замечание. Если (7.2) имеет место для всех x, то распределение P_ξ абсолютно непрерывно. (Без доказательства.)

Теорема 2. Пусть распределение P_ξ случайной величины ξ абсолютно непрерывно, $p_\xi(x)$ — плотность распределения, φ — борелевская функция. Тогда

$$\mathsf{E}\varphi(\xi(\omega)) = \int\limits_{\mathrm{R}} \varphi(x)p_\xi(x)\,dx.$$

Математическое ожидание существует в том и только в том случае, когда существует интеграл в смысле Лебега в правой части.

Доказательство повторяет рассуждения, приведенные при доказательстве теоремы 1.

Примеры.

1. Равномерное распределение на отрезке $[a,b]$. По определению

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

Тогда

$$F_\xi(x) = \begin{cases} 0, & x \le a, \\ \dfrac{x-a}{b-a}, & a \le x \le b, \\ 1, & b \le x. \end{cases}$$

2. Нормальное (гауссовское) распределение $\xi \sim \mathcal{N}(a,\sigma^2)$, $a \in \mathbf{R}$, $\sigma > 0$:

$$p_\xi(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-a)^2}{2\sigma^2}\right).$$

Стандартное нормальное распределение $\mathcal{N}(0,1)$:

$$p_\xi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

Для подсчета нормирующей константы $1/\sqrt{2\pi}$ нужно вычислить интеграл $\int\limits_{\mathrm{R}} \exp(-x^2/2)\,dx$. Квадрат этого интеграла легко вычисляется при переходе к полярным координатам :

$$\int\limits_{\mathrm{R}} \int\limits_{\mathrm{R}} e^{-(\frac{x^2}{2}+\frac{y^2}{2})}\,dx\,dy = \int\limits_0^{2\pi} \int\limits_0^{\infty} e^{-\frac{r^2}{2}}\, r\,dr\,d\varphi = 2\pi.$$

Упражнения.

1. Пусть $\xi \sim \mathcal{N}(0,1)$. Показать, что $\eta = \sigma\xi + a \sim \mathcal{N}(a,\sigma^2)$.

2. Преобразования при сдвиге и скейлинге. Пусть m и $\sigma > 0$ некоторые константы. Доказать, что:

а) $F_{\xi+m}(x) = F_\xi(x-m)$;

б) $F_{\sigma\xi}(x) = F_\xi(x/\sigma)$;

в) $p_{\xi+m}(x) = p_\xi(x-m)$;

г) $p_{\sigma\xi+m}(x) = \sigma^{-1} p_\xi((x-m)/\sigma)$.

Многомерные распределения

Определение 4. Совместной функцией распределения случайных величин $\xi_1, \ldots, \xi_n$, заданных на одном вероятностном пространстве, назовем функцию n вещественных переменных

$$F_{\xi_1 \ldots \xi_n}(x_1, \ldots, x_n) \stackrel{\text{def}}{=} \mathsf{P}\{\omega : \xi_1(\omega) \le x_1, \ldots, \xi_n(\omega) \le x_n\}.$$

Определение 5. Совместное распределение случайных величин $\xi_1, \ldots, \xi_n$ называется *абсолютно непрерывным*, если существует неотрицательная измеримая функция $p_{\xi_1 \ldots \xi_n}(x_1, \ldots, x_n)$, называемая совместной плотностью, такая, что для всех множеств B из борелевской σ-алгебры $\mathcal{B}(\mathbf{R}^n)$

$$\mathsf{P}\{\omega : (\xi_1, \ldots, \xi_n) \in B\} = \int\limits_B p_{\xi_1 \ldots \xi_n}(x_1, \ldots, x_n)\, dx_1 \cdots dx_n.$$

Очевидные свойства плотности многомерного распределения:

$$\int\limits_{\mathrm{R}^n} p_{\xi_1 \ldots \xi_n}(x_1, \ldots, x_n)\, dx_1 \cdots dx_n = 1;$$

для всех $x_1, \ldots, x_n$

$$F_{\xi_1 \ldots \xi_n}(x_1, \ldots, x_n) = \int\limits_{-\infty}^{x_1} \cdots \int\limits_{-\infty}^{x_n} p_{\xi_1 \ldots \xi_n}(x_1, \ldots, x_n)\, dx_1 \cdots dx_n.$$

Аналогично теореме 2 для любой борелевской функции $g : \mathbf{R}^n \to \mathbf{R}$ имеет место формула

$$\mathsf{E} g(\xi_1, \ldots, \xi_n) = \int\limits_{\mathrm{R}^n} g(x_1, \ldots, x_n) p_{\xi_1 \ldots \xi_n}(x_1, \ldots, x_n)\, dx_1 \cdots dx_n. \tag{7.4}$$

Обе части конечны одновременно и равны в том случае, когда существуют. Доказательство аналогично.

Предложение 2. Совместная функция распределения независимых случайных величин есть произведение функций распределения этих случайных величин :

$$F_{\xi_1 \ldots \xi_n}(x_1, \ldots, x_n) = F_{\xi_1}(x_1) \cdots F_{\xi_n}(x_n). \tag{7.5}$$

Доказательство. Из независимости получаем

$$\mathsf{P}\{\xi_1 \le x_1, \ldots, \xi_n \le x_n\} = \mathsf{P}\{\xi_1 \le x_1\} \cdots \mathsf{P}\{\xi_n \le x_n\}.$$

Замечание. Верно и обратное утверждение. Для того чтобы случайные величины $\xi_1, \ldots, \xi_n$ были независимы, необходимо и достаточно, чтобы для всех $x_1, \ldots, x_n$ было выполнено (7.5). Доказательство можно посмотреть в учебнике Ширяева, глава II, §5.

Следствие. Если совместное распределение случайных величин $\xi_1, \ldots, \xi_n$ абсолютно непрерывно, то

$$\xi_1, \ldots, \xi_n \text{ независимы } \iff p_{\xi_1 \ldots \xi_n}(x_1, \ldots, x_n) \equiv p_{\xi_1}(x_1) \cdots p_{\xi_n}(x_n).$$

Свертка

Рассмотрим следующую задачу. Найти функцию распределения $F(x)$ суммы двух независимых случайных величин ξ_1 и ξ_2, если известны функции распределения слагаемых, $F_1(x)$ и $F_2(x)$ соответственно.

Определение 6. Сверткой двух функций распределения F_1 и F_2 называется функция

$$F_1 * F_2(x) \stackrel{\text{def}}{=} \int\limits_{\mathrm{R}} F_1(x-y)\, dF_2(y).$$

Предложение 3. Функция распределения суммы двух независимых случайных величин с функциями распределения F_1 и F_2 есть свертка их функций распределения :

$$F_{\xi_1+\xi_2}(x) = F_1 * F_2(x). \tag{7.6}$$

Доказательство. Обозначим $I_{\{\xi_1+\xi_2 \le x\}}(\omega)$ индикатор события $\{\xi_1 + \xi_2 \le x\}$. Тогда

$$F_{\xi_1+\xi_2}(x) = \mathsf{P}(\xi_1 + \xi_2 \le x) = \int\limits_{\Omega} I_{\{\xi_1+\xi_2 \le x\}}(\omega)\, d\mathsf{P}(\omega) =$$

$$= \int\limits_{y+z \le x}\!\!\int d\mathsf{P}_{\xi_1\xi_2} = \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{x-z} d\mathsf{P}_{\xi_1}(y)\, d\mathsf{P}_{\xi_2}(z) = \int\limits_{\mathrm{R}} F_1(x-z)\, dF_2(z).$$

Замечание. Из доказательства предложения 3 видно, что определение свертки симметрично:

$$F_1 * F_2(x) = \int\limits_R F_1(x-y)\ dF_2(y) = \int\limits_R F_2(y-x)\ dF_1(x).$$

Предложение 4. Пусть случайные величины ξ_1 и ξ_2 независимы и их распределения абсолютно непрерывны с плотностями $p_{\xi_1}(x)$ и $p_{\xi_2}(x)$. Тогда $\xi_1+\xi_2$ имеет абсолютно непрерывное распределение с плотностью

$$p_{\xi_1+\xi_2}(x) = \int\limits_{-\infty}^{+\infty} p_{\xi_1}(x-y)p_{\xi_2}(y)\,dy = \int\limits_{-\infty}^{+\infty} p_{\xi_2}(y-x)p_{\xi_1}(x)\,dx.$$

Доказательство. Для того чтобы сделать рассуждения более понятными, предположим, что плотности $p_{\xi_1}(x)$ и $p_{\xi_2}(x)$ непрерывны. Известно, что для непрерывных ограниченных функций интегралы Лебега и Римана совпадают. Обозначим $D=\{(x_1,x_2)\in\mathbf{R}^2 : x_1+x_2\le x\}$. Тогда

$$F_{\xi_1+\xi_2}(x) = \mathsf{P}\{\xi_1+\xi_2\le x\} = \mathsf{P}\{(\xi_1,\xi_2)\in D\} =$$

$$= \int\limits_D p_{\xi_1\xi_2}(x_1,x_2)\,dx_1dx_2 = \int\limits_{x_1+x_2\le x}\!\!\int p_{\xi_1}(x_1)p_{\xi_2}(x_2)\,dx_1dx_2 =$$

$$= \int\limits_{-\infty}^{\infty}\Bigg(\int\limits_{-\infty}^{x-x_2} p_{\xi_1}(x_1)\,dx_1\Bigg)p_{\xi_2}(x_2)\,dx_2 =$$

$$= \int\limits_{-\infty}^{x}\Bigg(\int\limits_{-\infty}^{\infty} p_{\xi_1}(x_1-x_2)p_{\xi_2}(x_2)\,dx_2\Bigg)\,dx_1.$$

Для доказательства формулы свертки в случае измеримых плотностей общего вида следует повторить приведенные выше выкладки, обосновывая их применением теоремы Фубини и теорем о предельном переходе под знаком интеграла Лебега.

Лемма 1. Если свернуть произвольную функцию распределения F с функцией распределения Φ_σ гауссовской случайной величины со средним 0 и дисперсией σ^2, то полученное распределение $F_\sigma=\Phi_\sigma * F$ имеет плотность.

Доказательство. Согласно свойству 4 плотностей и формуле (7.6) плотность свертки равна

$$p_\sigma(x) = \frac{d}{dx}\int\limits_{\mathrm{R}} \Phi_\sigma(x-z)dF(z) =$$

$$= \int\limits_{\mathrm{R}} \frac{d}{dx}\Phi_\sigma(x-z)dF(z) = \int\limits_{\mathrm{R}} \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-z)^2}{2\sigma^2}\right)dF(z). \quad (7.7)$$

Упражнение. Обоснуйте законность дифференцирования под знаком интеграла в приведенном выше доказательстве.

Лемма 2. В обозначениях леммы 1, если $\sigma \to 0+$, то $F_\sigma \to F$ в каждой точке непрерывности F .

Интуитивно утверждение леммы 2 ясно, т.к. к ξ прибавляется гауссова величина η_σ, которая с большой вероятностью близка к нулю. Приведем строгое доказательство.

Доказательство. Из неравенства Чебышева следует, что для любого данного $\varepsilon > 0$ при всех достаточно малых значениях σ

$$\mathsf{P}(|\eta_\sigma| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon} \leq \varepsilon.$$

Далее,

$$F_\sigma(x) = \mathsf{P}(\xi \leq x-\eta_\sigma) = \mathsf{P}(\xi \leq x-\eta_\sigma, |\eta|_\sigma \leq \varepsilon)+\mathsf{P}(\xi \leq x-\eta_\sigma, |\eta|_\sigma \geq \varepsilon).$$

Но

$$\mathsf{P}(\xi \leq x-\eta_\sigma, |\eta|_\sigma \leq \varepsilon) \leq \mathsf{P}(\xi \leq x+\varepsilon),$$

а для достаточно малых σ

$$\mathsf{P}(\xi \leq x-\eta_\sigma, |\eta|_\sigma \geq \varepsilon) \leq \mathsf{P}(|\eta|_\sigma \geq \varepsilon) \leq \varepsilon.$$

Итак,

$$F_\sigma(x) \leq F(x+\varepsilon)+\varepsilon. \quad (7.8)$$

С другой стороны,

$$\mathsf{P}(\xi \leq x-\eta_\sigma) \geq \mathsf{P}(\xi \leq x-\eta_\sigma, |\eta_\sigma| \leq \varepsilon) \geq \mathsf{P}(\xi \leq x-\varepsilon, |\eta_\sigma| \leq \varepsilon) =$$

$$= \mathsf{P}(\xi \leq x-\varepsilon) - \mathsf{P}(\xi \leq x-\varepsilon, |\eta_\sigma| > \varepsilon) \geq F(x-\varepsilon)-\varepsilon. \quad (7.9)$$

Окончательно получаем:

$$F(x-\varepsilon)-\varepsilon \leq F_\sigma(x) \leq F(x+\varepsilon)+\varepsilon,$$

и если x — точка непрерывности функции F, то отсюда следует утверждение леммы.

