

Лекция 3

3. Последовательности независимых испытаний

3.1. Схема Бернулли

Одинаковые независимые между собой испытания называются *схемой Бернулли* или *испытаниями Бернулли*, если при каждом испытании имеется только два возможных исхода, причём вероятности этих исходов положительны и неизменны для всех испытаний.

Исходы испытаний обычно называют *успехом* и *неудачей* и обозначают соответственно буквами p и q . Очевидно, что $p > 0$, $q > 0$ и

$$p + q = 1. \quad (1.1)$$

Опишем вероятностное пространство, соответствующее n испытаниям Бернулли. Будем считать, что исходами каждого испытания являются либо 0, либо 1. Пусть успеху соответствует 1, а неудаче — 0. Пространством элементарных событий является множество

$$\Omega = \{\omega : \omega = (a_1, \dots, a_n), a_i = Y, H\}.$$

Вероятность элементарного события задаётся формулой

$$P(\omega) = p^{\sum a_i} q^{n - \sum a_i}.$$

Теорема 1.1. Если μ_n — число успехов в n испытаниях Бернулли, то

$$P(\mu_n = m) = C_n^m p^m q^{n-m}. \quad (1.2)$$

Доказательство. Если число успехов элементарного события ω равно m , то

$$P(\omega) = p^m q^{n-m}.$$

Очевидно, что количество таких событий равно C_n^m . ■

Пример 1. Монета подбрасывается 5 раз. Найти вероятность события A , состоящее в выпадении 3 гербов.

Решение. Выпадение герба будем считать успехом. Так как $p = q = \frac{1}{2}$, то из формулы (3.2) получаем

$$P(A) = C_5^3 \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^2 = \frac{10}{2^5} = \frac{5}{16} = 0,3125.$$

Одинаковые независимые между собой испытания называются *полиномиальной схемой*, если при каждом испытании имеется только k возможных исходов, причём вероятности этих исходов положительны и неизменны для всех испытаний. Схема Бернулли является частным случаем полиномиальной схемы при $k = 2$.

Пусть E_1, E_2, \dots, E_k — исходы испытаний и $p_i = P(E_i)$, $i = 1, \dots, k$. Очевидно, что $p_i > 0$ и

$$p_1 + \dots + p_k = 1. \quad (1.3)$$

Теорема 1.2. Рассмотрим полиномиальную схему, состоящую из n испытаний с возможными исходами E_1, E_2, \dots, E_k . Вероятность того, что событие E_i произойдет r_i раз, $i = 1, \dots, k$, равна

$$P_n(r_1, \dots, r_k) = \frac{n!}{r_1! \dots r_k!} p_1^{r_1} \dots p_k^{r_k}. \quad (1.4)$$

Доказательство. Вероятность элементарного события $\omega = (E_{i_1}, \dots, E_{i_n})$, в котором E_i встречается r_i раз, $i = 1, \dots, k$, равна

$$P(\omega) = p_1^{r_1} \dots p_k^{r_k}.$$

Из теоремы о полиномиальных коэффициентах следует, что количество таких событий равно $\frac{n!}{r_1! \dots r_k!}$. ■

Пример 2. Игральная кость бросается 12 раз. Найти вероятность события A , состоящее в том, что каждая грань выпадет дважды.

Решение. Пусть E_1, \dots, E_6 соответствуют шести граням. По условию все $p_i = 1/6$ и все $r_i = 2$. В силу (1.4) получаем

$$P(A) = \frac{12!}{(2!)^6} \left(\frac{1}{6}\right)^{12} = \frac{1925}{559872} = 0,003438 \dots$$

3.2. Локальная предельная теорема

Будем рассматривать схему Бернулли с вероятностью успеха p и вероятностью неудачи $q = 1 - p$. Число испытаний будем обозначать через n . Мы хотим получить асимптотическую оценку $P_n(m)$ при $n \rightarrow \infty$. Вывод этой оценки существенно опирается на формулу Стирлинга.

