Формула полной вероятности. Формула Бейеса

Пусть событие A может произойти только с одним из событий $H_1, H_2, ..H_n$, образующих полную систему попарно несовместных событий (см. рис. 1.3). Тогда вероятность события A вычисляется по формуле *полной вероятности*

$$P(A) = \sum_{l=1}^{n} P(H_i) \cdot P(A/H_i) \quad . \qquad (2.10)$$

В тесной связи с формулой полной вероятности находится так называемая *формула Бейеса*

$$P(H_i/A) = \frac{P(H_i) \cdot P(A/H_i)}{P(H)P(A/H_1) + ... + P(H_n)P(A/H_n)}, \qquad (2.11)$$

где $P(H_i/A)$ - вероятность гипотезы H_i после того, как имело место событие A.

Формула Бейеса позволяет переоценить вероятность гипотез, принятых до испытания, по результатам уже произведённого испытания [8].

Задача 1. Имеются две урны № 1, три урны № 2 и пять урн № 3. Урны внешне не отличаются одна от другой. В урне № 1 имеется 1 белый и 4 черных шара; в урне № 2 – 5 белых и 3 черных шара, в урне № 3 – 6 белых и 9 черных шаров. Наугад берут одну из урн и из нее вынимают шар. Какова вероятность того, что вынутый шар окажется белым?

Решение. Пусть событие A – появление белого шара из урны, взятой наудачу. Это событие будет происходить совместно с выбором урны, из которой извлекается шар. Пусть события $H_1,\ H_2,\ H_3$ состоят в том, что будут выбраны урны № 1, № 2, № 3 соответственно.

Определим вероятности гипотез:

$$P(H_1) = \frac{2}{10}; \quad P(H_2) = \frac{3}{10}; \quad P(H_3) = \frac{5}{10}.$$

Найдем условные вероятности появления белого шара из соответствующих урн:

$$P(A/H_1) = \frac{1}{5}; \; P(A/H_2) = \frac{5}{8}; \; P(A/H_3) = \frac{6}{15}.$$

Искомая вероятность

$$P(A) = P(H_1) \cdot P\left(A/H_1\right) + P(H_2) \cdot P(A/H_2) + P(H_3) \cdot P\left(A/H_3\right) =$$
$$= \frac{2}{10} \cdot \frac{1}{5} + \frac{3}{10} \cdot \frac{5}{8} + \frac{5}{10} \cdot \frac{6}{15} = 0{,}4275.$$

Задача 2. Имеется две партии деталей, причём известно, что в одной партии все детали удовлетворяют техническим условиям, а в другой партии $1/4$ деталей недоброкачественных. Деталь, взятая из наудачу выбранной партии, оказалась доброкачественной. Определить вероятность того, что вторая деталь из этой же партии окажется недоброкачественной, если первая деталь после проверки возвращена в партию.

Решение. Пусть событие *A* – “первая деталь доброкачественная”.

Гипотезы:

H_1 – “взята партия, содержащая недоброкачественные детали”;

H_2 – “взята партия доброкачественных деталей”.

По условию задачи,

$$P(H_1) = P(H_2) = 1/2,\ P(A/H_1) = 3/4,\ P(A/H_2) = 1;$$

$$P(A) = \frac{1}{2}\cdot\frac{3}{4} + \frac{1}{2}\cdot 1 = \frac{7}{8} \approx 0{,}875.$$

После первого испытания вероятность того, что партия содержит недоброкачественные детали, равна

$$P(H_1/A) = \frac{P(H_1)P(A/H_1)}{P(A)} = \frac{\frac{1}{2}\cdot\frac{3}{4}}{0{,}875} \approx 0{,}4286\ ;$$

$$P(H_2/A) = \frac{P(H_2)P(A/H_2)}{P(A)} = \frac{\frac{1}{2}\cdot 1}{0{,}875} \approx 0{,}5714\ .$$

Пусть событие *B* состоит в том, что при втором испытании деталь оказалась недоброкачественной. Вероятность данного события также находится по формуле полной вероятности. Если P_1 и P_2 – вероятности гипотез H_1 и H_2 после испытания, то согласно предыдущим вычислениям

$$P_1 = 0{,}4286\ ;\ P_2 = 0{,}5714.$$

Кроме того, $P(B/H_1) = 1/4$; $P(B/H_2) = 0$.

