

## 13. Ковариация и корреляция

### 13.1. Предварительные сведения

Пусть  $\xi$  и  $\eta$  — случайные величины, имеющие конечные вторые моменты:  $\mathbb{E}\xi^2 < \infty$ ,  $\mathbb{E}\eta^2 < \infty$ .

Ковариацией  $\text{Cov}(\xi, \eta)$  называется число  $\mathbb{E}(\xi - \mathbb{E}\xi)(\eta - \mathbb{E}\eta) = \mathbb{E}\xi\eta - \mathbb{E}\xi\mathbb{E}\eta$ .

Ковариация линейна по каждому своему аргументу (следует из свойств математического ожидания).

Для дискретных случайных величин с совместным распределением  $p_{ij} = \mathbb{P}(\xi = x_i, \eta = y_j)$  и одномерными распределениями  $p_{i\cdot} = \mathbb{P}(\xi = x_i) = \sum_j p_{ij}$  и  $p_{\cdot j} = \mathbb{P}(\eta = y_j) = \sum_i p_{ij}$  ковариация находится по формуле

$$\text{Cov}(\xi, \eta) = \sum_{ij} x_i y_j p_{ij} - \sum_i x_i p_{i\cdot} \sum_j y_j p_{\cdot j}.$$

Для абсолютно непрерывных случайных величин с совместной плотностью  $p_{\xi, \eta}(x, y)$  и одномерными плотностями  $p_\xi(x) = \int_{-\infty}^{\infty} p_{\xi, \eta}(x, y) dy$  и  $p_\eta(y) = \int_{-\infty}^{\infty} p_{\xi, \eta}(x, y) dx$  ковариация находится по формуле

$$\text{Cov}(\xi, \eta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x y p_{\xi, \eta}(x, y) dx dy - \int_{-\infty}^{\infty} x p_\xi(x) dx \int_{-\infty}^{\infty} y p_\eta(y) dy.$$

Коэффициентом корреляции случайных величин  $\xi$  и  $\eta$  называется число

$$\rho(\xi, \eta) = \frac{\mathbb{E}\xi\eta - \mathbb{E}\xi\mathbb{E}\eta}{\sqrt{\text{Var}\xi \text{Var}\eta}} = \frac{\text{Cov}(\xi, \eta)}{\sqrt{\text{Var}\xi \text{Var}\eta}}.$$

- $\xi$  и  $\eta$  независимы  $\Rightarrow \rho(\xi, \eta) = 0 \Leftrightarrow \xi$  и  $\eta$  некоррелированы;
- $|\rho(\xi, \eta)| \leq 1$  и  $\rho(\xi, \eta) = \pm 1 \Leftrightarrow \xi$  и  $\eta$  связаны линейно:  $\eta = a\xi + b$ ,  $a, b \in \mathbb{R}$ ,  $a \neq 0$ .

### 13.2. Практическое занятие

**1.** Игральная кость подброшена два раза. Пусть  $\xi$  — количество выпавших единиц, а  $\eta$  — количество выпавших шестерок. Выпишите совместное распределение  $(\xi, \eta)$  и найдите ковариацию  $\text{Cov}(\xi, \eta)$  и коэффициент корреляции  $\rho(\xi, \eta)$ . Согласуется ли знак коэффициента корреляции с вашими представлениями о характере зависимости  $\xi$  и  $\eta$ ?

**Ответ.**  $\text{Cov}(\xi, \eta) = -1/18$ ,  $\rho(\xi, \eta) = -1/5$ .

**2.** Случайная точка с координатами  $(\xi, \eta)$  равномерно распределена в треугольнике с вершинами в точках с координатами  $(0, 0)$ ,  $(1, 0)$  и  $(0, 1)$ . Найдите совместную плотность распределения  $p_{\xi, \eta}(x, y)$ , одномерные плотности  $p_\xi(x)$  и  $p_\eta(y)$  и вычислите ковариацию  $\text{Cov}(\xi, \eta)$  и коэффициент корреляции  $\rho(\xi, \eta)$ . Согласуется ли знак  $\rho(\xi, \eta)$  с вашими представлениями о характере зависимости  $\xi$  и  $\eta$ ?

**Ответ.**  $p_{\xi, \eta}(x, y) = \begin{cases} 2, & x \geq 0, y \geq 0, x + y \leq 1, \\ 0, & \text{иначе,} \end{cases}$   $p_\xi(x) = \begin{cases} 0, & x < 0 \text{ или } x > 1, \\ 2 - 2x, & 0 \leq x \leq 1, \end{cases}$   
 $p_\eta(y) = \begin{cases} 0, & y < 0 \text{ или } y > 1, \\ 2 - 2y, & 0 \leq y \leq 1, \end{cases}$   $\text{Cov}(\xi, \eta) = -1/36$ ,  $\rho(\xi, \eta) = -1/2$ .

