Лекция 4

Схема независимых испытаний Бернулли. Формула Бернулли и её асимптотики.

Событие A связано с экспериментом $\mathcal{E}$, $\mathsf{P}\{A\} = p$, $\mathsf{P}\{\overline{A}\} = 1 - p = q$.

Эксперимент $\mathcal{E}$ повторяется n раз. Пусть $P_n(k)$ - вероятность того, что в серии из n экспериментов событие A произошло ровно k раз, $k = 0, 1, \ldots, n$. Легко доказать, что

$$P_n(k) = C_n^k p^k q^{n-k}$$

- *формула Бернулли.*

Пример 1

Что более вероятно: выиграть у равносильного противника 2 партии из 4 или 4 из 8?

Решение: Будем считать, что ничьих быть не может. Встанем на сторону одного из игроков, так что вероятность выигрыша будет $p = \frac{1}{2}$ и вероятность проигрыша $q = 1 - p = \frac{1}{2}$. Тогда

Вероятность выиграть 2 из 4:

$$P_4(2) = C_4^2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^2 = 6 \cdot \frac{1}{16} = \frac{3}{8}$$

Вероятность выиграть 4 из 8:

$$P_8(4) = C_8^4 \left(\frac{1}{2}\right)^8 = 70 \cdot \frac{1}{256} = \frac{35}{128}$$

$\Rightarrow$ вероятнее выиграть 2 из 4, ибо $\frac{3}{8} > \frac{35}{128}$.

Пусть теперь ничьи возможны. В этом случае вероятность выигрыша $p = \frac{1}{3}$, а вероятность невыигрыша $q = 1 - \frac{1}{3} = \frac{2}{3}$ (вероятность ничьи и вероятность проигрыша по $\frac{1}{3}$). В этом случае

Вероятность выиграть 2 из 4:

$$P_4(2) = C_4^2 \left(\frac{1}{3}\right)^2 \left(\frac{2}{3}\right)^2 = 6 \cdot \frac{4}{81} = \frac{8}{27}$$

Вероятность выиграть 4 из 8:

$$P_8(4) = C_8^4 \left(\frac{1}{3}\right)^4 \left(\frac{2}{3}\right)^4 = 70 \cdot \frac{16}{6561} = \frac{1120}{6561}$$

Как видно, и в этом случае вероятнее выиграть 2 из 4.

Наиболее вероятное число успехов в серии из n испытаний

В формуле Бернулли

$$P_n(k) = C_n^k p^k q^{n-k}$$

при фиксированных n и p (а значит, и q) найдём такое k, при котором $P_n(k)$ будет максимальным.

Наиболее вероятное число успехов $k_{\text{н.в.}}$ - значение k, при котором (при фиксированных n и p) вероятность

$$P_n(k) = C_n^k p^k q^{n-k}$$

принимает наибольшее значение.

Поскольку

$$\frac{P_n(k+1)}{P_n(k)} = \frac{(n-k)p}{(k+1)q},$$

то

$$\begin{aligned} P_n(k+1) > P_n(k) &\Leftrightarrow k < np - q \\ P_n(k+1) = P_n(k) &\Leftrightarrow k = np - q \\ P_n(k+1) < P_n(k) &\Leftrightarrow k > np - q \end{aligned} \tag{1}$$

Обозначим $\alpha \stackrel{\text{def}}{=} np - q$. Тогда из соотношений (1) следует:

- если α - не целое, то $k_{\text{н.в.}} = [\alpha] + 1$;
- если α - целое, то $k_{\text{н.в.}} = \alpha$ и $k_{\text{н.в.}} = \alpha + 1$;

где $[\alpha]$ - целая часть числа α. Таким образом, в итоге можем написать:

$$np - q \leq k_{\text{н.в.}} \leq np + p. \tag{2}$$

Если $np-q$ не целое, то неравенствам (2) удовлетворяет единственное $k_{\text{н.в.}}$; если же $np - q$ целое, то будет два значения $k_{\text{н.в.}}$: $k_{\text{н.в.}} = np - q$ и $k_{\text{н.в.}} = np + p$.

Обобщенная формула Бернулли

Пусть $A_1, A_2, \ldots, A_k$ - полная группа событий, связанных с экспериментом $\mathcal{E}$. Пусть $P_n(m_1, m_2, \ldots, m_k)$ - вероятность того, что при проведении n экспериментов
событие A_1 произошло m_1 раз,
...
событие A_k произошло m_k раз,
так что $m_1 + m_2 + \cdots + m_k = n$. Обозначим
$\mathsf{P}\{A_1\} = p_1, \ldots, \mathsf{P}\{A_k\} = p_k$.

Тогда

$$P_n(m_1, m_2, \ldots, m_k) = \frac{n!}{m_1! m_2! \ldots m_k!} p_1^{m_1} p_2^{m_2} \ldots p_k^{m_k} \tag{3}$$

- *обобщенная формула Бернулли.*

Её доказательство аналогично доказательству классической формулы Бернулли. Последняя будет частным случаем формулы (3) при $k = 2$. В этом случае $A_1 = A$, $A_2 = \overline{A}$, $p_1 = p$, $p_2 = 1 - p = q$, $m_1 = m$, $m_2 = n - m$ и

$$P_n(m_1, m_2) = P_n(m) = \frac{n!}{m!(n-m)!} p^m q^{n-m}$$

Асимптотики формул Бернулли

Локальная теорема Муавра-Лапласа

Если вероятность наступления некоторого события в n независимых испытаниях постоянна и равна p $(0 < p < 1)$, то при $n \to \infty$

$$\frac{\sqrt{npq} P_n(k)}{\frac{1}{\sqrt{2\pi}} e^{-x^2/2}} \to 1$$

равномерно для всех k, для которых $x = x_{nk} = \frac{k - np}{\sqrt{npq}}$ находится в каком-либо конечном интервале.

Интегральная теорема Муавра-Лапласа

Для фиксированных z_1 и z_2 при $n \to \infty$

$$P_n(np + z_1\sqrt{npq} \le k \le np + z_2\sqrt{npq}) \to \Phi(z_2) - \Phi(z_1)$$

где $\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$ - функция Лапласа.

Теорема Пуассона

Если p_n - вероятность успеха в серии из n испытаний и $p_n \to 0$ при $n \to \infty$, то

$$P_n(k) - \frac{\lambda_n^k}{k!} e^{-\lambda_n} \to 0, \quad \text{где } \lambda_n = np_n$$