Формула Стирлинга:

$$n! = \sqrt{2\pi n} \left(\frac{n}{e}\right)^n \exp(\theta_n), \quad \frac{1}{12n+1} < \theta_n < \frac{1}{12n}. \quad (2.1)$$

Теорема 2.1. (Локальная теорема Муавра-Лапласа.) Если в схеме Бернулли $n \rightarrow \infty$, вероятность успеха p , $0 < p < 1$, постоянна, то для любого конечного промежутка $[a, b]$, $a \leq b$,

$$P_n(m) = \frac{1}{\sqrt{2\pi} \sqrt{npq}} e^{-x^2/2} (1 + O(n^{-1/2})). \quad (2.2)$$

равномерно для $x \in [a, b]$ вида

$$x = x_{mn} = \frac{m - np}{\sqrt{npq}}. \quad (2.3)$$

где m — целое неотрицательное число.

Доказательство. Так как

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

то, используя формулу Стирлинга, запишем

$$P_n(m) = \frac{\sqrt{2\pi n} e^{-n} n^n e^{(\theta_n - \theta_m - \theta_{n-m})} p^m q^{n-m}}{\sqrt{2\pi m} e^{-m} m^m \sqrt{2\pi(n-m)} e^{-(n-m)} (n-m)^{n-m}} =$$

$$= \frac{1}{\sqrt{2\pi}} \sqrt{\frac{n}{m(n-m)}} \left(\frac{np}{m}\right)^m \left(\frac{nq}{n-m}\right)^{n-m} e^\theta,$$

где $\theta = \theta_n - \theta_m - \theta_{n-m}$.

Из (2.3) следуют равенства

$$m = np + x \sqrt{npq} = np \left(1 + x \sqrt{\frac{q}{np}}\right),$$

$$n - m = nq - x \sqrt{npq} = nq \left(1 - x \sqrt{\frac{p}{nq}}\right). \quad (2.4)$$

Так как x принадлежит ограниченному промежутку $[a, b]$, то $m = O(n)$, $n - m = O(n)$ при $n \rightarrow \infty$ равномерно по $x \in [a, b]$. Следовательно,

$$|\theta| \leq |\theta_n| + |\theta_m| + |\theta_{n-m}| \leq \frac{1}{12n} + \frac{1}{12m} + \frac{1}{12(n-m)} = O(n^{-1}).$$

Поэтому,

$$e^\theta = 1 + O(n^{-1}) \quad (2.5)$$

при $n \rightarrow \infty$ равномерно по $x \in [a, b]$.

Рассмотрим величину

$$\ln A_n = \ln \left\{ \left(\frac{np}{m}\right)^m \left(\frac{nq}{n-m}\right)^{n-m} \right\} = m \ln \left(\frac{np}{m}\right) + (n-m) \ln \left(\frac{nq}{n-m}\right).$$

Из равенств (2.4) следует

$$\ln A_n = -np \left(1 + x \sqrt{\frac{q}{np}}\right) \ln \left(1 + x \sqrt{\frac{q}{np}}\right) - nq \left(1 - x \sqrt{\frac{p}{nq}}\right) \ln \left(1 - x \sqrt{\frac{p}{nq}}\right).$$

Величины $x \sqrt{q/np}$ и $x \sqrt{p/nq}$ есть величины $O(n^{-1/2})$ при $n \rightarrow \infty$ равномерно по $x \in [a, b]$. Следовательно,

$$\ln \left(1 + x \sqrt{\frac{q}{np}}\right) = x \sqrt{\frac{q}{np}} - \frac{1}{2} \frac{qx^2}{np} + O(n^{-3/2})$$

$$\ln \left(1 - x \sqrt{\frac{p}{nq}}\right) = -x \sqrt{\frac{p}{nq}} - \frac{1}{2} \frac{px^2}{nq} + O(n^{-3/2})$$