Литература

Ширяев А. Н. Вероятность, глава 2, §§3, 4. Колмогоров А. Н., Фомин С. В. Элементы теории функций и функционального анализа, глава 6. Феллер В. Введение в теорию вероятностей и ее приложения, том 2, глава1, §1; глава 3, §1; глава 5, §1.

Лекция 8
Характеристические функции

Определение 1. *Характеристической функцией* случайной величины ξ называется (комплекснозначная) функция на $\mathbf{R}$:

$$f(t) = f_\xi(t) = \mathsf{E}\exp(it\xi) = \int\limits_\Omega \exp(it\xi)d\mathsf{P} = \int\limits_\mathrm{R} \exp(itx)d\mathsf{P}_\xi(x) =$$

$$= \int\limits_\mathrm{R} \cos tx\, dF_\xi(x) + i \int\limits_\mathrm{R} \sin tx\, dF_\xi(x).$$

Свойства характеристической функции.

1. $f(0) = 1$, $|f(t)| \le 1$.

2. $f(t)$ непрерывна.

3. Выпуклость. Пусть $f_1, ..., f_n$ — характеристические функции, соответствующие функциям распределения $F_1, ..., F_n$. Тогда для любых $\alpha_1, ..., \alpha_n$ таких, что $\sum \alpha_i = 1$, функция $f(t) = \sum \alpha_i f_i(t)$ является характеристической функцией случайной величины с функцией распределения $\sum \alpha_i F_i$.

4. Характеристическая функция суммы двух независимых случайных величин равна произведению характеристических функций этих случайных величин. Если $\xi = \xi_1 + \xi_2$, где ξ_1 и ξ_2 независимы, то $f_{\xi_1+\xi_2}(t) = f_{\xi_1}(t) f_{\xi_2}(t)$.

5. Неотрицательная определенность: функция $f(t)$ является неотрицательно определенной, т.е. для любых $t_1, ..., t_n$ матрица $(f_{ij}) = (f(t_i - t_j))$ неотрицательно определена. В самом деле, для любого набора $(c_1, ..., c_n)$, $c_i \in \mathbf{C}$,

$$\sum_{k,j=1}^n f(t_k - t_j) c_k \bar{c}_j = \sum_{k,j=1}^n \int e^{i(t_k - t_j)x} c_k \bar{c}_j d\mathsf{P}_\xi(x) =$$

$$= \int \left| \sum_{k=1}^n c_k e^{it_k x} \right|^2 d\mathsf{P}_\xi(x) \ge 0.$$

Упражнение 1. Доказать свойства 1 – 4.

Различные частные случаи и свойства.

1. Если ξ имеет плотность $p(x)$, то по теореме 2 лекции 7

$$f(t) = \int p(x) \exp(itx)\, dx$$

есть просто преобразование Фурье плотности; при этом формула обращения (т.е. формула, восстанавливающая плотность по характеристической функции) есть обратное преобразование Фурье :

$$p(x) = \frac{1}{2\pi} \int_{\mathrm{R}} f(t) \exp(-itx)\, dt.$$

Например, характеристическая функция стандартной гауссовой величины $\xi \sim \mathcal{N}(0,1)$ равна

$$\int_{\mathrm{R}} \exp(itx) p(x)\, dx = \frac{1}{\sqrt{2\pi}} \int_{\mathrm{R}} \exp\left(itx - \frac{x^2}{2}\right) dx = \exp\left(-\frac{t^2}{2}\right). \quad (8.1)$$

(Для вычисления интеграла выделить полный квадрат в экспоненте и сделать замену переменных.) Таким образом, с точностью до константы эта плотность самодвойственна.

2. Если $|\xi| \le C$, то $f_\xi(t)$ аналитична на всей комплексной плоскости, т.е. разлагается в абсолютно сходящийся ряд

$$f_\xi(t) = \sum_{k=0}^{\infty} m_k \frac{1}{k!} (it)^k,$$

где $m_k = \mathsf{E}\xi^k$ — k-й момент случайной величины ξ. Отсюда, например, получаем полезные формулы для моментов случайной величины

$$\mathsf{E}\xi^k = \frac{1}{i^k} \frac{d^k}{dt^k} f_\xi(t) \Big|_{t=0}.$$

Наряду с моментами, важными характеристиками вероятностного распределения являются *семиинварианты.* Разлагая аналогичным образом в ряд логарифм характеристической функции, получим в качестве коэффициентов разложения семиинварианты s_k:

$$\log f_\xi(t) = \sum_{k=1}^{\infty} s_k \frac{1}{k!} (it)^k; \quad s_k = \frac{1}{i^k} \frac{d^k}{dt^k} \log f_\xi(t) \Big|_{t=0}.$$

3. Если ξ принимает целочисленные значения n с вероятностями p_n, то

$$f(\lambda) = \sum_{n=-\infty}^{\infty} p_n e^{in\lambda}.$$

В этом случае удобнее рассматривать *производящую функцию*

$$\Phi(z) = \sum_{n=-\infty}^{\infty} p_n z^n,$$

где $z = e^{i\lambda}$. Формула обращения в этом случае дается интегралом Коши

$$p_n = \frac{1}{2\pi i} \int\limits_{|z|=1} \frac{\Phi(z)}{z^{n+1}}\, dz.$$

4. Пусть случайная величина имеет областью значений отрезок $[-\pi, \pi]$ и имеет гладкую плотность $p(x)$. Тогда плотность разлагается в ряд Фурье:

$$p(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx},$$

а значения характеристической функции в целых точках есть просто коэффициенты ряда Фурье :

$$f_\xi(n) = \mathsf{E} \exp(in\xi) = \int\limits_{-\pi}^{\pi} \exp(inx) p(x)\, dx = c_n.$$

Упражнение 2. Доказать перечисленные выше утверждения 1 – 4.

Формула обращения. Теорема единственности

Докажем, что функция распределения однозначно определяется своей характеристической функцией и приведем формулу обращения явно. Частные случаи этой формулы мы привели выше в утверждениях 1 – 4.

Теорема 1. Пусть $F(x)$ — функция распределения случайной величины ξ и $f(t)$ — её характеристическая функция. Для любых двух точек x, y $(x \leq y)$, в которых F непрерывна, имеет место равенство:

$$F(x) - F(y) = \lim_{\sigma \to 0} \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} f(t) \frac{e^{-itx} - e^{-ity}}{it} \exp\left(-\frac{\sigma^2 t^2}{2}\right) dt.$$

Доказательство. Рассмотрим свертку F_σ функции распределения F с функцией распределения гауссовской случайной величины с нулевым средним и дисперсией σ^2 (см. лекцию 7). В силу (8.1) и свойства 4 характеристическая функция свертки $f_\sigma(t)$ равна

$$f_\sigma(t) = f(t) \exp\left(-\frac{t^2\sigma^2}{2}\right),$$

где $f(t)$ — характеристическая функция для распределения F. Поскольку по лемме 1 лекции 7 существует плотность p_σ для свертки F_σ, то из свойства 1 получаем

$$p_\sigma(x) = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \exp(-itx) \exp\left(-\frac{t^2\sigma^2}{2}\right) f(t)\, dt.$$

Поэтому

$$F_\sigma(y) - F_\sigma(x) = \frac{1}{2\pi} \int\limits_{x}^{y} \int\limits_{-\infty}^{\infty} \exp(-itz) f(t) \exp\left(-\frac{t^2\sigma^2}{2}\right) dt\, dz =$$

$$= \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} f(t) \frac{e^{-ity} - e^{-itx}}{-it} \exp\left(-\frac{t^2\sigma^2}{2}\right) dt.$$

Переходя теперь к пределу при $\sigma \to 0$ и пользуясь леммой 2 из лекции 7, получим утверждение теоремы.

Следствие (Теорема единственности). Функция распределения случайной величины однозначно восстанавливается по ее характеристической функции.

Доказательство. Так как F непрерывна справа, то она однозначно восстанавливается по своим значениям в точках непрерывности, следовательно, она единственным образом восстанавливается по своей характеристической функции.

Слабая сходимость

В предыдущих лекциях мы определили сходимость последовательности случайных величин по вероятности и с вероятностью 1. Еще один вид сходимости — слабая сходимость — определяется в терминах функций распределения.

Определение 2. Последовательность функций распределения $\{F_n\}$ *слабо сходится* к функции распределения F, если $F_n(x) \to F_n(x)$ в любой точке непрерывности функции F.

Замечание 1. Обычная поточечная сходимость функций распределения может плохо отражать стремление ξ_n к ξ при $n \to \infty$. Так, равномерно распределенная на $[a, b]$ случайная величина, очевидно, стремится к постоянной, равной a, при $b \downarrow a$. Но в точке a поточечной сходимости соответствующих функций распределения нет. В самом деле, $F_n(a) = 0$ для всех n, но $F(a) = 1$.

Слабую сходимость функций распределения можно эквивалентно задать в терминах слабой сходимости соответствующих мер на прямой.

Определение 3. Последовательность вероятностных мер $\{P_n\}$ на прямой слабо сходится к вероятностной мере P (обозначаем $P_n \Rightarrow P$), если для любой ограниченной непрерывной функции φ

$$\int\limits_{\mathrm{R}} \varphi \, dP_n \longrightarrow \int\limits_{\mathrm{R}} \varphi \, dP.$$

Лемма 1. Пусть задана последовательность функций распределения $\{F_n\}$ и пусть F также функция распределения. Пусть $\{P_n\}$ и P соответствующие им вероятностные меры на прямой. Тогда определения 2 и 3 эквивалентны, т.е. $F_n(x) \to F(x)$ слабо тогда и только тогда, когда $P_n \Rightarrow P$.

Эквивалентность этих видов сходимости представляется естественной в силу следующих рассуждений.

1. Пусть $P_n \Rightarrow P$. Пусть x — точка непрерывности F. Приблизим индикатор $I_{(-\infty,x]}$ полуинтервала $(-\infty, x]$ непрерывными ограниченными функциями $\varphi_{\varepsilon,\delta}$, $\varphi_{\varepsilon,\delta} < C < \infty$. Конечно, из-за разрыва индикатора в точке x это приближение не будет равномерным, но для заданных $\varepsilon > 0, \delta > 0$ функцию $\varphi_{\varepsilon,\delta}$ можно выбрать так, что окрестность M точки x, где $\varphi_{\varepsilon,\delta}$ отличается от $I_{(-\infty,x]}$ больше, чем на ε, имеет лебегову меру, не превосходящую $\delta > 0$:

$$M = \{y : |\varphi_{\varepsilon,\delta}(y) - I_{(-\infty,x]}(y)| \geq \varepsilon\}, \quad \mu(M) \leq \delta.$$

Тогда $F_n(x) = \int\limits_{\mathrm{R}} I_{(-\infty,x]}\,dP_n$ и $\int\limits_{\mathrm{R}} \varphi_{\varepsilon,\delta}\,dP_n$ отличаются друг от друга не более чем на $\varepsilon + (C+1)\int\limits_{M} dP_n$. В силу слабой сходимости $P_n \Rightarrow P$ при больших n интеграл $\int\limits_{\mathrm{R}} \varphi_{\varepsilon,\delta}\,dP_n$ мало отличается от $\int\limits_{\mathrm{R}} \varphi_{\varepsilon,\delta}\,dP$, а $\int\limits_{\mathrm{R}} \varphi_{\varepsilon,\delta}\,dP$ отличается от $\int\limits_{\mathrm{R}} I_{(-\infty,x]}\,dP$ не более чем на $\varepsilon + (C+1)\int\limits_{M} dP$. Заметим, что $\int\limits_{M} dP \to 0$ при $\delta \to 0$, так как x — точка непрерывности для F, а $\int\limits_{M} dP_n \to 0$ при $\delta \to 0$, иначе это противоречило бы условию $P_n \Rightarrow P$. Осталось заметить, что ε и δ выбирались произвольно малыми.

2. Пусть $F_n \to F$ слабо, и пусть x — точка непрерывности F. Тогда для индикаторов $I_{(-\infty,x]}(y)$ имеет место сходимость интегралов

$$\int\limits_{\mathrm{R}} I_{(-\infty,x]}\,dP_n \longrightarrow \int\limits_{\mathrm{R}} I_{(-\infty,x]}\,dP.$$

С помощью линейной комбинации таких индикаторов мы можем получить произвольную ступенчатую функцию. Пусть задана ограниченная непрерывная функция φ. Тогда на любом фиксированном отрезке $[a,b]$, где a и b точки непрерывности F, мы можем равномерно приблизить φ ступенчатой функцией L. При этом, если отрезок достаточно велик, то, в силу свойств функции распределения, его дополнение имеет малую P-меру, а в силу слабой сходимости $F_n \to F$ и малую P_n-меру. Тогда

$$\Big|\int\limits_{[a,b]} \varphi\,dP_n - \int\limits_{[a,b]} \varphi\,dP\Big| \le$$

$$\le \Big|\int\limits_{[a,b]} (\varphi - L)\,dP_n\Big| + \Big|\int\limits_{[a,b]} L\,dP_n - \int\limits_{[a,b]} L\,dP)\Big| + \Big|\int\limits_{[a,b]} (\varphi - L)\,dP\Big|.$$

Правую часть можно сделать произвольно малой за счет выбора L и достаточно больших n. Интегралы $\int\limits_{\overline{[a,b]}} \varphi\,dP_n$ и $\int\limits_{\overline{[a,b]}} \varphi\,dP$ по внешности $\overline{[a,b]}$ отрезка $[a,b]$ становятся малыми при увеличении отрезка $[a,b]$.

