Тема 12. Независимость

Независимость событий

Определение. *Условной вероятностью* события A при условии события B называется

$$\mathbf{P}(A\,|\,B) = \frac{\mathbf{P}(A \bigcap B)}{\mathbf{P}(B)}, \tag{1}$$

где предполагается, что $\mathbf{P}(B) > 0$.

Пример 1. Из колоды, состоящей из 36 карт, наудачу выбирается одна карта. Рассмотрим события A $=$ {вынута пика} и B $=$ {вынута пика или трефа или бубна}. Тогда

$$\mathbf{P}(A \bigcap B) = \mathbf{P}(A) = 9/36 = 1/4, \qquad \mathbf{P}(B) = 27/36 = 3/4, \qquad \mathbf{P}(A\,|\,B) = \frac{1/4}{3/4} = 1/3.$$

Определение. Событие A *не зависит* от события B, если $\mathbf{P}(A\,|\,B) = \mathbf{P}(A)$, что (при $\mathbf{P}(B) > 0$) равносильно условию

$$\mathbf{P}(A \bigcap B) = \mathbf{P}(A) \cdot \mathbf{P}(B), \tag{2}$$

которое симметрично относительно A и B. Его и будем считать определением *независимости* событий A и B, пригодным даже в случае $\mathbf{P}(A) = 0$ или $\mathbf{P}(B) = 0$.

<u>Не перепутайте</u> независимые события с несовместными, т. е. с непересекающимися множествами!

Например, событие A $=$ {вынута пика} и событие C $=$ {вынут туз} являются независимыми:

$$\mathbf{P}(A \bigcap C) = 1/36, \qquad \mathbf{P}(A) = 9/36 = 1/4, \qquad \mathbf{P}(C) = 4/36 = 1/9, \qquad 1/36 = (1/4) \cdot (1/9).$$

Однако если добавить в колоду джокера (карту без масти и наименования), то эти события уже станут формально зависимыми: $1/37 \neq (4/37) \cdot (9/37)$.

Независимость случайных величин

Понятие независимости случайных величин играет определяющую роль в теории вероятностей, выделяя вероятностные задачи из проблем теории меры и математического анализа. Даже существует такое шуточное определение:

***теория вероятностей* = *теория меры* + *независимость*.**

Определение. <u>Дискретные</u> случайные величины X и Y называются *независимыми*, если для любой пары их значений (x_i, y_j) выполняется равенство

$$\mathbf{P}(X = x_i, Y = y_j) = \mathbf{P}(X = x_i) \cdot \mathbf{P}(Y = y_j). \tag{3}$$

Иначе говоря, для произвольных чисел x_i и y_j события $\{X = x_i\}$ и $\{Y = y_j\}$ являются независимыми. Равенство (3) также означает, что совместное распределение получается как произведение частных.

Пример 2. Пусть в урне находятся m занумерованных шаров. Шары с номерами 1, 2, … , l белого цвета, а остальные шары — чёрного. Наудачу с возвращением извлекают шары. Появление белого шара назовём «успехом», чёрного — «неудачей». Тогда доля белых шаров в урне $p = l/m$ задаёт вероятность «успеха» при каждом извлечении шара. Положим $q = 1 - p$.

Определим случайную величину L_1 как число извлечений до первого «успеха» (включительно) и случайную величину L_2 как число извлечений после первого до второго «успеха» (включительно). Каждая из случайных величин может принимать счётное множество значений: 1, 2, … . Нетрудно понять, что совместное распределение случайных величин L_1 и L_2 имеет вид

$$\mathbf{P}(L_1 = i, L_2 = j) = \frac{(m-l)^{i-1} l (m-l)^{j-1} l}{m^{i+j}} = q^{i-1} p q^{j-1} p, \quad i, j = 1, 2, \ldots .$$

Суммируя эти вероятности по j от 1 до ∞ и учитывая, что сумма всех вероятностей геометрического распределения равна 1, находим частное распределение случайной величины L_1:

$$\mathbf{P}(L_1 = i) = \sum_{j=1}^{\infty} \mathbf{P}(L_1 = i, L_2 = j) = q^{i-1} p \sum_{j=1}^{\infty} q^{j-1} p = q^{i-1} p.$$

Аналогично получаем, что $\mathbf{P}(L_2 = j) = q^{j-1} p$. Заметим, что для произвольных i и j верно равенство

$$\mathbf{P}(L_1 = i, L_2 = j) = \mathbf{P}(L_1 = i) \cdot \mathbf{P}(L_2 = j),$$

что и доказывает независимость случайных величин L_1 и L_2.

Независимость дискретных случайных величин без труда обобщается на n-мерный случай.