Таким образом, получаем

$$np \left(1 + x \sqrt{\frac{q}{np}}\right) \ln \left(1 + x \sqrt{\frac{q}{np}}\right) = x \sqrt{npq} + \frac{qx^2}{2} + O(n^{-1/2}),$$

$$nq \left(1 - x \sqrt{\frac{p}{nq}}\right) \ln \left(1 - x \sqrt{\frac{p}{nq}}\right) = -x \sqrt{npq} + \frac{px^2}{2} + O(n^{-1/2}),$$

и поэтому

$$\ln A_n = -\frac{x^2}{2} + O(n^{-1/2}). \quad (2.6)$$

Осталось рассмотреть множитель $\sqrt{\frac{n}{m(n-m)}}$. Из равенств (2.4) следует

$$m(n-m) = n^2 pq (1 + O(n^{-1/2})).$$

Отсюда получаем

$$\sqrt{\frac{n}{m(n-m)}} = \frac{1}{\sqrt{npq}} ((1 + O(n^{-1/2}))). \quad (2.7)$$

Теперь утверждение теоремы следует из (2.5), (2.6), (2.7). ■

Пример 1. Найти вероятность того, что событие A наступит ровно 80 раз в 400 испытаниях, если вероятность этого события в каждом испытании равна 0,2.

Решение. По условию $n = 400$, $m = 80$, $p = 0,2$, $q = 0,8$. Из локальной теоремы Муавра-Лапласа имеем

$$P_{400}(80) \approx \frac{1}{\sqrt{npq}} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} = \frac{1}{\sqrt{400 \cdot 0,2 \cdot 0,8}} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} = \frac{1}{8\sqrt{2\pi}} e^{-x^2/2}.$$

Находим x :

$$x = \frac{m - np}{\sqrt{npq}} = 0.$$

В результате

$$P_{400}(80) \approx \frac{1}{8\sqrt{2\pi}} = 0,0498678.$$

Точное значение

$$P_{400}(80) = \frac{400!}{80! 320!} 0,2^{80} 0,8^{320} = 0,049813272.... \quad \blacksquare$$

3.3. Интегральная предельная теорема

Будем рассматривать схему Бернулли с вероятностью успеха p и вероятностью неудачи q . Число испытаний будем обозначать через n . Пусть μ — число успехов в n испытаниях.

Теорема 3.1. (Интегральная теорема Муавра-Лапласа.) *Если в схеме Бернулли $n \rightarrow \infty$, вероятность успеха p , $0 < p < 1$, постоянна, то равномерно по $a \in \mathbb{R}$, $b \in \mathbb{R}$, $a < b$, имеет место соотношение*

$$P \left\{ a \leq \frac{\mu - np}{\sqrt{npq}} \leq b \right\} \rightarrow \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx. \quad (3.1)$$

Доказательство. Предположим сначала, что $|a| \leq C$, $|b| \leq C$. Пусть $]x[$ — наименьшее целое число такое, что $x \leq]x[$, а $[x]$ — наибольшее целое число такое, что $[x] \leq x$. Пусть

$$m_1 =]np + a\sqrt{npq}[, \quad m_2 = [np + b\sqrt{npq}].$$

Тогда

$$P \left\{ a \leq \frac{\mu - np}{\sqrt{npq}} \leq b \right\} = \sum_{m=m_1}^{m_2} P \{ \mu = m \}. \quad (3.2)$$

Обозначим $m = np + x_m \sqrt{npq}$, тогда

$$\Delta x_m = x_{m+1} - x_m = \frac{1}{\sqrt{npq}}.$$

По локальной предельной теореме запишем (3.2) в виде

$$\mathbf{P} \left\{ a \leq \frac{\mu - np}{\sqrt{npq}} \leq b \right\} = \sum_{m=m_1}^{m_2} \frac{1}{\sqrt{2\pi}} e^{-x_m^2/2} \Delta x_m (1 + O(n^{-1/2})). \quad (3.3)$$

Справа в (3.3) стоит интегральная сумма, сходящаяся равномерно по a, b при $n \rightarrow \infty$ к интегралу

$$\frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx.$$

Следовательно, теорема доказана при $|a| \leq C, |b| \leq C$.