Приведем строгое доказательство леммы 1 "в одну сторону": если $P_n \Rightarrow P$, то $F_n \to F$ слабо.

Доказательство. Для любых фиксированных $\varepsilon > 0$ и $x > 0$ рассмотрим непрерывную функцию $\varphi_\varepsilon(t)$, равную 1 при $t \le x$, 0 при $t \geq x + \varepsilon$ и линейную на $[x, x+\varepsilon]$. Тогда

$$F_n(x) \leq \int\limits_{-\infty}^{x} \varphi_\varepsilon \, dP_n \leq \int\limits_{-\infty}^{\infty} \varphi_\varepsilon \, dP_n \longrightarrow \int\limits_{-\infty}^{\infty} \varphi_\varepsilon \, dP \leq F(x+\varepsilon).$$

Устремляя $\varepsilon \to 0$ и учитывая непрерывность F справа, получим, что для любого $\delta > 0$ для достаточно больших n

$$F_n(x) \leq F(x) + \delta.$$

Возьмем далее непрерывную функцию f_ε, равную 1 при $t \leq x - \varepsilon$, 0 при $t \geq x$ и линейную на $[x-\varepsilon, x]$. Аналогично получаем, что что для любого $\delta > 0$ для достаточно больших n

$$F_n(x) \geq F(x-0) - \delta.$$

Так как x есть точка непрерывности, то имеем:

$$|F_n(x) - F(x)| \leq \delta.$$

Упражнение 3. Провести полное доказательство пункта 2 леммы 1.

В дальнейшем нам понадобится следующий результат.

Теорема 2 (Теорема непрерывности). Пусть $\{F_n\}$ — функции распределения и $\{f_n\}$ — их характеристические функции, причем существует

$$f(t) = \lim_{n\to\infty} f_n(t),$$

и $f(t)$ непрерывна в нуле. Тогда F_n слабо сходятся к некоторой функции распределения F.

Замечание 2. В формулировке теоремы 2 неявно содержится утверждение о том, что $f(t)$ является характеристической функцией соответствующей функции распределения F.

Замечание 3. Из слабой сходимости $P_n \Rightarrow P$ следует сразу сходимость характеристических функций $f_n \to f$; в теореме 2 утверждается, что если $f_n \to f$ и f — характеристическая функция, то $P_n \Rightarrow P$ (и тогда $F_n \to F$ слабо).

Для доказательства теоремы 2 нам понадобится определение слабой компактности семейства вероятностных мер и несколько сопутствующих результатов.

Слабая компактность

Определение 4. Множество вероятностных мер $\{P_n\}$ на прямой называется *слабо компактным*, если из любого его бесконечного подмножества можно выбрать слабо сходящуюся подпоследовательность $\{P_{n_k}\}$.

Следующая лемма позволит яснее представить себе критерий слабой компактности, который будет приведен немного позже.

Лемма 2 (Теорема Хелли или "принцип выбора"). Если на конечном отрезке $[a,b]$ прямой задана бесконечная последовательность вероятностных мер $\{P_n\}$, то она слабо компактна.

Доказательство. Рассмотрим множество всех мономов вида x^n, $n = 0, 1, 2, \ldots$ Выберем поднабор $\{n_0\}$, для которого сходится подпоследовательность $\Big\{\int\limits_{[a,b]} x^0\, dP_{n_0}\Big\}$, из него выделим поднабор $\{n_1\}$, для которого сходится $\Big\{\int\limits_{[a,b]} x^1\, dP_{n_1}\Big\}$ и так далее. Построим диагональную последовательность $\{P_{n'}\}$, взяв первый член из $\{P_{n_0}\}$, второй из $\{P_{n_1}\}$ и т.д. Тогда для любого $m = 0, 1, 2\ldots$ имеет место сходимость последовательности $\Big\{\int\limits_{[a,b]} x^m\, dP_{n'}\Big\}$.

Следовательно, для любого полинома $\pi(x)$ имеет место сходимость последовательности $\Big\{\int\limits_{[a,b]} \pi(x)\, dP_{n'}(x)\Big\}$.

Из теории функций хорошо известно, что на отрезке любую непрерывную ограниченную функцию φ можно равномерно приблизить полиномами. Тогда для данного $\varepsilon > 0$ найдется полином $\pi(x)$ степени n такой, что $\sup\limits_{[a,b]} |\varphi(x) - \pi(x)| < \varepsilon$. Следовательно, $\int\limits_{[a,b]} \pi(x)\, dP_{n'}(x)$ и $\int\limits_{[a,b]} \varphi(x)\, dP_{n'}$ отличаются друг от друга не более чем на ε. Отсюда следует сходимость последовательности $\Big\{\int\limits_{[a,b]} \varphi(x)\, dP_{n'}\Big\}$.

Теорема Прохорова (Критерий слабой компактности). Последовательность мер $\{P_n\}$ на прямой слабо компактна тогда и только тогда, когда для любого $\varepsilon > 0$ существует интервал $K(\varepsilon)$ такой, что для всех n, кроме конечного числа,

$$P_n(\mathbf{R} \setminus K(\varepsilon)) < \varepsilon.$$

Доказательство аналогично доказательству предыдущей леммы. Его можно найти в книге Гихман И.И., Скороход А.В. Теория случайных процессов, том 1, гл. 6, §1. М., Наука, 1971.

Вернемся к теореме 2. Непрерывность характеристической функции $f(t)$ при $t = 0$ влечет слабую компактность семейства мер $\{P_n\}$. Это вытекает из следующей леммы.

Лемма 3. Если ξ — случайная величина с характеристической функцией f, то при любом $u > 0$

$$P\left(|\xi| > \frac{2}{u}\right) \le \frac{2}{u}\int\limits_0^u (1 - Ref(t))\,dt. \tag{8.2}$$

Доказательство. Подставив в правую часть (8.2)

$$Ref(t) = \int\limits_{-\infty}^{\infty} \cos tx\,dF(x),$$

где F — функция распределения ξ, и меняя порядок интегрирования, имеем:

$$\frac{2}{u}\int\limits_0^u (1 - ef(t))\,dt = \frac{2}{u}\int\limits_0^u \int\limits_{-\infty}^{\infty} (1 - \cos tx)\,dF(x)\,dt =$$

$$= 2\int\limits_{-\infty}^{\infty} \Big(\frac{1}{u}\int\limits_0^u (1 - \cos tx)\,dt\Big)\,dF(x) = 2\int\limits_{-\infty}^{\infty} \left(1 - \frac{\sin ux}{ux}\right)\,dF(x).$$

Последний интеграл оценивается снизу так :

$$2\int\limits_{-\infty}^{\infty} \left(1 - \frac{\sin ux}{ux}\right)\,dF(x) \ge 2\int\limits_{|ux|\ge 2} \left(1 - \frac{1}{|ux|}\right)\,dF(x) \ge$$

$$\ge \int\limits_{|x|\ge 2/u} dF(x) \ge P\left(|\xi| > \frac{2}{u}\right).$$

Перейдем к доказательству теоремы непрерывности.

Доказательство теоремы 2. Так как f непрерывна в нуле и $f(0) = 1$, то для любого $\varepsilon > 0$ можно найти u такое, что

$$0 \le \frac{1}{u}\int\limits_0^u (1 - Ref(t))\,dt \le \varepsilon.$$

Но в силу леммы 3 и конечности u

$$P_n\left(\mathbf{R}\setminus\left[-\frac{2}{u},\frac{2}{u}\right]\right)\le\int\limits_{|x|\ge 2/u} dF_n(x)\le$$

$$\le\frac{2}{u}\int\limits_0^u(1-Ref_n(t)\,dt\longrightarrow\frac{2}{u}\int\limits_0^u(1-Ref(t))\,dt\le 2\varepsilon.$$

Таким образом, по теореме Прохорова получаем, что семейство $\{P_n\}$ слабо компактно. Покажем, что P_n слабо сходится к некоторой вероятностной мере. Выберем подпоследовательность $\{P_{n'}\}$, которая в силу слабой компактности сходится к некоторой вероятностной мере P. Пусть $\{P_n\}$ не сходится слабо к P, т.е. для некоторой ограниченной непрерывной функции φ интеграл $\int\varphi\,dP_n$ не сходится к $\int\varphi\,dP$. Выберем подпоследовательность $\{P_{\tilde n}\}$ такую, что $\int\varphi\,dP_{\tilde n}\to C\neq\int\varphi\,dP$. Тогда из этой подпоследовательности можно извлечь другую подпоследовательность $P_{n''}$, слабо сходящуюся к некоторой мере Q, отличной от P. Следовательно, в силу существования предела последовательности характеристических функций $\lim\limits_n f_n$

$$\lim_{n'}\int\limits_{\mathrm{R}} e^{itx}\,dP_{n'}=\lim_{n''}\int\limits_{\mathrm{R}} e^{itx}\,dP_{n''},$$

то есть

$$\int\limits_{\mathrm{R}} e^{itx}\,dP=f(t)=\int\limits_{\mathrm{R}} e^{itx}\,dQ,\qquad t\in\mathbf{R}.$$

Тогда f есть характеристическая функция двух разных мер, что невозможно согласно теореме единственности (см. следствие теоремы 1).

Литература

Ширяев А. Н. Вероятность, глава 2, §12, глава 3, §§1, 2.

Лекция 9

Доказательство предельных теорем методом характеристических функций. Центральная предельная теорема. Многомерное нормальное распределение

В этой лекции мы приведем ряд важных теорем, доказательство которых использует теорему непрерывности, сформулированную на прошлой лекции.

Блуждание на окружности

Рассмотрим независимые одинаково распределенные случайные величины $\{\xi_j\}, j = 1, 2, \ldots$ с областью значений $[-\pi, \pi)$ и гладкой на $\mathbf{R}$ плотностью $p(x)$, что, конечно, влечет $p(-\pi) = p(\pi) = 0$. Плотность разлагается в ряд Фурье

$$p(x) = \sum_{k=-\infty}^{\infty} c_k \, e^{ikx}, \tag{9.1}$$

где, очевидно, $c_0 = 1$.

Коэффициенты $\{c_k\}$ однозначно определяют плотность $p(x)$, но, с другой стороны, коэффициенты c_k есть значения характеристической функции f в целых точках: $c_k = f(-k)$, поэтому значения f в целых точках однозначно определяют плотность p.

Обозначим

$$S_n = \sum_{j=1}^{n} \xi_j; \quad S_n^0 = S_n \ (\bmod 2\pi).$$

Таким образом, если рассматривать ξ_j как величину скачка в момент j блуждающей по окружности частицы, то S_n^0 описывает положение этой частицы в момент n. Пусть $p_n^0(x)$ — плотность случайной величины S_n^0. Она представима в виде ряда Фурье

$$p_n^0(x) = \sum_{k=-\infty}^{\infty} c_k^n \, e^{ikx},$$

где c_k^n есть n-я степень коэффициента c_k из (9.1).

Упражнение 1. Доказать, что p_n^0 имеет указанное разложение Фурье.

Лемма 1. Пусть плотность p случайных величин ξ_j достаточно гладкая. Тогда последовательность случайных величин S_n^0 слабо сходится при $n \to \infty$ к равномерно распределенной на единичной окружности случайной величине r.

Доказательство. Напомним, что плотность случайной величины r равна $p_r(x) = 1/(2\pi)$, а характеристическая функция

$$\varphi_r(t) = \frac{\sin \pi t}{\pi t}, \quad t \in \mathbf{R}.$$

Мы покажем, что имеет место сходимость характеристических функций случайных величин S_n^0 к φ_r. Характеристическая функция S_n^0 равна

$$\varphi_n^0(t) = \int_{-\pi}^{\pi} e^{itx} \, p_n^0(x) \, dx = \sum_{k=-\infty}^{\infty} c_k^n \, \frac{\sin \pi(t+k)}{\pi(t+k)}.$$

Если $|c_k| < 1$ при $k \neq 0$, то при $n \to \infty$ значения φ_r и $\lim \varphi_n^0$ в целых точках совпадают, что влечет совпадение распределений.

Для больших k малость соответствующих коэффициентов Фурье следует при достаточной гладкости p из формулы интегрирования по частям :

$$c_k = \int_{-\pi}^{\pi} p(x) \, e^{-itx} \, dx = p(x) \, \frac{e^{-ikx}}{-ik} \bigg|_{-\pi}^{\pi} + \int_{-\pi}^{\pi} p'(x) \, \frac{e^{-itx}}{ik} \, dx,$$

откуда $|c_k| \to 0$ при $k \to \infty$.

Задача. Найти условия на плотность p, достаточные для того, чтобы $|c_k| < 1$ при $k \neq 0$.

Теорема Пуассона

Определение 1. Случайная величина ξ, принимающая неотрицательные целые значения с вероятностями

$$p_k = \mathsf{P}(\xi = k) = \frac{\lambda^k}{k!} e^{-\lambda},$$

где $\lambda > 0$ некоторое число, имеет *пуассоновское* распределение с параметром λ.