Определение. Дискретные случайные величины $X_1, \ldots, X_n$ называются *независимыми*, если для произвольного набора их значений $x_1, \ldots, x_n$ выполняется равенство

$$\mathbf{P}(X_1 = x_1, \ \ldots, X_n = x_n) = \prod_{i=1}^{n} \mathbf{P}(X_i = x_i). \tag{4}$$

Пример 3. Обобщим пример 2. Определим случайную величину L_i как число извлечений от $(i-1)$-го «успеха» (исключительно) до i-го «успеха» (включительно). Суммируя совместные вероятности как в примере 2, получим, что каждая случайная величина L_i имеет геометрическое распределение, и случайные величины $L_1, \ldots, L_n$ являются независимыми.

Пример 4. В условиях примера 2 обозначим через I_k индикатор «успеха» при k-м извлечении шара. Легко проверить, что $\mathbf{P}(I_k = 1) = p$, $\mathbf{P}(I_k = 0) = q$. Эти два равенства можно записать одной формулой:

$$\mathbf{P}(I_k = x) = p^x q^{1-x}, \tag{5}$$

где переменная x принимает только значения 0 или 1. Напомним, что в теме 5 случайные величины $I_1, I_2, \ldots$ назывались *испытаниями Бернулли* или *схемой Бернулли* с вероятностью «успеха» p. Применяя формулу (5) нетрудно установить, что испытания Бернулли $I_1, \ldots, I_n$ являются независимыми случайными величинами (*выведите формулу типа формулы* (5) *для совместной вероятности* $\mathbf{P}(I_1 = x_1, \ldots, I_n = x_n)$.)

Теперь дадим определение независимости случайных величин $X_1, \ldots, X_n$ в общем случае.

Определение. Компоненты случайного вектора $\boldsymbol{X} = (X_1, \ldots, X_n)$ называются *независимыми*, если для произвольных действительных чисел $x_1, \ldots, x_n$ выполняется равенство

$$F_{\boldsymbol{X}}(x_1, \ldots, x_n) = \mathbf{P}(X_1 \le x_1, \ldots, X_n \le x_n) = \prod_{i=1}^{n} \mathbf{P}(X_i \le x_i) = \prod_{i=1}^{n} F_{X_i}(x_i). \qquad (6)$$

Пример 5. Пусть точка $\boldsymbol{Y} = (Y_1, \ldots, Y_n)$ выбирается наудачу в *n*-мерном единичном кубе $[0, 1]^n$. Тогда для $0 \le x_1 \le 1, \ldots, 0 \le x_n \le 1$ совместная функция распределения $F_{\boldsymbol{Y}}(x_1, \ldots, x_n) = x_1 \cdot x_2 \cdot \ldots \cdot x_n$. При этом

$$F_{Y_i}(x_i) = F_{\boldsymbol{Y}}(1, \ldots, 1, x_i, 1, \ldots, 1) = x_i.$$

Следовательно, каждая случайная величина Y_i имеет равномерное распределение на отрезке [0, 1], и случайные величины $Y_1, \ldots, Y_n$ независимы.

Критерием независимости компонент случайного вектора $\boldsymbol{X} = (X_1, \ldots, X_n)$, имеющего плотность $f_{\boldsymbol{X}}(x_1, \ldots, x_n)$, служит выполнение для произвольных чисел $x_1, \ldots, x_n$ равенства

$$f_{\boldsymbol{X}}(x_1, \ldots, x_n) = \prod_{i=1}^{n} f_{X_i}(x_i). \qquad (7)$$

Здесь $f_{X_i}(x)$ обозначает плотность случайной величины X_i.

Пример 6. Пусть случайные величины $T_1, \ldots, T_n$ одинаково распределены по показательному закону с параметром $\lambda > 0$ и независимы. Тогда их совместная плотность задаётся формулой

$$f_{T_1, \ldots, T_n}(x_1, \ldots, x_n) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i} I_{\{x_i \ge 0\}} = \lambda^n e^{-\lambda \sum x_i} I_{\{\min\{x_1, \ldots, x_n\} \ge 0\}},$$

где I_{A} обозначает индикатор множества A.

Свойства независимых случайных величин

1) Пусть $X_1, \ldots, X_n$ — независимые случайные величины, f и g — непрерывные функции от k и $n-k$ переменных соответственно. Тогда случайные величины $Y = f(X_1, \ldots, X_k)$ и $Z = g(X_{k+1}, \ldots, X_n)$ также являются независимыми.

2) Если случайные величины X и Y с конечными математическими ожиданиями независимы, то

$$\mathrm{M}XY = \mathrm{M}X \cdot \mathrm{M}Y.$$

Это свойство можно переформулировать так: ковариация независимых случайных величин равна 0. Обратное утверждение неверно (см. задачу 12.4).

