Поведение отклонений

В следующих задачах можно считать, что случайные величины принимают лишь конечное число значений. Для случайной величины ξ, принимающей значения a_i с вероятностями p_i, её *математическим ожиданием* называется величина $\mathbb{E}\xi := \sum_i p_i a_i$.

Пусть в эксперименте наблюдаются случайные величины ξ и η, принимающие значения a_i и b_j с вероятностями p_i и q_j соответственно. Естественно сказать, что они независимы, если вероятность того, что ξ принимает значение a_i, а η значение b_j, равна $p_i q_j$.

Задача 1. Если ξ и η независимы, то

$$\mathbb{E}(\xi\eta) = \mathbb{E}\xi \cdot \mathbb{E}\eta.$$

Задача 2 («Теорема Пифагора»)**.** Если ξ и η независимы, и *имеют нулевые математические ожидания*, то

$$\mathbb{E}(\xi+\eta)^2 = \mathbb{E}\xi^2 + \mathbb{E}\eta^2.$$

Дисперсией называется среднее значение квадрата отклонения от матожидания:

$$\mathbb{D}\xi := \mathbb{E}(\xi - \mathbb{E}\xi)^2$$

Задача 3. Дисперсия суммы независимых случайных величин равна сумме дисперсий.

Задача 4.

$$\mathbb{D}\xi = \mathbb{E}(\xi^2) - (\mathbb{E}\xi)^2.$$

Замечание. Задача 2 использует, что если случайные величины с нулевым средним независимы, то они «перпендикулярны» относительно скалярного произведения $\langle \xi, \eta \rangle_0 = \mathbb{E}\xi\eta$. То же самое правда без условия нулевого среднего, если сначала вычесть из случайной величины её среднее — но получающееся скалярное произведение,

$$\langle \xi, \eta \rangle := \mathbb{E}((\xi - \mathbb{E}\xi)(\eta - \mathbb{E}\eta)) = \mathbb{E}(\xi - \mathbb{E}\xi) \cdot \mathbb{E}(\eta - \mathbb{E}\eta).$$

получается полуопределённым: «скалярный квадрат» константы будет равен нулю.

Задача 5. Пусть ξ_i — независимые подбрасывания монетки (0 и 1 с вероятностями по $\frac{1}{2}$), а

$$S_n := \sum_{i=1}^{n} \xi_i.$$

Чему равна дисперсия S_n? А дисперсия доли орлов $\frac{S_n}{n}$?

Следующие несколько задач посвящены простейшей версии *центральной предельной теоремы* — теореме Муавра-Лапласа. А именно: предположим, что мы подбрасываем честную монету $2n$ раз, и интересуемся отклонением ξ числа орлов от его среднего значения n.

Задача 6. а) Найдите вероятность $p_k = \mathbb{P}(\xi = k)$ того, что количество орлов равно $n+k$.
Ответ. $p_k = \binom{2n}{n+k} \cdot 2^{-2n}$.
б) При каком k эта вероятность максимальна?

Задача 7. а) Упростите выражение $\binom{N}{a} / \binom{N}{a+1}$.
б) Запишите отношение p_k/p_0 в виде произведения k сомножителей, близких к 1 при $|k| \ll n$.
Ответ.

$$\frac{p_k}{p_0} = \left(1 - \frac{1}{n+1}\right) \dots \left(1 - \frac{2k-1}{n+k}\right). \tag{1}$$

Задача 8. Куда стремится выражение (1) при одновременном стремлении $n, k \to \infty$, таком, что $\frac{k}{\sqrt{n}} \to x \in \mathbb{R}$?
Указание. Общий принцип: *«Видишь длинное произведение? Прологарифмируй!»*
Указание. $\ln(1+\varepsilon) \approx \varepsilon$.
Ответ. $\exp(-x^2)$.

Задача 9. Найдите интеграл $\int_{-\infty}^{\infty} e^{-x^2}\, dx$.
Указание. Возведите интеграл в квадрат и перейдите в полярные координаты
Ответ. $\sqrt{\pi}$.

Задача 10 (Центральная предельная теорема — теорема Муавра–Лапласа)**.**
а) Докажите, что $p_0 \sim \frac{1}{\sqrt{\pi n}}$.
Указание. $\sum_k p_k = 1$; превратите эту сумму в (почти) интегральную сумму Римана.
б) Пусть заданы $x_1 < x_2$; найдите предел вероятности события

$$\mathbb{P}(x_1 \leq \frac{\xi}{\sqrt{n}} \leq x_2)$$

Ответ. $\int_{x_1}^{x_2} \frac{1}{\sqrt{\pi}} e^{-x^2}\, dx$.

Замечание. Обычно в центральной предельной используют другую нормировку ответа в центральной предельной теореме — деля на корень из дисперсии и получая предельный закон с дисперсией, равной 1. Оно называется *стандартным гауссовым (или нормальным) распределением*, а его плотность равна

$$\rho(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}};$$

Урна Пойа

Задача 11. а) Найдите вероятность любого пути из (a, b) в (m, n) в урне Пойа; в частности, проверьте, что все эти пути равновероятны.
Ответ.

$$\frac{(m-1)!}{(a-1)!} \cdot \frac{(n-1)!}{(b-1)!} \cdot \frac{(a+b-1)!}{(m+n-1)!}.$$

б) С какой вероятностью, начав из $(1, 1)$, после n шагов мы попадём в $(k, n+2-k)$?
Ответ. $\frac{1}{n+1}$.

в) Найдите вероятность попасть в (m, n), начав из (a, b). Как ведёт себя эта вероятность при $m, n \to \infty$, $\frac{m}{m+n} \to x$?
Ответ.

$$p_{(a,b)\to(m,n)} = \frac{(m-1)!}{(a-1)!} \cdot \frac{(n-1)!}{(b-1)!} \cdot \frac{(a+b-1)!}{(m+n-1)!} \cdot \frac{(m+n-a-b)!}{(m-a)!(n-b)!}.$$

$$p_{(a,b)\to(m,n)} \sim \frac{1}{m+n} \cdot \frac{(a+b-1)!}{(a-1)!(b-1)!} \cdot x^{a-1}(1-x)^{b-1}.$$

Задача 12. Найдите интеграл $I_{a,b} := \int_0^1 x^{a-1}(1-x)^{b-1}\,dx$.
Ответ.

$$\frac{(a-1)!(b-1)!}{(a+b-1)!}.$$

Задача 13. Как выглядит функция плотности $\rho_{a,b}(x) = \frac{1}{I_{a,b}} x^{a-1}(1-x)^{b-1}$ при больших a, b и при $\frac{a}{a+b} = s$? Где у неё максимум, в окрестности какого радиуса вокруг этого максимума сосредоточена половина полной массы?