Упражнение 2. Доказать, что математическое ожидание, дисперсия и характеристическая функция пуассоновской случайной величины с параметром λ равны, соответственно,

$$\mathsf{E}\xi = \lambda, \quad \mathsf{D}\xi = \lambda, \quad \varphi_\xi(t) = \exp\{\lambda(e^{it} - 1)\}. \tag{9.2}$$

Теорема 1 (Теорема Пуассона). Рассмотрим схему Бернулли с n испытаниями. Пусть вероятность успеха p зависит от n, причем так, что $p(n) \to 0$ при $n \to \infty$, а $np(n) \to \lambda, \lambda > 0$. Пусть, как обычно, $S_n = \sum\limits_{i=1}^{n} \xi_i$ — число успехов в серии длины n. Тогда

$$P(S_n = m) \longrightarrow \frac{e^{-\lambda}\lambda^m}{m!}$$

для любого фиксированного $m, m = 0, 1, 2, \ldots$

Доказательство. Характеристическая функция бернуллиевской случайной величины равна $p \exp^{it} + q$. Характеристическая функция суммы n независимых бернуллиевских случайных величин

$$\varphi_{S_n}(t) = \mathsf{E}e^{itS_n} = \prod_{k=1}^{n}(pe^{it} + q) = \prod_{k=1}^{n}(1 + p(e^{it} - 1)) \sim$$

$$\sim \left(1 + \frac{\lambda}{n}\left(e^{it} - 1\right)\right)^n \longrightarrow \exp\{\lambda(e^{it} - 1)\},$$

при $n \to \infty$, если $p \to 0, \;\; p \cdot n \to \lambda$.

Центральная Предельная Теорема

Теорема 2 (Центральная Предельная Теорема). Пусть случайные величины $\{\xi_j\}_{j=1}^{\infty}$ *независимы, одинаково распределены*, имеют нулевое среднее и ограничены: $\mathsf{E}\xi_j = 0$, $\mathsf{P}(|\xi_j| \le C) = 1$ для некоторого $0 < C < \infty$. Тогда

$$\frac{S_n}{\sqrt{\mathsf{D}S_n}} \longrightarrow \mathcal{N}(0, 1)$$

в смысле слабой сходимости распределений. Таким образом, для любого открытого множества $A \subset \mathbf{R}^1$

$$\mathsf{P}\left(\frac{S_n}{\sqrt{\mathsf{D}S_n}} \in A\right) \longrightarrow \Phi(A) = \frac{1}{\sqrt{2\pi}} \int\limits_A e^{-\frac{x^2}{2}}\,dx \qquad \text{при} \quad n \to \infty.$$

Замечание 1. Поскольку $\mathsf{D}S_n = D \cdot n$, где $D = \mathsf{D}\xi_i$, то из центральной предельной теоремы следует, что S_n имеет порядок $\sqrt{n}$ при $n \to \infty$.

Замечание 2. Центральная предельная теорема (ЦПТ) существует в большом количестве вариантов, например, для некоторых случаев, когда случайные величины являются слабо зависимыми. Заметим, что получить предельное распределение с помощью сверток можно довольно редко, также как и вычислить в общем случае (например, отказавшись от независимости) характеристическую функцию для S_n, но в *независимом* случае (тем более, когда случайные величины ограничены) это легко.

Доказательство ЦПТ. Фиксируем $t \in \mathbf{R}$ и рассмотрим характеристическую функцию суммы S_n в точке t. Имеем

$$\varphi_{S_n}(t) = \mathsf{E}\exp(itS_n/\sqrt{\mathsf{D}S_n}) = \prod_{k=1}^{n} \mathsf{E}\exp(it\xi_k/\sqrt{\mathsf{D}S_n}) = f^n\left(\frac{t}{\sqrt{\mathsf{D}S_n}}\right),$$

где $f(t) = \mathsf{E}\exp it\xi_1$ — характеристическая функция случайной величины ξ_1. Обозначим $\sigma^2 = \mathsf{D}\xi_1$, тогда $\sqrt{\mathsf{D}S_n} = \sigma\sqrt{n}$. Разложив экспоненту в ряд Тейлора, получим:

$$f(t/\sqrt{\mathsf{D}S_n}) = \mathsf{E}\Big(1 + \frac{i\xi_1 t}{\sigma\sqrt{n}} + \frac{1}{2!}\Big(\frac{i\xi_1 t}{\sigma\sqrt{n}}\Big)^2 + \ldots\Big) =$$

$$= 1 + 0 - \frac{t^2}{2n} + \mathsf{E}\Big(\sum_{k=3}^{\infty} \frac{(it\xi_1)^k}{k!\sigma^k n^{k/2}}\Big),$$

причем легко видеть, что в силу ограниченности ξ_1 последнее слагаемое

$$\left|\mathsf{E}\Big(\sum_{k=3}^{\infty} \frac{(it\xi_1)^k}{k!\sigma^k n^{k/2}}\Big)\right| \le \sum_{k=3}^{\infty} \frac{(tC)^k}{k!\sigma^k n^{k/2}} = o\Big(\frac{1}{n}\Big) \qquad \text{при} \quad n \to \infty.$$

Поэтому

$$f\left(t/\sqrt{\mathsf{D}S_n}\right) = 1 - \frac{t^2}{2n} + o\Big(\frac{1}{n}\Big).$$

Тогда

$$f^n(t/\sqrt{\mathsf{D}S_n}) \to \exp\left(-\frac{t^2}{2}\right) \quad \text{при} \quad n \to \infty.$$

Остается только вспомнить, что полученная функция является характеристической для стандартного нормального распределения.

Скейлинги

В осях пространства-времени (t, x) изменение масштаба (scaling) означает преобразование $(t, x) \to (\alpha t, \beta x)$.

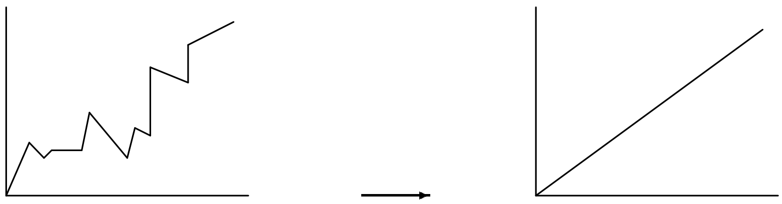
Сумму S_n независимых одинаково распределенных случайных величин, фигурирующих в законе больших чисел и центральной предельной теореме, можно рассматривать как положение случайного блуждания на вещественной прямой в момент времени n, при этом ξ_j есть скачок в момент j (вообще говоря, не целочисленный!). Рассмотрим такое блуждание в осях (n, S_n).

Закон больших чисел приобретает наглядность при уменьшении масштаба по времени. Будем измерять время не единичными отрезками, а большими интервалами. Единицу времени положим равной N. Таким образом, $n = \tau N$ для некоторого вещественного τ. Тогда отрезок времени $n = \tau N$ в новом масштабе равен τ, т.е. произведена замена (скейлинг) $N\tau \to \tau$. Зафиксируем τ, а масштаб N будем предполагать переменной величиной, в дальнейшем переходя к пределу $N \to \infty$. Одновременно уменьшим масштаб по пространству в N раз: $S_{[\tau N]} \to \frac{1}{N} S_{[\tau N]}$, как при составлении географической карты. Зафиксируем число τ, называемое макроскопическим временем, и устремим N к бесконечности: $N \to \infty$. Такой предельный переход называется скейлинговым (масштабным) пределом и обозначается $N\tau \xrightarrow{s} \tau$, $Nx \xrightarrow{s} x$. Предельное пространство-время теперь непрерывно. При этом

$$\lim_{N\to\infty} \frac{1}{N} S_{[\tau N]} = m\tau,$$

т.е. в макроскопическом пространстве-времени случайное блуждание превращается в детерминированное равномерное движение со скоростью m (см. рисунок 9.1).

Совсем другой скейлинг делается в центральной предельной теореме. Рассмотрим случай отсутствия сноса, т.е. $\mathsf{E}\xi_k = 0$. Как и прежде, изменим масштаб по времени линейным образом: $n = \tau N$. Однако масштаб по пространству уменьшим лишь в $\sqrt{N}$ раз:



$(n\tau, nx) \to (\tau, x)$

Рис. 9.1.

$S_{[\tau N]} \to \frac{1}{\sqrt{N}} S_{[\tau N]}$, т.е. масштаб времени меняется примерно как квадрат изменения масштаба пространства. Перейдем теперь к масштабному пределу $N \to \infty$ (такой предельный переход будем обозначать так: $N\tau \xrightarrow{s} \tau$, $\sqrt{N}x \xrightarrow{s} x$). При этом мы получим набор случайных величин

$$w(\tau) = \lim_{N\to\infty} \frac{1}{\sqrt{N}} S_{[\tau N]},\ \tau \in [0, \infty),$$

который является знаменитым винеровским процессом (броуновским движением).

Задача. Доказать, что для любых $\tau, \sigma \in [0, \infty)$ пара случайных величин $(w(\tau), w(\sigma))$ есть двумерный нормальный вектор (см. ниже), и ковариация его компонент равна

$$\mathsf{E} w(\tau) w(\sigma) = \min(\tau, \sigma).$$

Многомерные нормальные распределения

Мы будем строить произвольные гауссовы системы, исходя из системы $\xi_1, \xi_2, \ldots$ независимых одинаково распределенных гауссовых величин с нулевым средним и единичной дисперсией (стандартного гауссовского вектора). Напомним, что гауссова случайная величина с нулевым средним и единичной дисперсией имеет плотность $(1/\sqrt{2\pi})\exp(-x^2/2)$, а ее характеристическая функция равна $\exp(-t^2/2)$. Дадим определение характеристической функции случайного вектора (т.е. конечного набора случайных величин).

Определение 1. Характеристической функцией случайного вектора $\zeta = (\zeta_1, \zeta_2, ..., \zeta_n)$ называется следующая функция n вещественных переменных $(\tau_1, \tau_2, ..., \tau_n)$:

$$f_\zeta(\tau) = f_\zeta(\tau_1, \tau_2, ..., \tau_n) = \mathsf{E}\exp(i\sum_{k=1}^{n}\tau_k\zeta_k) = \mathsf{E}e^{i(\tau,\zeta)}.$$

Отметим, что для многомерных характеристических функций справедливы аналоги доказанных выше теорем непрерывности и единственности. Обозначим через $\xi = (\xi_1, \xi_2, ..., \xi_m)$ вектор из независимых стандартных гауссовых величин, $\xi_i \sim \mathcal{N}(0,1)$.

Определение-теорема. Случайный вектор $\eta = (\eta_1, \eta_2, ..., \eta_n)$ называется *гауссовским с нулевым средним*, если выполнено одно из следующих эквивалентных условий.

1. $\eta = C\xi$, т.е. η_j являются линейными комбинациями случайных величин ξ_k, $j = 1, \ldots, n$, $k = 1, \ldots, m$:

$$\eta_j = \sum_{k=1}^{m} c_{jk}\xi_k,$$

где $C = (c_{jk})$ — некоторая вещественная матрица размера $n \times m$.

2. Характеристическая функция вектора η имеет вид

$$f_\eta(t) = \exp\Big(-\frac{1}{2}(Bt, t)\Big), \tag{9.3}$$

где симметрическая квадратная матрица B порядка n неотрицательно определена, $t = (t_1, t_2, ..., t_n) \in \mathbf{R}^n$.

3. Плотность p_η вектора η (т.е. плотность совместного распределения случайных величин η_i, $i = 1, \ldots, n$) имеет вид

$$p_\eta(x) = \Big((\sqrt{2\pi})^n\sqrt{\det B}\Big)^{-1}\exp\Big(-\frac{1}{2}(B^{-1}x, x)\Big), \tag{9.4}$$

где

$$x = (x_1, x_2, ..., x_n) \in \mathbf{R}^n,$$

причем симметрическая матрица $B = (b_{kj})$ порядка n есть матрица ковариаций случайных величин системы η : $b_{kj} = \mathsf{E}\eta_k\eta_j$ (и, значит, она является неотрицательно определенной).

Упражнение. Проверить, что

$$\int\limits_{\mathbf{R}^n} p_\eta(x)\,dx = 1.$$

Доказательство определения-теоремы.