**3.** Случайная величина  $\xi$  равномерно распределена на отрезке  $[-\pi, \pi]$ . Докажите, что случайные величины  $\eta_1 = \sin \xi$  и  $\eta_2 = \cos \xi$  некоррелированы, хотя и являются зависимыми.

4. Некоторая величина отклоняется от своего среднего значения под воздействием двух случайных факторов  $A$  и  $B$ . Стандартное отклонение, вызванное фактором  $A$ , равно 1.2, а фактором  $B$  — 1.1. Коэффициент корреляции между этими отклонениями равен  $1/3$ . Найдите стандартное отклонение, вызываемое совместным действием обоих факторов.

**Указание.** Вычислите дисперсию отклонения, вызванного совместным действием обоих факторов.

**Ответ.**  $\sqrt{1.2^2 + 1.1^2 + 2 \cdot \frac{1}{3} \cdot 1.2 \cdot 1.1} \approx 1.88$ .

### 13.3. Домашнее задание

5. В одной из групп 19 человек получили оценку за контрольную работу по теории вероятностей шкале от 2 до 5 баллов и экзаменационную оценку на зимней сессии. Пусть  $(\xi, \eta)$  — случайный вектор, представляющий собой пару оценок, полученных наудачу выбранным студентом. Распределение этого случайного вектора имеет вид:

$\xi \setminus \eta$	2	3	4	5
2	0	1/19	0	0
3	1/19	1/19	2/19	3/19
4	0	0	1/19	3/19
5	0	0	0	7/19

Найдите  $E\xi$ ,  $E\eta$ ,  $\text{Var } \xi$ ,  $\text{Var } \eta$ ,  $\text{Cov}(\xi, \eta)$  и  $\rho(\xi, \eta)$ . Согласуется ли знак  $\rho(\xi, \eta)$  с вашими представлениями о характере зависимости  $\xi$  и  $\eta$ ?

**Ответ.**  $E\xi = \frac{74}{19} \approx 3.89$ ,  $E\eta = \frac{85}{19} \approx 4.47$ ,  $\text{Var } \xi = \frac{338}{361} \approx 0.94$ ,  $\text{Var } \eta = \frac{280}{361} \approx 0.78$ ,  $\text{Cov}(\xi, \eta) = \frac{189}{361} \approx 0.52$ ,  $\rho(\xi, \eta) = \frac{189}{\sqrt{338 \cdot 280}} \approx 0.61$ . Положительная корреляция должна быть ожидаема.

6. Случайные величины  $\xi$  и  $\eta$  независимы и имеют стандартное нормальное распределение. Найдите коэффициент корреляции величин  $\alpha\xi + \beta\eta$  и  $\alpha\xi - \beta\eta$ ,  $\alpha, \beta \in \mathbb{R}$ .

**Ответ.**  $\frac{\alpha^2 - \beta^2}{\alpha^2 + \beta^2}$ .

7. На отрезок  $[0, 1]$  наудачу брошены две точки. Пусть  $\zeta$  — координата левой точки,  $\eta$  — координата правой точки. Найдите плотность совместного распределения  $\zeta$  и  $\eta$  и вычислите коэффициент корреляции  $\rho(\zeta, \eta)$ .

**Ответ.**  $p_{\zeta, \eta}(x, y) = \begin{cases} 2, & 0 < x < y < 1, \\ 0, & \text{иначе,} \end{cases} \quad \rho(\zeta, \eta) = 1/2$ .

8. Зависимые случайные величины  $\xi$  и  $\eta$  обладают конечными дисперсиями:  $\text{Var } \xi = \sigma_\xi^2$ ,  $\text{Var } \eta = \sigma_\eta^2$ . Укажите верхнюю и нижнюю границы, в которых может находиться дисперсия  $\text{Var}(\xi + \eta)$ . Приведите примеры, когда указанные границы достигаются.

**Указание.** Используйте ограниченность коэффициента корреляции.

**Ответ.**  $(\sigma_\xi - \sigma_\eta)^2 \leq \text{Var}(\xi + \eta) \leq (\sigma_\xi + \sigma_\eta)^2$ .

9. В продукции завода брак вследствие дефекта  $A$  составляет 3%, а вследствие дефекта  $B$  — 4.5%. Годная продукция составляет 95%. Найдите коэффициент корреляции дефектов  $A$  и  $B$ .

**Указание.** Пусть  $\xi = 1$ , если изделие имеет дефект  $A$ , и  $\xi = 0$ , если изделие не имеет дефекта  $A$ . Аналогично введите величину  $\eta$ , указывающую на присутствие дефекта  $B$ . По данным из условия задачи составьте таблицу совместного распределения  $(\xi, \eta)$ .

**Ответ.**  $\approx 0.669$ .