Избавимся теперь от ограничения $|a| \leq C, |b| \leq C$. Обозначим $\xi_n = \frac{\mu - np}{\sqrt{npq}}$. Имеем равенство

$$\mathbf{P} \{ |\xi_n| > C \} = 1 - \mathbf{P} \{ |\xi_n| \leq C \}. \quad (3.4)$$

Из анализа известно, что

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2} dx = 1.$$

Поэтому

$$\frac{1}{\sqrt{2\pi}} \int_{-C}^C e^{-x^2/2} dx = 1 - \frac{1}{\sqrt{2\pi}} \int_{|x|>C} e^{-x^2/2} dx. \quad (3.5)$$

Из (3.4) и (3.5) получаем

$$\left| \mathbf{P} \{ |\xi_n| > C \} - \frac{1}{\sqrt{2\pi}} \int_{|x|>C} e^{-x^2/2} dx \right| = \left| \mathbf{P} \{ |\xi_n| \leq C \} - \frac{1}{\sqrt{2\pi}} \int_{-C}^C e^{-x^2/2} dx \right|. \quad (3.6)$$

Пусть задано $\varepsilon > 0$. Тогда найдётся такое C , что

$$\frac{1}{\sqrt{2\pi}} \int_{|x|>C} e^{-x^2/2} dx < \frac{\varepsilon}{8}. \quad (3.7)$$

Зафиксируем C . По доказанному выше найдётся такое n_1 , что для всех $n \geq n_1$

$$\left| \mathbf{P} \{ |\xi_n| \leq C \} - \frac{1}{\sqrt{2\pi}} \int_{-C}^C e^{-x^2/2} dx \right| < \frac{\varepsilon}{8},$$

откуда, в силу (3.6) и (3.7), для тех же $n \geq n_1$ имеем

$$\mathbf{P} \{ |\xi_n| > C \} \leq \frac{\varepsilon}{4}. \quad (3.8)$$

Возьмём теперь произвольный интервал $[a, b]$. Обозначим $[A, B] = [a, b] \cap [-C, C]$. Так как $-C \leq A \leq B \leq C$, то, по уже доказанному, существует такое n_2 , что для всех $n \geq n_2$ справедливо неравенство

$$\left| \mathbf{P} \{ \xi_n \in [A, B] \} - \frac{1}{\sqrt{2\pi}} \int_A^B e^{-x^2/2} dx \right| < \frac{\varepsilon}{2}. \quad (3.9)$$

Из неравенства

$$\begin{aligned} \left| \mathbf{P} \{ \xi_n \in [a, b] \} - \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx \right| &\leq \mathbf{P} \{ |\xi_n| > C \} + \\ &+ \frac{1}{\sqrt{2\pi}} \int_{|x|>C} e^{-x^2/2} dx + \left| \mathbf{P} \{ \xi_n \in [A, B] \} - \frac{1}{\sqrt{2\pi}} \int_A^B e^{-x^2/2} dx \right| \end{aligned}$$

получаем, в силу (3.7)–(3.9), что при $n \geq n_0 = \max\{n_1, n_2\}$

$$\left| P\{\xi_n \in [a, b]\} - \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx \right| < \varepsilon$$

равномерно по всем $a \leq b$. Теорема доказана. ■

Следствие. (Закон больших чисел Бернулли). Пусть выполнены условия интегральной теоремы Муавра-Лапласа. Тогда для любого $\varepsilon > 0$ справедливо равенство

$$\lim_{n \rightarrow \infty} P\left\{\left|\frac{\mu}{n} - p\right| < \varepsilon\right\} = 1. \quad (3.10)$$

Доказательство. Имеем

$$P\left\{\left|\frac{\mu}{n} - p\right| < \varepsilon\right\} = P\left\{-\varepsilon \sqrt{\frac{n}{pq}} < \frac{\mu - np}{\sqrt{npq}} < \varepsilon \sqrt{\frac{n}{pq}}\right\}.$$