1 $\Rightarrow$ **2**. Найдем характеристическую функцию для η. Заметим, что характеристическая функция системы ξ независимых стандартных гауссовских случайных величин равна $f_\xi(t) = \exp(-(t,t)/2)$. Тогда

$$f_\eta(t) = \mathsf{E}e^{i(t,C\xi)} = \mathsf{E}e^{i(C^*t,\xi)} = e^{-\frac{1}{2}(C^*t,C^*t)} = e^{-\frac{1}{2}(CC^*t,t)},$$

где C^* — транспонированная матрица. Заметим, что CC^* симметричная неотрицательно определенная матрица, таким образом, имеет место (9.3) с $B = CC^*$. При этом

$$\mathsf{E}\eta_k\eta_j = \sum_{l=1}^{n} c_{kl}c_{jl} = b_{kj}.$$

1 $\iff$ **3**. Для любого борелевского множества A из $\mathbf{R}^n$

$$\mathsf{P}(\eta \in A) = \mathsf{P}(C\xi \in A) = \mathsf{P}(\xi \in C^{-1}A) =$$

$$= \int\limits_{C^{-1}A} \frac{1}{(2\pi)^{n/2}} \exp\Big\{-\frac{1}{2}(x,x)\Big\}\,dx. \tag{9.5}$$

Сделаем замену переменных $y = Cx$. Получим, что правая часть (9.5) равна

$$\frac{1}{(2\pi)^{n/2}} \int\limits_A \exp\Big\{-\frac{1}{2}(C^{-1}y, C^{-1}y)\Big\}|\det C^{-1}|\,dy =$$

$$= \frac{1}{(2\pi)^{n/2}} \int\limits_A \exp\Big\{-\frac{1}{2}((CC^*)^{-1}y, y)\Big\}\frac{1}{\sqrt{\det CC^*}}\,dy,$$

и, таким образом, имеет место (9.4). Эту цепочку равенств можно пройти и в обратном направлении.

2 $\iff$ **3**. Покажем, что функция $f_\eta(t)$ получается Фурье-преобразованием $p_\eta(x)$. Рассмотрим

$$\int\limits_{\mathrm{R}^n} e^{i(t,x)} \frac{1}{\sqrt{\det B}(2\pi)^{n/2}} e^{-\frac{1}{2}(B^{-1}x,x)}\,dx.$$

Сделаем замену переменных $x = \mathcal{O}u$, $t = \mathcal{O}v$, где $\mathcal{O}$ — ортогональная матрица, приводящая B к диагональному виду: $\mathcal{O}^*B\mathcal{O} = D$ и $\det B = \det D = d_{11}d_{22}d_{33}...d_{nn}$, где $D = (d_{ij})$. Тогда $(B^{-1}x, x) =$

$(B^{-1}(\mathcal{O}u), \mathcal{O}u) = (D^{-1}u, u)$ и рассматриваемый интеграл преобразуется к виду :

$$\frac{1}{(2\pi)^{n/2}} \frac{1}{(d_{11}...d_{nn})^{1/2}} \int\limits_{\mathrm{R}^n} \exp\Big\{ i(v,u) - \frac{1}{2}(D^{-1}u, u) \Big\}\, du =$$

$$= \prod_{k=1}^{n} \frac{1}{\sqrt{2\pi d_{kk}}} \int\limits_{\mathrm{R}} \exp\Big\{ iv_k u_k - \frac{1}{2}\frac{u_k^2}{d_{kk}} \Big\}\, du_k =$$

$$= \prod_{k=1}^{n} \exp\Big\{ -\frac{v_k^2 d_{kk}}{2} \Big\} = \exp\Big\{ -\frac{1}{2}(Dv, v) \Big\} = \exp\Big\{ -\frac{1}{2}(Bt, t) \Big\}.$$

Так же проверяется, что обратным преобразованием Фурье от f_η является p_η.

Замечание 3. Гауссовский вектор η' с произвольным математическим ожиданием $m = (m_1, \ldots, m_n)$ может быть получен из гауссовского вектора η с нулевым средним сдвигом на вектор средних: $\eta' = \eta + m$. Для него используется обозначение $\eta' \sim \mathcal{N}(m, B)$. Плотность и характеристическую функцию вектора η' легко найти, зная плотность и характеристическую функцию для η.

Замечание 4. Вычисляя значение характеристической функции $f_\eta(t')$ в точке $t' = (0, ..., 1, 0, ..., 0)$, где единица стоит на k-м месте, получим, что характеристическая функция случайной величины η_k равна $\exp(-t_k^2 b_{kk}/2)$, т.е. $\eta \sim \mathcal{N}(0, b_{kk})$, где b_{kk} — диагональный элемент матрицы $B = (b_{ij})$.

Замечание 5. Элементы матрицы ковариаций B вектора η можно вычислить, применив $\frac{\partial^2}{\partial t_k \partial t_j}$ к обеим частям равенства

$$f_\eta(t) = \int e^{i(t,x)} p_\eta(x)\, dx$$

и положив $t = 0$ (обосновать законность дифференцирования под знаком интеграла).

Замечание 6. Если η_k не коррелирует с остальными компонентами вектора η (это значит, что ковариация этой случайной величины с любой другой компонентой есть нуль), то она независима от них. В этом легко убедиться, так как в этом случае характеристическая функция f_η вектора η расщепляется в произведение характеристической функции случайной величины η_k и характеристической функции вектора, полученного удалением η_k из

η. Матрица ковариаций этого вектора получается удалением из ковариационной матрицы B строки и столбца с номером k.

Литература

Ширяев А. Н. Вероятность, глава 3, §§3, 4, глава 2, §13. Феллер В. Введение в теорию вероятноятей и ее приложения, том 1, глава 3, §§4 – 6.

Лекция 10

Пуассоновская мера. Условное математическое ожидание

Рассмотрим n-мерный куб $\Lambda \subset \mathbf{R}^n$ с центром в начале координат, длина стороны которого равна N. Объем куба будем обозначать $|\Lambda|$: $\quad |\Lambda| = N^n$. Бросим наугад точку ("частицу") в куб Λ. Ее координаты имеют равномерное распределение в Λ (см. лекцию 7) с плотностью

$$p(x_1, \dots x_n) = \begin{cases} 1/|\Lambda|, & \text{если} \quad (x_1, \dots, x_n) \in \Lambda; \\ 0, & \text{если} \quad (x_1, \dots, x_n) \notin \Lambda. \end{cases} \tag{10.1}$$

Выберем в кубе борелевское подмножество $A \in \mathcal{B}(\Lambda^n)$. Тогда вероятность того, что частица попала в множество A, равна

$$\mathsf{P}\{\text{частица} \in A\} = \frac{|A|}{|\Lambda|}, \tag{10.2}$$

где $|A|$ — объем подмножества A. Бросим теперь в куб M частиц, каждую независимо от остальных. Тогда вероятность того, что все частицы попали в A, равна

$$\mathsf{P}\{\text{все частицы попали в}\, A\} = \left(\frac{|A|}{|\Lambda|}\right)^M . \tag{10.3}$$

Таким образом, мы имеем M независимых равномерно распределенных на Λ n-мерных случайных величин $\xi_1, \dots, \xi_M$, и плотность каждой из них задана в (10.1). При этом

$$\mathsf{P}\{\xi_1 \in A\} = \int\limits_A p(x_1, \dots, x_n)\, dx_1 \dots dx_n = \frac{|A|}{|\Lambda|} \,,$$

$$\mathsf{P}\{\xi_1 \in A, \dots, \xi_M \in A\} = \prod_{i=1}^M P\{\xi_i \in A\} = \left(\frac{|A|}{|\Lambda|}\right)^M .$$

Вероятность того, что в A попало ровно k частиц из данных M, равна

$$P_M^{(\Lambda)}(A, k) = P(A, k) = C_M^k \left(\frac{|A|}{|\Lambda|}\right)^k \left(1 - \frac{|A|}{|\Lambda|}\right)^{M-k}.$$

Вспомним, что это в точности есть вероятность наступления ровно k успехов в схеме Бернулли из M испытаний. Здесь испытание состоит в бросании частицы в Λ, а успех достигается при попадании в A.

Нам бы хотелось построить вероятностную меру, описывающую распределение бесконечного числа частиц в бесконечном пространстве. Заметим, что в пределе $|\Lambda| \to \infty$ введенные нами равномерно распределенные на Λ случайные величины уже не могут пониматься в обычном смысле в силу противоречия (в пределе) между требованиями

$$1/|\Lambda| \to 0 \quad \text{и} \quad \int\limits_{\Lambda} p(x_1, \ldots, x_n)\, dx_1 \ldots dx_n = 1.$$

Однако при некоторых условиях вероятности $P_M^{(\Lambda)}(A, k)$ имеют пределы.

Предложение 1. Рассмотрим произвольную последовательность кубов $\Lambda_0 \subset \Lambda_1 \subset \ldots \subset \Lambda_s \subset \ldots$, расширяющихся к $\mathbf{R}^n$. Стороны кубов равны, соответственно, $N_0 < N_1 < \ldots < N_s < \ldots$ Таким образом, $N_j \uparrow \infty$ при $s \to \infty$. Пусть число частиц M_s в кубе Λ_s увеличивается с ростом s так, что

$$\frac{M_s}{|\Lambda_s|} = \frac{M_s}{N_s^n} = \lambda > 0, \tag{10.4}$$

где λ некоторая константа. Тогда

$$\lim_{\substack{M\to\infty \\ |\Lambda|\to\infty}} P_M^{(\Lambda)}(A, k) = \lim_{s\to\infty} P_{M_s}^{(\Lambda_s)}(A, k) = \frac{(\lambda|A|)^k}{k!} e^{-\lambda|A|}. \tag{10.5}$$

Предельный переход по последовательности кубов, расширяющихся к $\mathbf{R}^n$, в статистической механике называется *термодинамическим* предельным переходом, а условие (10.4) есть условие постоянной плотности частиц.

Доказательство. Предложение 1 является переформулировкой теоремы Пуассона (см. лекцию 9). Число испытаний в рассматриваемой схеме Бернулли равно M, вероятность успеха равна $|A|/|\Lambda|$

и, по условию,

$$M\frac{|A|}{|\Lambda|} = \lambda|A|.$$

Предложение 1 можно доказать и непосредственно, вычисляя асимптотику функций $P_M^{(\Lambda)}(A,k)$. Например, при $k=0$

$$P_M^{(\Lambda)}(A,0) = \left(1-\frac{|A|}{|\Lambda|}\right)^{\lambda|\Lambda|} \longrightarrow e^{-\lambda|A|}.$$

При $k=1$

$$P_M^{(\Lambda)}(A,1) = M\frac{|A|}{|\Lambda|}\left(1-\frac{|A|}{|\Lambda|}\right)^{M-1} \longrightarrow \lambda|A|\cdot e^{-\lambda|A|}.$$

Упражнение 1. Продолжить доказательство предложения 1, вычислив асимптотику $P_M^{(\Lambda)}(A,k)$.

Как мы знаем, получающиеся в пределе числа

$$P(A,k) = \frac{(\lambda|A|)^k}{k!}e^{-\lambda|A|} \tag{10.6}$$

задают распределение Пуассона, $\sum\limits_k P(A,k)=1$, то есть мы можем воспринимать их как предельное распределение вероятностей.

Чтобы корректно задать вероятностную меру в "бесконечном объеме" (то есть меру, описывающую расположение бесконечного числа частиц во всем пространстве), воспользуемся аксиоматикой Колмогорова.

В качестве пространства элементарных исходов возьмем пространство Ω всех счетных подмножеств $\mathbf{R}^n$, являющихся *локально конечными*, т.е. пересечение такого подмножества с любым шаром конечного радиуса состоит из конечного числа точек (может быть, и пустого). Эти подмножества описывают расположение частиц в пространстве. В трехмерном случае они интерпретируются как молекулы идеального газа.

Пусть $\mathcal{A}=\{A_1,\dots,A_m\}$ — произвольный набор m попарно непересекающихся связных ограниченных борелевских множеств в $\mathbf{R}^n$. Пусть $\kappa=(k_1,\dots,k_m)$ — произвольный вектор натуральных чисел. Рассмотрим множество

$$D(\mathcal{A},\kappa) = \{\omega\in\Omega : \#(\omega\cap A_1)=k_1,\dots,\#(\omega\cap A_m)=k_m\},$$

состоящее из таких счетных множеств $\omega \in \Omega$, что для любого $i = 1, \ldots, m$ в множество A_i попало ровно k_i точек (символ $\#$ обозначает мощность множества). Множества $D(\mathcal{A}, \kappa)$ будем называть *базовыми событиями*, при этом *носителем* события $D(\mathcal{A}, \kappa)$ назовем множество $A = \cup_{i=1}^m A_i$. Определим вероятность события $D(\mathcal{A}, \kappa)$ следующим образом:

$$\mathsf{P}(D(\mathcal{A}, \kappa)) = \prod_{i=1}^{m} \frac{(\lambda |A_i|)^{k_i}}{k_i!} e^{-\lambda |A_i|}\,. \tag{10.7}$$

Пусть B — несвязное ограниченное борелевское множество. Определим вероятность того, что в B попало ровно k частиц, следующим образом:

$$\mathsf{P}\{D(B, k)\} = P(B, k) = \frac{(\lambda |B|)^k}{k!} e^{-\lambda |B|}\,. \tag{10.8}$$

Упражнение 2. Проверить, что определение (10.8) согласуется с (10.7).

Итак, вероятность того, что в заданное борелевское множество A попадет ровно k частиц, равна $P(A, k)$ (см. (10.6)), и такие события независимы, если их носители не пересекаются.

Построим алгебру $\widetilde{\Sigma}$, содержащую все базовые события $D(\mathcal{A}, \kappa)$ для всех $m \geq 1$, $\mathcal{A}$ и κ.