Поэтому искомая вероятность:

$$P(B) = 0{,}4286\cdot\frac{1}{4} + 0{,}5714\cdot 0 \approx 0{,}107.$$

§5. Последовательность независимых испытаний (схема Бернулли). Теорема Пуассона. Локальная и интегральная теоремы Муавра – Лапласа

1. Последовательность независимых испытаний (схема Бернулли)

Испытания называются *независимыми* относительно события *A*, если вероятность появления события *A* в каждом из этих испытаний не зависит от результата, полученного в других испытаниях.

Пусть эксперимент состоит в проведении n независимых испытаний, в каждом из которых может произойти некоторое событие *A* (назовем его “успехом”, тогда $\overline{A}$ соответственно “неуспех”). Вероятность неуспеха равна $q = 1 - p$.

Рассмотрим общий случай в рамках схемы Бернулли – нахождение вероятности того, что в n испытаниях произойдёт ровно κ успехов $(\kappa \le n)$. Обозначим эту вероятность $P_n(\kappa)$. Событию B (произошло κ успехов в n испытаниях) благоприятствуют те элементарные события, в которые входит κ множителей A и $n-\kappa$ множителей $\overline{A}$; вероятности событий равны $p^{\kappa} \cdot q^{n-\kappa}$, а их число, как нетрудно видеть, равно числу способов, сколькими можно выбрать κ элементов из n без учёта порядка, т. е. C_n^{κ}. Согласно определению вероятности,

$$P_n(\kappa) = P(B) = p^{\kappa} \cdot q^{n-\kappa} + p^{\kappa} \cdot q^{n-\kappa} + \ldots + p^{\kappa} \cdot q^{n-\kappa} = C_n^{\kappa} p^{\kappa} q^{n-\kappa}, \quad (2.12)$$

где $q = 1 - p$.

Формулу (2.12) называют *формулой Бернулли* [8].

2. Локальная теорема Муавра – Лапласа

Пусть в схеме Бернулли $p \ne 0;1$, тогда $\dfrac{\sqrt{npq}}{\varphi(x)} P_n(\kappa) \to 1$ при $n \to \infty$, где $x = \dfrac{\kappa - np}{\sqrt{npq}}$; $\varphi(x) = \dfrac{1}{\sqrt{2\pi}} \ell^{-\frac{x^2}{2}}$. Следовательно, при больших n

$$P_n(\kappa) = \frac{1}{\sqrt{npq}} \varphi(x). \quad (2.13)$$

Для значений функции $\varphi(x)$ составлено прил. 1.

3. Интегральная теорема Муавра – Лапласа

Пусть в схеме Бернулли κ - число успехов в n испытаниях и $P_n(\kappa_1;\kappa_2) = P_n(\kappa_1 < \kappa < \kappa_2)$.

Тогда при больших n

$$P_n(\kappa_1;\kappa_2) = \frac{1}{\sqrt{2\pi}} \int\limits_{x_1}^{x_2} \ell^{\frac{-x^2}{2}} dx, \text{ где } x_1 = \frac{\kappa_1 - np}{\sqrt{npq}},\ x_2 = \frac{\kappa_2 - np}{\sqrt{npq}}.$$

Если обозначить $\Phi(x) = \dfrac{1}{\sqrt{2\pi}} \int\limits_0^x \ell^{\frac{-x^2}{2}} dx,$ то получаем формулу для вычислений

$$P_n(\kappa_1;\kappa_2) = \frac{1}{\sqrt{2\pi}} \int\limits_0^{x_2} \ell^{\frac{-x^2}{2}} dx - \frac{1}{\sqrt{2\pi}} \int\limits_0^{x_1} \ell^{\frac{-x^2}{2}} dx = \Phi(x_2) - \Phi(x_1). \quad (2.14)$$