В силу интегральной теоремы Муавра-Лапласа

$$\lim_{n \rightarrow \infty} P\left\{\left|\frac{\mu}{n} - p\right| < \varepsilon\right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2} dx = 1.$$

Следствие доказано. ■

3.4. Применение интегральной теоремы Муавра-Лапласа

Введём функции $\varphi(x)$ и $\Phi(x)$ равенствами:

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad \Phi(x) = \int_{-\infty}^x \varphi(t) dt. \quad (4.1)$$

Функция $\Phi(x)$ называется *нормальной функцией распределения*, а функция $\varphi(x)$ — *плотностью нормального распределения*.

Значения функций $\varphi(x)$, $\Phi(x)$ при $x > 0$ можно найти в специальных таблицах. При $x < 0$ можно воспользоваться тождеством

$$\Phi(x) + \Phi(-x) \equiv 1, \quad x \in \mathbb{R}. \quad (4.2)$$

Докажем (4.2). Пусть $x \in \mathbb{R}$, тогда

$$1 = \int_{-\infty}^{+\infty} \varphi(t) dt = \int_{-\infty}^x \varphi(t) dt + \int_x^{+\infty} \varphi(t) dt.$$

Первый интеграл равен $\Phi(x)$. Во втором интеграле сделаем замену $s = -t$. Получим

$$\int_x^{+\infty} \varphi(t) dt = - \int_{-x}^{-\infty} \varphi(-s) ds = \int_{-\infty}^{-x} \varphi(s) ds = \Phi(-x).$$

Здесь мы учли чётность функции $\varphi(x)$.

Предположим, что нам нужно вычислить вероятность $P\{m_1 \leq \mu \leq m_2\}$ в схеме Бернулли с n независимыми испытаниями и с вероятностью успеха p . Вычислим

$$x_1 = \frac{m_1 - np}{\sqrt{npq}}, \quad x_2 = \frac{m_2 - np}{\sqrt{npq}}. \quad (4.3)$$

Далее

$$P\{m_1 \leq \mu \leq m_2\} \approx \frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-t^2/2} dt = \Phi(x_2) - \Phi(x_1). \quad (4.4)$$

Пример 1. В партии из $n = 22500$ изделий, каждое изделие независимо от других может быть бракованным с вероятностью $p = 1/5$. Найти вероятность того, что число μ бракованных изделий находится между 4380 и 4560.

Решение. Значение $npq = 3600$ велико, поэтому можно воспользоваться интегральной теоремой Муавра-Лапласа. Вычисляем по формулам (4.3) значения x_1, x_2 :

$$x_1 = \frac{4380 - 4500}{60} = -2, \quad x_2 = \frac{4560 - 4500}{60} = 1.$$

Из интегральной теоремы Муавра-Лапласа получаем

$$P\{4380 \leq \mu \leq 4560\} = \Phi(1) - \Phi(-2).$$

Из таблицы находим: $\Phi(1) = 0,841345$, $\Phi(2) = 0,9772499$. По формуле (4.2) $\Phi(-2) = 0,0227501$. В результате

$$P\{4380 \leq \mu \leq 4560\} = 0,841345 - 0,0227501 = 0,8185949.$$

Точный расчёт даёт значение 0,8213141718.

При использовании формул (4.3), (4.4) мы допускаем погрешность. Эту погрешность можно значительно уменьшить, если немного изменить формулы (4.3), (4.4). Положим

$$x'_1 = \frac{m_1 - 1/2 - np}{\sqrt{npq}}, \quad x'_2 = \frac{m_2 + 1/2 - np}{\sqrt{npq}}, \quad (4.3^*)$$

$$P\{m_1 \leq \mu \leq m_2\} \approx \Phi(x'_2) - \Phi(x'_1). \quad (4.4^*)$$

Пример 2. Решим пример 1, используя формулы (4.3*), (4.4*).