Определим меру P на алгебре $\tilde{\Sigma}$ как аддитивную функцию такую, что события с непересекающимися носителями независимы. А именно, пусть D_1 и D_2 — базовые события с непересекающимися носителями. Тогда

$$\mathsf{P}(D_1 \cup D_2) = \mathsf{P}(D_1) + \mathsf{P}(D_2) - \mathsf{P}(D_1 \cap D_2);$$

$$\mathsf{P}(D_1 \cap D_2) = \mathsf{P}(D_1) \cdot \mathsf{P}(D_2)$$

Предложение 2. Пусть D_1 и D_2 — базовые события с пересекающимися носителями. Тогда $D_1 \cap D_2$ представимо в виде объединения конечного числа базовых событий с непересекающимися носителями.

Доказательство. Проверим утверждение на примере событий с $m = 1$. Пусть $D_1 = D_1(A_1, k_1)$; $D_2 = D_2(A_2, k_2)$. Пусть $k_1 \leq k_2$. Обозначим $A_3 = A_1 \cap A_2 \neq \emptyset$. Тогда

$$D_1 \cap D_2 = \{\omega : \#(A_1 \cap \omega) = k_1,\ \#(A_2 \cap \omega) = k_2\} =$$

$$= \bigcup_{k_3=0}^{k_1} \{\omega : \#((A_1 \setminus A_3) \cap \omega) = k_1 - k_3,$$

$$\#((A_2 \setminus A_3) \cap \omega) = k_2 - k_3, \#(A_3 \cap \omega) = k_3\}.$$

Случай $m > 1$ полностью аналогичен.

Предложение 3. Мера P является счетно-аддитивной на $\tilde{\Sigma}$.

Упражнение 3. Доказать предложение 3.

Построим теперь наименьшую σ-алгебру Σ, содержащую $\tilde{\Sigma}$. По теореме о продолжении меры мера P продолжается однозначно на σ-алгебру Σ. Таким образом, мы построили вероятностное пространство $(\Omega, \Sigma, \mathsf{P})$, описывающее распределение бесконечного числа частиц в бесконечном пространстве. В размерности $n > 1$ это вероятностное пространство называется идеальным газом, или *пуассоновским случайным полем.*

Замечание 1. Пусть $B_s \in \mathcal{B}(\mathbf{R}^n)$ — последовательность борелевских множеств такая, что $B_s \uparrow B \in \mathcal{B}(\mathbf{R}^n)$ (или $B_s \downarrow B \in \mathcal{B}(\mathbf{R}^n)$) при $s \to \infty$. Тогда $|B_s| \to |B|$ и, очевидно, для любого $0 \leq k \leq \infty$

$$P(B, k) = \lim_{s \to \infty} P(B_s, k).$$

В частности, из (10.6) видно, что вероятность того, что в заданной точке окажется частица, равна нулю. Аналогично, для множества $B \in \mathcal{B}(\mathbf{R}^n)$ с $|B| = \infty$ вероятность $P(B, k)$ того, что в B попало ровно k частиц ($k < \infty$), равна нулю. Заметим, кроме того, что при малых $\lambda|B|$

$$P(B, 1) \sim \lambda|B|\,, \quad \mathsf{P}\Big(\bigcup_{j \geq 2} D(B, j)\Big) = o(\lambda|B|),$$

$$P(B, 0) = 1 - \lambda|B| + o(\lambda|B|).$$

В размерности $n = 1$ пуассоновское случайное поле называют *пуассоновским потоком.* При этом числовая ось $\mathbf{R}$ интерпретируется как ось времени, а нахождение частицы в точке $t \in \mathbf{R}$ как наступление в момент t некоторого события, повторяющегося через случайные промежутки времени. Нам интересно узнать вероятностные характеристики этих случайных промежутков времени.

Определение 1. Случайная величина имеет *экспоненциальное распределение*, если ее функция распределения F и плотность p равны соответственно

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{при } x \geq 0, \\ 0 & \text{при } x < 0; \end{cases} \qquad p(x) = \begin{cases} \lambda e^{-\lambda x} & \text{при } x \geq 0, \\ 0, & \text{при } x < 0; \end{cases}$$

где $\lambda > 0$.

Лемма 1. Расстояние между соседними частицами пуассоновского потока распределено экспоненциально.

Пусть в точке x находится частица. Выберем произвольное $y > x$. Тогда соседняя с нею справа частица находится правее, чем y (т.е. отстоит от данной частицы не менее, чем на $y - x$), если в полуинтервале $(x, y]$ нет частиц. Таким образом, чтобы доказать лемму 1, нужно показать, что условная вероятность

$$\mathsf{P}\,(\text{ в } (x,y] \text{ нет частиц} \mid \text{в } x \text{ есть частица}) =$$

$$= \mathsf{P}\{D((x,y],0) \mid D(x,1)\} = \exp\{-\lambda(y-x)\}. \tag{10.9}$$

Приведем интуитивное доказательство леммы 1, опирающееся на замечание 1. Выберем в интервале (x, y) точку $z \in (x, y)$. Устремим z к x и запишем условную вероятность (10.9) как предел:

$$\mathsf{P}\{D((x,y],0) \mid D(x,1)\} = \lim_{z\downarrow x} \mathsf{P}\{D((z,y],0) \mid D([x,z],1)\}.$$

Поскольку носители событий $D((z,y],0)$ и $D([x,z],1)$ не пересекаются, эти события независимы. Имеем:

$$\mathsf{P}\{D((z,y],0) \mid D([x,z],1)\} = \frac{\mathsf{P}\{D((z,y],0) \cap D([x,z],1)\}}{\mathsf{P}\{D([x,z],1)\}} =$$

$$= \frac{\mathsf{P}\{D((z,y],0)\mathsf{P}D([x,z],1)\}}{\mathsf{P}\{D([x,z],1)\}} = \mathsf{P}\{D((z,y],0)\}.$$

Поэтому

$$\mathsf{P}\{D((x,y],0) \mid D(x,1)\} = \lim_{z\downarrow x} \mathsf{P}\{D((z,y],0)\} = \lim_{z\downarrow x} e^{-\lambda(y-z)} = e^{\lambda(y-x)}.$$

Чтобы получить строгое доказательство леммы 1, нам нужно научиться работать с условными вероятностями при условии события, имеющего нулевую вероятность. До сих пор мы имели дело лишь с условными вероятностями при условии события, имеющего *положительную* вероятность. Для того, чтобы определить условную вероятность при условии события, которое имеет *нулевую* вероятность, нам понадобится понятие *условного математического ожидания*.

Условное математическое ожидание

В лекции 5 мы ввели понятие условного математического ожидания одной дискретной случайной величины относительно другой. Напомним его.

Пусть ξ и η — случайные величины на пространстве Ω, принимающие конечное число значений. Пусть $\{x_1, x_2, ..., x_n\}$ есть множество значений ξ, а $\{y_1, y_2, ..., y_m\}$ есть множество значений η. Случайная величина ξ индуцирует разбиение $\mathcal{D}_\xi$ пространства элементарных событий Ω на n непересекающихся подмножеств:

$$\Omega = \bigcup_{i=1}^{n} \{\omega : \xi(\omega) = x_i\},$$

$$\{\omega : \xi(\omega) = x_i\} \cap \{\omega : \xi(\omega) = x_j\} = \emptyset \text{ при } i \neq j.$$

Аналогично, случайная величина η индуцирует разбиение $\mathcal{D}_\eta$. Заметим, что множество $\{\omega : \xi(\omega) = x_i\}$, вообще говоря, может не быть элементом алгебры, порожденной разбиением $\mathcal{D}_\eta$ (т.е. минимальной алгебры, содержащей все элементы $\mathcal{D}_\eta$). Обозначим $\mathcal{E}_\eta$ σ-алгебру, порожденную элементами разбиения $\mathcal{D}_\eta$, т.е. минимальную σ-алгебру, содержащую множества $\{\omega : \eta(\omega) = y_i\}$, $i = 1, \ldots, m$. Поскольку система $\mathcal{D}_\eta$ конечна, то порожденная ею алгебра совпадает с $\mathcal{E}_\eta$. Таким образом, мы можем сказать, что случайная величина ξ не обязана быть измеримой относительно сигма-алгебры $\mathcal{E}_\eta$. Очевидно, функции $\xi(\omega)$ и $\eta(\omega)$ являются постоянными на элементах разбиения $\mathcal{D}_\xi$ и $\mathcal{D}_\eta$ соответственно.

Зафиксируем значение y_j случайной величины η.

Определение 2. *Условным математическим ожиданием случайной величины* ξ *при условии* $\eta = y_j$ называется число

$$\mathsf{E}(\xi \mid \{\eta = y_j\}) = \sum_{i=1}^{n} x_i \mathsf{P}(\xi = x_i | \eta = y_j).$$

Если условие $\eta = y_j$ не фиксировать, а допустить, что оно может меняться, получим функцию от y_j:

$$\mathsf{E}(\xi|\eta) = \sum_{j=1}^{m} \sum_{i=1}^{n} x_i \mathsf{P}(\xi = x_i | \eta = y_j) I_{\{\eta = y_j\}}(\omega), \tag{10.10}$$

где I_A есть индикатор множества A :

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

Функция $\mathsf{E}(\xi|\eta)$, очевидно, является случайной величиной, постоянной на каждом элементе разбиения $\mathcal{D}_\eta$, а значит, измеримой относительно $\mathcal{E}_\eta$.

Определение 3. *Условным математическим ожиданием* ξ *относительно* η называется случайная величина $\mathsf{E}(\xi|\eta)$, определенная в (10.10).

Выражение (10.10) обозначают также $\mathsf{E}(\xi|\mathcal{E}_\eta)$ и называют условным математическим ожиданием ξ относительно σ-алгебры $\mathcal{E}_\eta$.

Таким образом, $\mathsf{E}(\xi|\eta)$ есть усреднение ξ на каждом множестве уровня $\{\omega : \eta(\omega) = y_j\}$ функции η. Это сразу влечет равенство

$$\int\limits_M \xi\, d\mathsf{P} = \int\limits_M \mathsf{E}(\xi|\eta)\, d\mathsf{P} \qquad \text{для любого } M \in \mathcal{E}_\eta \tag{10.11}$$

(см. лекцию 5). В нашем случае интегралы в (10.11) есть в действительности просто конечные суммы. Но мы записываем это свойство в общем виде, так как оно имеет место и для произвольных случайных величин. В частности,

$$\mathsf{E}\xi = \int\limits_\Omega \mathsf{E}(\xi|\eta)\, d\mathsf{P} = \mathsf{E}\,\mathsf{E}(\xi|\eta) \tag{10.12}$$

(обобщение формулы полной вероятности).

Перейдем теперь к построению условного математического ожидания в общем случае. Для этого требуется дополнительная техника.

Абсолютная непрерывность мер

Теорема-определение (Теорема Радона — Никодима).

Пусть на измеримом пространстве (Ω, Σ), имеются две меры (не обязательно вероятностные), μ и ν. Мера μ называется *абсолютно непрерывной* относительно ν, если выполнено любое из следующих двух эквивалентных условий.

1. Для любого $A \in \Sigma$ если $\mu(A) > 0$, то $\nu(A) > 0$.
2. Существует измеримая функция $f(\omega)$ на Ω, такая, что

$$\mu(A) = \int\limits_A f(\omega)\, d\nu$$

для любого $A \in \Sigma$. Эта функция называется *плотностью меры μ по мере ν*, или *производной Радона — Никодима*, и обозначается $f(\omega) = \frac{d\mu}{d\nu}$. Она единственна с точностью до значений на множествах μ-меры 0 (или, что эквивалентно, ν-меры 0).

Доказательство. То, что из 2 следует 1, очевидно. Доказательство обратного утверждения можно найти в книге А. Н. Колмогорова и С. В. Фомина.

Пусть задано вероятностное пространство (Ω, Σ, μ) и две случайные величины ξ и η на нем. Попробуем повторить определение условного математического ожидания, данное для дискретного случая, в общем случае. Каждой случайной величине η сответствует разбиение пространства Ω на подмножества $A_x, x \in \mathbf{R}$, уровня x : $A_x = \{\omega : \eta(\omega) = x\}$. Однако во многих случаях $\mu(A_x) = 0$. Так, например, если случайная величина η имеет непрерывное распределение, $\mu(A_x) = 0$ для любого x. Мы не умеем пока определять условные вероятности относительно событий, имеющих нулевую вероятность. Рассмотрим σ-алгебру $\mathcal{E}_\eta$, порожденную случайной величиной η. Ее элементами являются события вида $\{\omega : \eta(\omega) \in B\}$, $B \in \mathcal{B}(\mathbf{R})$.

Предположим, что $\xi(\omega) \geq 0$ для почти всех ω. Рассмотрим измеримое пространство $(\Omega, \mathcal{E}_\eta)$. На нем есть две меры : мера μ и мера ν, где $\nu(A) = \int\limits_A \xi(\omega)\, d\mu$. Очевидно, ν абсолютно непрерывна относительно μ. Поэтому по теореме Радона — Никодима существует плотность $u(\omega)$, *измеримая относительно $\mathcal{E}_\eta$*, такая что $\nu(A) = \int\limits_A u(\omega)\, d\mu$. Положим по определению $\mathsf{E}(\xi|\eta) = u(\omega)$. Фактически функция $\mathsf{E}(\xi|\eta)$ является функцией только от $y = \eta(\omega)$. Это результат следующей теоремы, которую мы приводим без доказательства.