Для значений функции $\Phi(x)$, соответствующих значениям аргумента $x \in [0;5]$, имеется прил. 2. Для отрицательных x значения $\Phi(x)$ можно получить, воспользовавшись нечётностью $(\Phi(-x) = -\Phi(x))$ этой функции, а при $x > 5$ можно считать $\Phi(x) = 0{,}5$, т. к. $\Phi(5) = 499997$; $\Phi(x) \xrightarrow[x \to \infty]{} 0{,}5$ и $\Phi(x)$ – функция возрастающая.

4. Теорема Пуассона

Если n достаточно велико, а p мало, то

$$P_n(\kappa) = \frac{\lambda^\kappa}{\kappa!}\ell^{-\lambda}, \qquad \text{где } \lambda = np. \qquad (2.15)$$

Задача 1. В урне 20 шаров: 15 белых и 5 чёрных. Вынули подряд 5 шаров, причём каждый вынутый шар возвращали в урну, и перед извлечением следующего шары в урне тщательно перемешивались. Найти вероятность того, что из пяти вынутых шаров будет два белых.

Решение. Вероятность появления белого шара в каждом испытании $p = 15/20$, а вероятность не появления белого шара $q = 1 - p = 1/4$. По формуле Бернулли (2.12) находим

$$P_5(2) = C_5^2 p^2 q^{5-2} = \frac{5!}{2!\cdot(5-2)!}\cdot\left(\frac{3}{4}\right)^2\cdot\left(\frac{1}{4}\right)^3 = \frac{5!}{2!\cdot 3!}\cdot\frac{3^2}{4^4} =$$

$$= \frac{1\cdot 2\cdot 3\cdot 4\cdot 5}{1\cdot 2\cdot 1\cdot 23}\cdot\frac{3^2}{4^4} = \frac{45}{512}.$$

Задача 2. Найти вероятность того, что из 100 независимых выстрелов будет 75 попаданий, если вероятность попадания при одном выстреле равна 0,8.

Решение. Очевидно, мы находимся в рамках схемы Бернулли:

$n = 100$; $\kappa = 75$; $p = 0{,}8$; $q = 0{,}2$; n – достаточно велико, воспользуемся формулой (2.13):

$$\sqrt{npq} = \sqrt{100\cdot 0{,}8\cdot 0{,}2} = 4, \quad x = \frac{\kappa - np}{\sqrt{npq}} = \frac{75 - 100\cdot 0{,}8}{4} = -1{,}25.$$

По прил. 1 находим $\varphi(1{,}25) = 0{,}1826$, тогда $P_{100}(75) = \frac{1}{4}\cdot 0{,}1826 \approx 0{,}0456$.

Задача 3. Вероятность появления события при одном опыте равна 0,3. С какой вероятностью можно утверждать, что частота этого события при 100 опытах будет лежать в пределах от 0,2 до 0,4?

Решение. Для того, чтобы частота лежала в пределах от 0,2 до 0,4 в серии из 100 опытов, число появлений события m должно быть не менее 20 и не более 40 $(20 < m < 40)$.

Воспользуемся интегральной теоремой Муавра-Лапласа, формулой (2.14)

$$P(20 \le m \le 40) = \Phi\left(\frac{40 - n\cdot p}{\sqrt{n\cdot p\cdot q}}\right) - \Phi\left(\frac{20 - n\cdot p}{\sqrt{n\cdot p\cdot q}}\right),$$

где $n = 100$, $p = 0{,}3$, $q = 0{,}7$. Следовательно,

$$\Phi\left(\frac{40 - n\cdot p}{\sqrt{n\cdot p\cdot q}}\right) = \Phi\left(\frac{40 - 30}{\sqrt{100\cdot 0{,}3\cdot 0{,}7}}\right) = \Phi(2{,}18) = 0{,}4854,$$