3) Если случайные величины с конечными дисперсиями $X_1, \ldots, X_n$ независимы, то

$$\mathrm{D}(X_1 + \ldots + X_n) = \mathrm{D}X_1 + \ldots + \mathrm{D}X_n.$$

Это важное свойство вытекает из предыдущего свойства и свойства 3 ковариации из темы 11.

Формулы свёртки

Формулы свёртки позволяют вычислять распределение суммы независимых случайных величин. В частности, распределение суммы двух независимых <u>дискретных</u> случайных величин X и Y вычисляется с помощью *дискретной формулы свёртки*:

$$\mathbf{P}(X+Y=x)=\sum_j \mathbf{P}(X=x-y_j)\,\mathbf{P}(Y=y_j). \qquad (8)$$

Эта формула немедленно следует из аксиомы счётной аддитивности и определения независимости:

$$\mathbf{P}(X+Y=x)=\sum_j \mathbf{P}(X+Y=x,\,Y=y_j)=\sum_j \mathbf{P}(X=x-y_j,\,Y=y_j)=\sum_j \mathbf{P}(X=x-y_j)\,\mathbf{P}(Y=y_j).$$

В свою очередь, плотность распределения суммы независимых случайных величин X и Y, <u>имеющих плотности</u> $f_X(x)$ и $f_Y(x)$ соответственно, вычисляется согласно формуле

$$f_{X+Y}(x)=\int_{-\infty}^{+\infty} f_X(x-y)\,f_Y(y)\,dy. \qquad (9)$$

Отметим, что справа в формуле (9) стоит несобственный интеграл, зависящий от параметра x. В явном виде его удаётся найти только для некоторых наиболее простых плотностей.

Домашнее задание

Если **второй** буквой вашего **имени** служит:

А, Б, В, Г, Д, Е, Ё, то «своими» являются задачи 12.1 и 12.5;
Ж, З, И, Й, К, Л, М, то «своими» являются задачи 12.2 и 12.6;
Н, О, П, Р, то «своими» являются задачи 12.3 и 12.7;
С, Т, У, Ф, Х, Ц, Ч, Ш, Щ, Ъ, Ы, Ь, Э, Ю, Я, то «своими» являются задачи 12.4 и 12.8.

12.1) Из урны, содержащей 6 занумерованных шаров, извлекают наудачу 2 шара. Пусть N_i — номер i-го шара, i = 1, 2. Являются ли случайные величины N_1 и N_2 независимыми, если шары извлекаются: а) с возвращением; б) без возвращения?

12.2) Доказать, что если событие A не зависит от события B, то:
а) A не зависит от $\overline{\mathrm{B}}$; б) $\overline{\mathrm{A}}$ не зависит от $\overline{\mathrm{B}}$. (*Указание*. Используйте аддитивность вероятности.)

12.3) Случайные величины Y_1 и Y_2 имеют равномерное распределение на отрезке [0, 1] и независимы. Найти формулу и построить график плотности случайной величины Y_1+Y_2.
(*Указание.* Используйте индикаторы в формуле свёртки и рассмотрите 2 случая: $0\le x\le 1$ и $1\le x\le 2$.)

12.4) Пусть случайная величина X принимает значения 0, $\pi/2$, π с вероятностями 1/3 каждое. Доказать, что случайные величины $Y=\sin X$ и $Z=\cos X$ некоррелированы, но зависимы.

12.5) Случайные величины Y_1 и Y_2 имеют равномерное распределение на отрезке [0, 1] и независимы. Найти формулу и построить график плотности случайной величины Y_1-Y_2.
(*Указание*. Найдите формулу для функции распределения и плотности случайной величины $-Y_2$, затем используйте формулу свёртки. Рассмотрите два случая: $-1\le x\le 0$ и $0\le x\le 1$.)

12.6) Пусть случайные величины N_1 и N_2 независимы и имеют пуассоновские распределения с параметрами λ и μ соответственно. Найти (без знака суммы) распределение случайной величины $N_1 + N_2$. (*Указание*. Используйте дискретную формулу свёртки и бином Ньютона.)

12.7) Доказать свойство 2 независимых случайных величин в дискретном случае.

12.8) Случайные величины T_1 и T_2 имеют одинаковое экспоненциальное распределение с параметром λ. Найти формулу и построить график плотности случайной величины $T_1 - T_2$. (*Указание*. Найдите формулу для функции распределения и плотности случайной величины $-T_2$, затем используйте формулу свёртки. Рассмотрите два случая: $x < 0$ и $x \geq 0$.)

12.9)* Найти по индукции формулу плотности суммы n независимых и одинаково распределённых показательных случайных величин с параметром λ. (*Указание*. Используйте формулу свёртки.)