Теорема 1. Пусть (Ω, Σ, μ) — вероятностное пространство, на котором заданы случайные величины ζ и η. Пусть случайная величина ζ измерима относительно σ-алгебры $\mathcal{E}_\eta$, порожденной случайной величиной η. Тогда найдется борелевская функция φ такая, что ζ есть композиция функций φ и η, то есть $\zeta(\omega) = \varphi(\eta(\omega))$ для каждого ω.

Учитывая теорему 1, мы можем определить условное математическое ожидание случайной величины ξ при условии, что $\eta = y$:

$$\mathsf{E}(\xi|\{\eta = y\}) = \varphi(y). \tag{10.13}$$

Заметим, что если $\mathsf{E}\xi = \infty$, то случайная величина $\mathsf{E}(\xi|\eta)$, вообще говоря, является расширенной.

Дадим теперь формальное определение условного математического ожидания.

Определение 4. *Условным математическим ожиданием* неотрицательной случайной величины ξ относительно случайной величины η называется расширенная случайная величина $\mathsf{E}(\xi|\eta)$, такая что $\mathsf{E}(\xi|\eta)$ является $\mathcal{E}_\eta$-измеримой и для любого $A \in \mathcal{E}_\eta$

$$\int\limits_A \xi\, d\mu = \int\limits_A \mathsf{E}(\xi|\eta)\, d\mu. \tag{10.14}$$

Для произвольной случайной величины ξ ее условное математическое ожидание считается определенным, если

$$\min(\mathsf{E}(\xi^+|\eta), \mathsf{E}(\xi^-|\eta) < \infty.$$

В этом случае

$$\mathsf{E}(\xi|\eta) = \mathsf{E}(\xi^+|\eta) - \mathsf{E}(\xi^-|\eta).$$

Замечание 2. Существование условного математического ожидания следует из теоремы Радона — Никодима. Оно определено единственным образом с точностью до множеств μ-меры 0.

Свойства условных математических ожиданий.

1. Если a, b, c - константы, то

$$\mathsf{E}(a\xi_1 + b\xi_2 + c|\eta) = a\mathsf{E}(\xi_1|\eta) + b\mathsf{E}(\xi_2|\eta) + c \qquad \text{(п. н.)}.$$

2. Если $\mu\{\omega : \xi_1(\omega) \le \xi_2(\omega)\} = 1$, то $\mathsf{E}(\xi_1|\eta) \le \mathsf{E}(\xi_2|\eta)$ (п.н.).
3. $|\mathsf{E}(\xi|\eta)| \le \mathsf{E}(|\xi|\ |\eta)$ (п.н.).
4. $\mathsf{E}\,\mathsf{E}(\xi|\eta) = \mathsf{E}\xi$.
5. Если $\mathcal{E}_{\eta_1} \subseteq \mathcal{E}_{\eta_2}$, то есть σ-алгебра $\mathcal{E}_{\eta_2}$ является измельчением σ-алгебры $\mathcal{E}_{\eta_1}$, то

$$\mathsf{E}(\mathsf{E}(\xi|\eta_2)|\eta_1) = \mathsf{E}(\xi|\eta_1) \qquad \text{(п.н.)}.$$

6. Если $\mathcal{E}_{\eta_1} \subseteq \mathcal{E}_{\eta_2}$, то

$$\mathsf{E}(\mathsf{E}(\xi|\eta_1)|\eta_2) = \mathsf{E}(\xi|\eta_1) \qquad \text{(п.н.)}.$$

7. Если случайная величина ξ измерима относительно $\mathcal{E}_\eta$, то

$$\mathsf{E}(\xi|\eta) = \xi \qquad \text{(п.н.)}.$$

8. Если случайные величины ξ и η независимы, то

$$\mathsf{E}(\xi|\eta) = \mathsf{E}\xi \qquad (\text{п.н.}).$$

9. Если ξ_1 $\mathcal{E}_\eta$-измерима и $\mathsf{E}|\xi_1| < \infty$, $\mathsf{E}|\xi_1\xi| < \infty$, то

$$\mathsf{E}(\xi_1\xi|\eta) = \xi_1\mathsf{E}(\xi|\eta) \qquad (\text{п.н.}).$$

Упражнение 4. Пользуясь определением и свойствами интеграла, доказать свойства 1, 3, 4, 6, 7.

Доказательства остальных свойств можно прочитать в учебнике Ширяева, глава 2, §7.

Можно рассматривать σ-алгебру $\Sigma_1 \subset \Sigma$, не связанную априори с какой-либо случайной величиной, и определить условное математическое ожидание ξ относительно Σ_1.

Определение 5. Пусть $(\Omega, \Sigma, \mathsf{P})$ — вероятностное пространство, на котором задана случайная величина ξ и пусть $\Sigma_1 \subset \Sigma$ есть σ-подалгебра σ-алгебры Σ. *Условным математическим ожиданием случайной величины* ξ *относительно* σ*-алгебры* Σ_1 называется расширенная случайная величина $\mathsf{E}(\xi|\Sigma_1)$, такая что $\mathsf{E}(\xi|\Sigma_1)$ является Σ_1-измеримой и для любого $A \in \Sigma_1$

$$\int\limits_A \xi\, d\mathsf{P} = \int\limits_A \mathsf{E}(\xi|\Sigma_1)\, d\mathsf{P}.$$

Мы видим из этого определения, что в частном случае, когда σ-алгебра Σ_1 порождена случайной величиной, то есть $\Sigma_1 = \mathcal{E}_\eta$, $\mathsf{E}(\xi|\mathcal{E}_\eta) = \mathsf{E}(\xi|\eta)$.

Пример. Условное математическое ожидание для непрерывно распределенных случайных величин. Пусть случайные величины ξ и η, заданные на вероятностном пространстве $(\Omega, \Sigma, \mathsf{P})$, имеют совместную плотность $p(x, y)$. Предположим, что $\mathsf{E}\xi < \infty, \mathsf{E}\eta < \infty$. Тогда плотность p_η равна $p_\eta(y) = \int\limits_R p(x, y)\, dx$. Обозначим

$$p_{\xi|\eta}(x|y) = \frac{p(x, y)}{p_\eta(y)}, \tag{10.15}$$

если $p_\eta(y) \neq 0$. Для тех y, где $p_\eta(y) = 0$, полагаем выражение равным нулю. Это выражение называется *условной плотностью случайной величины* ξ *относительно* η. С ее помощью условное математическое ожидание ξ относительно η можно записать явно

$$\mathsf{E}(\xi|\eta)(y) = \int\limits_{-\infty}^{\infty} x\, p_{\xi|\eta}(x|y)\, dx \tag{10.16}$$

как борелевскую функцию от значений y, принимаемых случайной величиной η.

Проверим (10.14), дважды воспользовавшись заменой переменных в интеграле Лебега (см. лекцию 7):

$$\int\limits_{\{\omega:\,\eta(\omega)\in B\}} \mathsf{E}(\xi|\eta)(\omega)\,d\mathsf{P}(\omega) = \int\limits_B\int\limits_{\mathrm{R}} x\,p_{\xi|\eta}(x|y)\,dx\,p_\eta(y)\,dy =$$

$$= \int\limits_B\int\limits_{\mathrm{R}} x\,\frac{p(x,y)}{p_\eta(y)}\,p_\eta(y)\,dx\,dy = \int\limits_B\int\limits_{\mathrm{R}} x\,p(x,y)\,dx\,dy =$$

$$= \int\limits_\Omega \xi\,I_{\{\eta\in B\}}\,d\mathsf{P} = \int\limits_{\{\omega:\,\eta(\omega)\in B\}} \xi\,d\mathsf{P}\,.$$

Поясним теперь, почему $p_{\xi|\eta}(x|y)$ называется условной плотностью. Во-первых, не вдаваясь в строгое обоснование предельного перехода, заметим, что $p_{\xi|\eta}(x|y)$ можно получить как предел при $\varepsilon\to 0$ условной вероятности

$$\mathsf{P}(\xi\in\mathcal{O}_\varepsilon(x)|\eta\in\mathcal{O}_\varepsilon(y)) = \frac{\int\limits_{\mathcal{O}_\varepsilon(x)}\int\limits_{\mathcal{O}_\varepsilon(y)} p(s,z)\,dz\,ds}{\int\limits_{\mathcal{O}_\varepsilon(y)} p_\eta(z)\,dz}\,,$$

где $\mathcal{O}_\varepsilon(x)$ и $\mathcal{O}_\varepsilon(y)$ — интервалы длины ε с центрами x и y соответственно. Само же условное математическое ожидание исходя из (10.15) и (10.16) можно представить как предел

$$\mathsf{E}(\xi|\eta=y) = \lim_{\varepsilon\to 0}\ \frac{\int\limits_R\int\limits_{\mathcal{O}_\varepsilon(y)} s\,p(s,z)\,dz\,ds}{\int\limits_{\mathcal{O}_\varepsilon(y)} p_\eta(z)\,dz}\,.$$

Определим теперь условную вероятность при условии события, вероятность которого может быть равной нулю.

Определение 6. Пусть $A\in\Sigma$. *Условной вероятностью события A при условии $\eta=y$* называется величина

$$\mathsf{P}(A|\{\eta=y\}) = \mathsf{E}(I_A|\{\eta=y\}),$$

где I_A — индикатор события A.

Упражнение 5. Проверить, что для дискретной случайной величины η определение 6 превращается в известное нам определение обычной условной вероятности.

Предложение 4. Пусть случайные величины ξ и η имеют совместную плотность $p(x, y)$. Тогда

$$P(\xi \in C|\eta = y) = \int\limits_C p_{\xi|\eta}(x|y)\,dx\,, \tag{10.17}$$

то есть $p_{\xi|\eta}(x|y)$, определенная в (10.15), есть не что иное, как плотность условного распределения ξ при условии $\eta = y$.

Мы не приводим формального доказательства этого предложения, опирающегося на определение 6. Заметим только, что другое доказательство может быть получено с помощью предельного перехода в тех случаях, когда он поддается обоснованию. Действительно, взяв фиксированное множество $C \subset \mathbf{R}$ и переходя к пределу $\varepsilon \to 0$ в равенстве

$$\mathsf{P}(\xi \in C|\eta \in \mathcal{O}_\varepsilon(y)) = \frac{\int\limits_C \int\limits_{\mathcal{O}_\varepsilon(y)} p(s, z)\,dz\,ds}{\int\limits_{\mathcal{O}_\varepsilon(y)} p_\eta(z)\,dz}\,,$$

получим (10.17).

Теперь мы можем вернуться к доказательству леммы 1.

Доказательство леммы 1. Рассмотрим случайную величину I_x, являющуюся индикатором события $D(\{x\}, 1)$, состоящего в том, что в точке находится частица:

$$I_x = \begin{cases} 1, & \text{если в } x \text{ есть частица;} \\ 0, & \text{в противном случае.} \end{cases}$$

Тогда

$$\mathsf{P}\,(\text{ в } (x, y] \text{ нет частиц } \mid \text{в } x \text{ есть частица}) =$$
$$= \mathsf{P}\{D((x, y], 0) \mid I_x = 1\} = \mathsf{E}(I_{D((x,y],0)} \mid I_x = 1).$$

Поскольку носители событий $D((x, y], 0)$ и $D(\{x\}, 1)$ не пересекаются, случайные величины $I_{D((x,y],0)}$ и I_x независимы. Поэтому

$$\mathsf{E}(I_{D((x,y],0)} \mid I_x = 1) = \mathsf{E}(I_{D((x,y],0)}) = P((x, y], 0) = e^{-\lambda(y-x)}.$$

Литература

Колмогоров А. Н., Фомин С. В. Элементы теории функций и функционального анализа, глава 6. Ширяев А. Н. Вероятность, глава 2, § 7.

Лекция 11

Задача перколяции на Z^2

В этой лекции мы рассмотрим задачу перколяции на квадратной решетке. Теория перколяции (протекания) возникла первоначально в физике твердого тела, но в последние годы находит все более широкие применения в самых различных естественных науках. Основной объект этой теории — случайные однородные множества на графах, решетках, группах, евклидовых пространствах. Изучаются глобальные свойства таких множеств (проблема связности, статистика ограниченных компонент и т.п.). Здесь мы рассмотрим лишь одну задачу — перколяционную задачу узлов на квадратной решетке.