где значение $\Phi(2{,}18)$ найдено по прил. 2.

$$\Phi\left(\frac{20 - n\cdot p}{\sqrt{n\cdot p\cdot q}}\right) = \Phi\left(\frac{20 - 30}{\sqrt{100\cdot 0{,}3\cdot 0{,}7}}\right) = \Phi(-2{,}18) = -\Phi(2{,}18) = -0{,}4854.$$

Следовательно, $P(20 \le m \le 40) = 0{,}4854 + 0{,}4854 = 0{,}97$.

Задача 4. Аппаратура содержит 2000 элементов, вероятность отказа каждой из них $p = 0{,}0005$. Какова вероятность отказа трех элементов, если отказы происходят независимо друг от друга?

Решение. p – мало. Воспользуемся теоремой Пуассона

$$\lambda = 2000 \cdot 0{,}0005 = 1;\; P_{2000}(3) = \frac{\lambda^3}{3!}\ell^{-\lambda} = \frac{1}{2 \cdot 3}\ell^{-1} \approx 0{,}06.$$

Задача 5. Испытанию подвергается партия транзисторов. Вероятность безотказной работы каждого транзистора равна 0,92. Определить, какое число транзисторов следует испытать, чтобы с вероятностью не менее 0,95 можно было зафиксировать хотя бы один отказ.

Решение. Обозначим количество испытуемых транзисторов через n. Тогда вероятность их безотказной работы равна $0{,}92^n$. События “все транзисторы работают безотказно” и “хотя бы один транзистор не работает” образуют полную группу событий. Значит, вероятность события “хотя бы один отказ” равна $1 - 0{,}92^n$. По условию задачи эта величина больше 0,95, т.е.

$$1 - 0.92^n \ge 0{,}95; -0{,}92^n \ge -0{,}5; 0{,}92^n \le 0{,}5; \ln 0{,}92^n \ge \ln 0{,}5;$$

$$n \ln 0{,}92 \ge \ln 0{,}5; n \ge \frac{\ln 0{,}5}{\ln 0{,}92} \approx 35{,}9.$$

Следовательно, $n \ge 36$.

В заключение рассмотрим задачу, иллюстрирующую все три формулы.

Задача 6. Работница обслуживает 800 веретен. Вероятность обрыва пряжи на каждом из них за промежуток времени t равна 0,005. Найти наиболее вероятное число обрывов и его вероятность.

Решение. Наиболее вероятное число обрывов будет $\lambda=np=4$. Точное значение вероятности четырех обрывов равно [см. формулу (2.12)]

$$P_{800}(4) = C_{800}^4 \cdot 0{,}005^4 \cdot (0{,}995)^{796} = 0{,}1945.$$

Пользуясь формулой Пуассона с $\lambda = np = 4$, получаем [см. формулу (2.15)]

$$P_{800}(4) \approx \frac{4^4}{4!} \cdot e^{-4} = \frac{256}{24} \cdot 0{,}0183 = 0{,}1954.$$

Вычисление по точной формуле дает 0,1945, так что ошибка при пользовании формулой Пуассона составляет 0,0009. Локальная предельная теорема Муавра–Лапласа дает для данного случая [см. формулу (2.13)]

$$P_{800}(4) \approx \frac{1}{\sqrt{2\pi \cdot 800 \cdot 0{,}005 \cdot 0{,}995}} \cdot e^{-\frac{x^2}{2}} \approx 0{,}2000,$$

ибо здесь $x = \dfrac{\kappa - np}{\sqrt{npq}} = \dfrac{4 - 800 \cdot 0{,}005}{\sqrt{800 \cdot 0{,}005 \cdot 0{,}995}} = 0$; $e^0=1$, так что ошибка составляет уже 0,0055, т. е. в шесть раз больше, чем при использовании формулы Пуассона, т. к. $np=4<10$.