$$x'_1 = \frac{4380 - 4500 - 1/2}{60} \approx -2,008, \quad x'_2 = \frac{4560 - 4500 + 1/2}{60} \approx 1,008.$$

Далее

$$P\{4380 \leq \mu \leq 4560\} = \Phi(1,008) - \Phi(-2,008).$$

Из таблицы находим: $\Phi(1,008) = 0,843273$, $\Phi(2) = 0,9776784$. По формуле (4.2) $\Phi(-2,008) = 0,0223216$. Получаем

$$P\{4380 \leq \mu \leq 4560\} = 0,843273 - 0,0223216 = 0,8209514.$$

Полученное значение ближе к точному значению 0,8213141718.

Часто вместо функции $\Phi(x)$ используют функцию

$$\Phi_0(x) = \int_0^x \varphi(t) dt = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt, \quad (4.5)$$

называемую *интегралом Лапласа*. Значения $\Phi_0(x)$ при положительных x приведены в специальных таблицах. Для отрицательных x следует воспользоваться нечётностью $\Phi_0(x)$: $\Phi_0(-x) = -\Phi_0(x)$. Отметим формулу:

$$\Phi(x) = \begin{cases} \frac{1}{2} + \Phi_0(x), & \text{при } x \geq 0, \\ \frac{1}{2} - \Phi_0(-x), & \text{при } x < 0. \end{cases} \quad (4.6)$$

3.5. Теорема Пуассона

Теорема 5.1. (Теорема Пуассона.) Если в схеме Бернулли $n \rightarrow \infty$, вероятность успеха $p \rightarrow 0$ так, что $np \rightarrow \lambda$, $0 < \lambda < \infty$, то

$$P_n(m) = C_n^m p^m q^{n-m} \rightarrow \frac{\lambda^m}{m!} e^{-\lambda}. \quad (5.1)$$

Доказательство. Положим $\lambda = np$. Имеем

$$\begin{aligned} P_n(m) &= \frac{n(n-1)\dots(n-m+1)}{m!} \left(\frac{\lambda_n}{n}\right)^m \left(1 - \frac{\lambda_n}{n}\right)^{n-m} = \\ &= \frac{\lambda_n^m}{m!} \left(1 - \frac{\lambda_n}{n}\right)^n \left(1 - \frac{\lambda_n}{n}\right)^{-m} \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{m-1}{n}\right). \end{aligned}$$

Отсюда при $n \rightarrow \infty$ получим утверждение теоремы. ■

Обозначим

$$P(m) = \frac{\lambda^m}{m!} e^{-\lambda}. \quad (5.2)$$

Полученное распределение вероятностей носит название *закона Пуассона*.

Пример 1. Вероятность попадания в цель при каждом выстреле равна $p = 0,001$. Найти вероятность попадания в цель двумя и более пулями, если число выстрелов равно 5000.

Решение. Будем считать, что каждый выстрел это испытание и попадание в цель это событие. Нужно вычислить $P\{\mu_n \geq 2\}$. В данном примере

$$\lambda = np = 0.001 \cdot 5000 = 5.$$

Искомая вероятность равна

$$P\{\mu_n \geq 2\} = \sum_{m=2}^n P_n(m) = 1 - P_n(0) - P_n(1).$$

По теореме Пуассона

$$P_n(0) \approx e^{-5}, \quad P_n(1) \approx 5e^{-5}.$$

Следовательно,

$$P\{\mu_n \geq 2\} \approx 1 - 6e^{-5} = 0,9595723180\dots$$

Точное вычисление даёт

$$\begin{aligned} P_n(0) &= (1-p)^n = 0.999^{5000} = 0,006721111960, \\ P_n(1) &= np(1-p)^{n-1} = 0.999^{5000} = 0,0336391990. \end{aligned}$$

В результате

$$P\{\mu_n \geq 2\} = 0,9596396890\dots$$

Ошибка от использования теоремы Пуассона составляет около 0,007%. ■