Обозначим множество узлов квадратной решетки $\mathbf{Z}^2$ через V, а множество ребер через E. Пара $v_1, v_2 \in \mathbf{Z}^2$ есть ребро, $(v_1, v_2) \in E$, только если $||v_1 - v_2|| = 1$. На множестве узлов (вершин) V рассмотрим независимое бернуллиевское случайное поле. Это значит, что каждой вершине $v \in V$ приписана случайная величина, принимающая значения $+1, -1$ и все эти случайные величины независимы. Пространством элементарных исходов будет $\Omega = \prod_{v \in V} \{+1, -1\}$. Произвольную точку из Ω мы будем называть конфигурацией и обозначать через $\omega = \{\omega(v)\}_{v \in V}$. В качестве σ-алгебры Σ будем рассматривать σ-алгебру, порожденную цилиндрическими множествами в Ω :

$$\{\omega : \omega(v_1) = e_1, \omega(v_2) = e_2, \dots, \omega(v_n) = e_n\}, \ v_i \in V, \ e_i = \pm 1.$$

Вероятностная мера на (Ω, Σ) задается произведением мер

$$\mathsf{P} = \mathsf{P}_p = \prod_{v \in V} \mu_v ,$$

где μ_v — распределение случайной величины, находящейся в вершине v:

$$\mu_v(\omega(v) = +1) = p, \quad \mu_v(\omega(v) = -1) = 1 - p = q.$$

Упражнение 1. Повторив рассуждения, проведенные при построении схемы Бернулли (см. лекцию 2), убедиться, что мера P определена корректно.

Определение 1. *Связным путем* длины n на решетке $\mathbf{Z}^2$ называется последовательность вершин $(v_1, v_2, \ldots, v_{n+1}), v_i \in V$, такая, что для любого $i = 1, \ldots, n$ пара (v_i, v_{i+1}) образует ребро: $(v_i, v_{i+1}) \in E$. Путь называется *самонепересекающимся*, если все v_i различны.

Если $\omega(v) = +1$, то вершину v назовем *занятой*, если же $\omega(v) = -1$, то *свободной*.

Определение 2. *Кластером* $W(v)$ занятой вершины v назовем множество занятых узлов $v' \in V$, достижимых из v по связным путям $(v, v_1, v_2, \ldots, v_k, v')$, состоящим только из занятых узлов.

Количество узлов в кластере $W(v)$ будем обозначать $|W(v)|$. Возможен случай, когда $|W(v)| = \infty$. Кластер — это случайный объект.

Один из главных вопросов в теории перколяции — распределение размера кластера $|W(v)|$. Особый интерес представляет *вероятность просачивания на бесконечность* (или перколяции):

$$\theta_p(v) = \mathsf{P}_p\{|W(v)| = \infty\}.$$

Ясно, что если $p = 0$, то $\theta_0(v) = 0$. Если же $p = 1$, то $\theta_1(v) = 1$. Если $0 < p < 1$, то $0 \le \theta_p(v) < 1$.

Определим критическую вероятность p_c:

$$p_c = \sup\{p \in [0,1] \,:\, \theta_p(v) = 0\}.$$

Поскольку меры μ_v одинаковы для всех v, то критическая вероятность p_c не зависит от v. Основной целью настоящей лекции является доказательство следующего утверждения.

Теорема 1. Для перколяционной задачи узлов на квадратной решетке

$$1/3 \le p_c \le 6/7,$$

то есть при $p < 1/3$ для любой вершины v кластер $W(v)$ конечен с вероятностью 1, а при $p > 6/7$ вершина v принадлежит бесконечному кластеру из занятых вершин с положительной вероятностью.

Доказательство. Разобьем доказательство этой теоремы на два этапа. Сначала мы строго докажем, что $p_c \ge 1/3$. Далее мы покажем, что $p_c \le 6/7$. В этой части наши рассуждения будут опираться на геометрические леммы. Интуитивно их понять нетрудно, однако их строгое доказательство, изложенное, например, в книге Кестена, представляет существенные трудности (не вероятностные), и мы его здесь проводить не будем.

Докажем, что $p_c \geq 1/3$. Рассмотрим событие $A_n = A_n(v)$, состоящее в том, что существует связный путь без самопересечений длины n, состоящий из занятых вершин и исходящий из v. Обозначим через b_n число всевозможных связных самонепересекающихся путей длины n, исходящих из v.

Лемма 1. Имеет место следующая оценка:

$$b_n \leq 4 \cdot 3^{n-1}. \tag{11.1}$$

Упражнение 2. Доказать лемму 1.

Пусть $(v, v_1, v_2, \ldots, v_{n-1}, v_n)$ есть путь без самопересечений длины n. Вероятность того, что все его вершины заняты, равна p^{n+1}. Учитывая лемму 1, получаем

$$\mathsf{P}(A_n) \leq p^{n+1} b_n \leq \frac{4}{9}(3p)^{n+1}. \tag{11.2}$$

Заметим, что семейство событий $\{A_n\}$ вложенное, то есть $A_{i+1} \subseteq A_i$ для любого i. Поскольку

$$\{\omega : \ |W(v)| = \infty\} = \bigcap_{n=1}^{\infty} A_n(v),$$

то

$$\mathsf{P}_p\{\omega : \ |W(v)| = \infty\} = \lim_{n\to\infty} \mathsf{P}_p(A_n(v)) \leq \lim_{n\to\infty} \frac{4}{9}(3p)^{n+1}. \tag{11.3}$$

Следовательно, $\mathsf{P}_p\{\omega : \ |W(v)| = \infty\} = 0$ при $p < 1/3$. Тем самым мы доказали, что $p_c \geq 1/3$.

Для доказательства оценки сверху рассмотрим перколяционную задачу на другом графе. Множество вершин (узлов) будет то же, $\mathbf{Z}^2$. Ребра же будут связывать все пары вершин, расстояние между которыми меньше или равно $\sqrt{2}$. Мы будем называть эту решетку *квадратной с двумя диагоналями* (основную решетку, для которой доказывается теорема 1, мы называем квадратной). Критическую вероятность для этой решетки мы обозначим $\tilde{p}_c$, она определяется аналогично. Повторяя наши рассуждения для квадратной решетки с двумя диагоналями, мы можем доказать, что $\tilde{p}_c \geq 1/7$.

Рассмотрим конкретное элементарное событие ω, то есть зафиксируем конфигурацию из $+1$ и -1, приписанных узлам решетки $\mathbf{Z}^2$.

Лемма 2 (геометрическая). Кластер $W(v)$ из занятых вершин, содержащий v, конечен тогда и только тогда, когда существует замкнутый путь из свободных вершин, объемлющий вершину v и такой, что любые две последовательные вершины этого пути являются соседними на квадратной решетке с двумя диагоналями. (Такой путь мы называем связным в смысле квадратной решетки с двумя диагоналями.)

Лемма 3 (геометрическая). Если существует лишь конечное число контуров (т.е. замкнутых путей), состоящих из свободных вершин, объемлющих вершину v и связных в смысле квадратной решетки с двумя диагоналями, то существует бесконечный кластер из занятых вершин (связный в смысле квадратной решетки).

Замечание. Кластер в лемме 3 не обязательно содержит вершину v.

Доказательство этих лемм можно прочитать в книге Х. Кестена "Теория просачивания для математиков", М., Мир, 1986. Мы его опускаем.

Лемма 4. Если $p > 6/7$, то с вероятностью 1 существует лишь конечное число замкнутых контуров из свободных вершин, объемлющих v и связных в смысле квадратной решетки с двумя диагоналями.

Доказательство. Пусть C_n — событие, состоящее в том, что существует замкнутый контур длины n (без самопересечений) из свободных вершин, связный в смысле квадратной решетки с двумя диагоналями и объемлющий v. Тогда

$$\mathsf{P}_p(C_n) \le (1-p)^{n+1} R_n,$$

где R_n — число таких контуров. Из комбинаторных соображений легко получаем, что

$$R_n \le n \cdot 8 \cdot 7^{n-1}.$$

Таким образом,

$$\mathsf{P}_p(C_n) \le \frac{8}{49} n (7(1-p))^{n+1}.$$

Отсюда следует, что при $p > 6/7$ сходится ряд из вероятностей

$$\sum_{n=1}^{\infty} \mathsf{P}(C_n) < \infty.$$

Теперь нам понадобится одно важное утверждение. Пусть имеется некоторая последовательность событий $C_1, C_2, \ldots$ заданных на одном вероятностном пространстве. Определим событие $\{C_n$ б.ч.$\}$, состоящее в том, что наступит бесконечно много событий из последовательности C_n:

$$\{C_n \text{ б.ч.}\} = \bigcap_{n=1}^{\infty} \bigcup_{k \geq n} C_k.$$

Лемма 5 (Лемма Бореля — Кантелли).

Если $\sum \mathsf{P}(C_n) < \infty$, то $\mathsf{P}\{C_n \text{ б.ч.}\} = 0$.

Если $\sum \mathsf{P}(C_n) = \infty$ и события $C_1, C_2, \ldots$ независимы, то $\mathsf{P}\{C_n \text{ б.ч.}\} = 1$.

Мы докажем лемму Бореля — Кантелли после завершения доказательства теоремы.

Применяя лемму Бореля — Кантелли, получаем, что с вероятностью 1 происходит лишь конечное число событий C_n, следовательно, с вероятностью 1 существует лишь конечное число замкнутых самонепересекающихся контуров, охватывающих v. Лемма 4 доказана.

Из лемм 3 и 4 следует, что при $p > 6/7$ с вероятностью 1 существует бесконечный кластер из занятых вершин. Напомним, что по определению кластер является множеством, связным в смысле квадратной решетки.

Для завершения доказательства теоремы нам осталось доказать следующее утверждение.

Лемма 6. Если $\mathsf{P}_p\{|W(v)| = \infty\} = 0$ для любой v, то вероятность того, что существует бесконечный кластер из занятых вершин, равна 0.

Доказательство. Действительно, событие, состоящее в том, что существует бесконечный кластер из занятых вершин, можно представить в виде объединения событий:

$$\{\omega : \exists \text{ бесконечный кластер }\} = \bigcup_{v \in V} \{\omega : |W(v)| = \infty\}.$$

Вероятность каждого события, стоящего в объединении, равна 0, следовательно, и вероятность объединения равна нулю:

$$\mathsf{P}_p\{\text{существует бесконечный кластер}\} = 0.$$

Завершаем доказательство теоремы 1. Итак, мы показали, что при $p > 6/7$

$$\mathsf{P}_p\{\omega : \text{существует бесконечный кластер}\} = 1.$$

Из леммы 6 следует, что существует v, такая что

$$\mathsf{P}_p\{|W(v)| = \infty\} > 0.$$

Но критическая вероятность не зависит от v в силу пространственной однородности рассматриваемого бернуллиевского поля. Тем самым доказана верхняя оценка для критической вероятности теоремы 1:

$$p_c \le 6/7.$$

Доказательство леммы Бореля – Кантелли. События $\bigcup_{k\ge n} C_k$ вложенные, поэтому в силу непрерывности вероятностной меры

$$\mathsf{P}\{C_n \text{ б.ч.}\} = \mathsf{P}\{\bigcap_{n=1}^{\infty} \bigcup_{k\ge n} C_k\} = \lim \mathsf{P}(\bigcup_{k\ge n} C_k) \le$$

$$\le \lim \sum_{k\ge n} \mathsf{P}(C_n),$$

откуда и следует первое утверждение леммы.

Если события $C_1, C_2, \ldots$ независимы, то также независимы и их дополнения $\bar{C}_1, \bar{C}_2, \ldots$ Тогда для любого $N \ge n$

$$\mathsf{P}(\bigcap_{k=n}^{N} \bar{C}_k) = \prod_{k=n}^{N} \mathsf{P}(\bar{C}_k).$$

Поэтому

$$\mathsf{P}(\bigcap_{k=n}^{\infty} \bar{C}_k) = \prod_{k=n}^{\infty} \mathsf{P}(\bar{C}_k).$$

Воспользовавшись неравенством

$$\log(1-x) \le -x, \quad 0 \le x < 1,$$

получим:

$$\log \prod_{k=n}^{\infty} (1 - \mathsf{P}(C_k)) = \sum_{k=n}^{\infty} \log(1 - \mathsf{P}(C_k)) \le$$

$$\le -\sum_{k=n}^{\infty} \mathsf{P}(C_k) = -\infty.$$

Следовательно, для любого n

$$\mathsf{P}(\bigcap_{k=n}^{N} \bar{C}_k) = 0.$$

Отсюда следует, что

$$\mathsf{P}\{C_n \text{ б.ч.}\} = 1.$$

Литература

Кестен Х. Теория просачивания для математиков. М., Мир, 1986. М.В. Меньшиков М. В., Молчанов С. А., Сидоренко А. Ф. Теория перколяции и некоторые приложения. В книге "Итоги науки и техники. Теория вероятностей, математическая статистика. Теоретическая кибернетика." М., ВИНИТИ, 1986, т. 24, с. 53-110.

Малышев Вадим Александрович

Меньшиков Михаил Васильевич

Петрова Елена Николаевна

ВВЕДЕНИЕ В ТЕОРИЮ ВЕРОЯТНОСТЕЙ

М.: Издательство механико-математического факультета МГУ,
1997.—117 с.

Подписано в печать 27.04.1997 г.
Формат 60×901/16. Объем 7,5 п.л.
Заказ 8. Тираж 500 экз.

Издательство механико-математического факультета МГУ
г. Москва, Ленинские горы.
Лицензия на издательскую деятельность ЛР N 020806,
от 23.08.1993г.

Отпечатано на типографском оборудовании механико-математического факультета и франко-русского центра им. А.М. Ляпунова.