ФОРМУЛЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ

Повторение испытаний

Если производятся испытания, в каждом из которых вероятность появления события A не зависит от исходов других испытаний, то такие испытания называются *независимыми относительно события* A.

Пусть производятся n независимых испытаний и пусть вероятность появления события A во всех испытаниях одинакова и равна p (следовательно, вероятность появления противоположного события $\overline{A}$ равна $q = 1 - p$). Тогда вероятность того, что событие A появится ровно k раз, находится по формуле:

$$p_n(k) = C_n^k \cdot p^k \cdot q^{n-k}.$$

Эта формула называется *формулой Бернулли*.

Число k называется *числом успехов*.

Вероятности $p_n(k)$ называются *биномиальными вероятностями* и удовлетворяют условию: $p_n(0) + p_n(1) + p_n(2) + \ldots + p_n(n) = 1$.

Пример

Найти вероятность того, что при семи бросаниях монеты герб выпадет три раза.

Так как производятся независимые испытания (бросания монеты) и вероятность выпадения герба во всех испытаниях одинакова (она равна $\frac{1}{2}$), то применима формула Бернулли.

В данном случае $n = 7$, $k = 3$, $p = \frac{1}{2}$, $q = 1 - p = 1 - \frac{1}{2} = \frac{1}{2}$.

Значит, искомая вероятность

$$p_7(3) = C_7^3 \cdot \left(\frac{1}{2}\right)^3 \cdot \left(\frac{1}{2}\right)^{7-3} = \frac{7!}{3! \cdot 4!} \cdot \frac{1}{8} \cdot \frac{1}{16} = \frac{35}{128}.$$

Еще пример

Число k_0, которому при заданном значении n соответствует максимальная биномиальная вероятность $p_n(k_0)$, называется *наиболее вероятным числом успехов*. Число k_0 удовлетворяет двойному неравенству:

$$(n+1)\cdot p-1\le k_0\le n\cdot p+p\,.$$

Найти наиболее вероятное число выпадений герба при 37 бросаниях монеты.

По условию, $n=37$, $p=\frac{1}{2}$. Тогда $(37+1)\cdot\frac{1}{2}-1\le k_0\le 37\cdot\frac{1}{2}+\frac{1}{2}$. Следовательно, $18\le k_0\le 19$. Полученному неравенству удовлетворяют два целых числа: $k_0'=18$, $k_0''=19$. Это означает, что наиболее вероятное число выпадений герба равно 18 или 19.

Новый пример

Пример . Вероятность попадания стрелка в мишень при одном выстреле равна 0,7 и не зависит от порядкового номера выстрела. Найти вероятность того, что при 6 выстрелах произойдет: а) не более двух промахов; б) хотя бы один промах.

а) По условию задачи успехом является не попадание, а промах стрелка, поэтому $p = 0{,}3$. Следовательно, $q = 1 - 0{,}3 = 0{,}7$.

Его продолжение

Событие «произошло не более двух промахов» является суммой трех несовместных событий: «не произошло ни одного промаха (произошло 0 промахов)», «произошел один промах», «произошло два промаха». Значит, по теореме сложения вероятностей искомая вероятность $p_6(0 \le k \le 2) = p_6(0) + p_6(1) + p_6(2)$.

Продолжение продолжается

Вычислим вероятности $p_6(0), p_6(1), p_6(2)$:

$$p_6(0) = C_6^0 \cdot 0{,}3^0 \cdot 0{,}7^{6-0} = \frac{6!}{0! \cdot 6!} \cdot 0{,}3^0 \cdot 0{,}7^6 = 1 \cdot 1 \cdot 0{,}117649 = 0{,}117649;$$

$$p_6(1) = C_6^1 \cdot 0{,}3^1 \cdot 0{,}7^{6-1} = \frac{6!}{1! \cdot 5!} \cdot 0{,}3^1 \cdot 0{,}7^5 = 6 \cdot 0{,}3 \cdot 0{,}16807 = 0{,}302526;$$

$$p_6(2) = C_6^2 \cdot 0{,}3^2 \cdot 0{,}7^{6-2} = \frac{6!}{2! \cdot 4!} \cdot 0{,}3^2 \cdot 0{,}7^4 = 15 \cdot 0{,}09 \cdot 0{,}2401 = 0{,}324135.$$

Окончательно получаем

$$p_6(0 \le k \le 2) = 0{,}117649 + 0{,}302526 + 0{,}324135 = 0{,}74431.$$

Еще чуть-чуть потерпите

б) События «произошел хотя бы один промах» $(1 \le k \le 6)$ и «не произошло ни одного промаха» $(k = 0)$ являются противоположными, значит

$$p_6(1 \le k \le 6) = 1 - p_6(0) = 1 - 0{,}117649 = 0{,}882351.$$

Формулы Лапласа

Очевидно, что при больших значениях n и k вычисление вероятностей $p_n(k)$ по формуле Бернулли превращается в технически сложную задачу.

Если n достаточно велико, то вероятность того, что событие A появится ровно k раз в n независимых испытаниях, находится по формуле:

$$p_n(k) \approx \frac{1}{\sqrt{npq}} \cdot \varphi(x), \text{ где } \varphi(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}, \; x = \frac{k-np}{\sqrt{npq}}.$$

Эта формула называется *локальной формулой Лапласа*.

Формулы Лапласа

Заметим, что использование локальной формулы Лапласа в случае, когда вероятность p близка к 0 или 1, обычно приводит к значительным погрешностям. Для того чтобы в этом случае формула дала результат с незначительной ошибкой, необходимо, чтобы число испытаний n было очень велико. На практике локальной формулой Лапласа пользуются при выполнении условия: $npq \geq 10$.

Формулы Лапласа

Вероятность того, что в n независимых испытаниях событие A появится не менее k_1 раз и не более k_2 раз, находится по формуле (при условии, что n достаточно велико):

$$p_n(k_1;k_2) = \Phi(x_2) - \Phi(x_1), \text{ где}$$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}}dt \text{ – функция Лапласа, } x_1 = \frac{k_1 - np}{\sqrt{npq}},\ x_2 = \frac{k_2 - np}{\sqrt{npq}}.$$

Эта формула называется *интегральной формулой Лапласа*.

Пример

Для данного сорта растения всхожими являются 75% семян. Найти вероятность того, что среди 1000 высеянных семян взойдет: а) 770 семян; б) от 720 до 770 семян.

По условию задачи $n = 1000$; $p = 0,75$; $q = 0,25$. Так как $npq = 1000 \cdot 0,75 \cdot 0,25 = 187,5 > 10$, то применимы формулы Лапласа.

а) Итак, $n = 1000$; $k = 770$; $p = 0,75$; $q = 0,25$.

Применим локальную формулу Лапласа.

Найдем значение x:

$$x = \frac{k - np}{\sqrt{npq}} = \frac{770 - 1000 \cdot 0,75}{\sqrt{1000 \cdot 0,75 \cdot 0,25}} \approx \frac{20}{13,69} \approx 1,46.$$

По таблице значений функции $\varphi(x)$ найдем $\varphi(1,46) = 0,1374$.

Тогда вероятность

$$p_{1000}(770) = \frac{1}{\sqrt{npq}} \cdot \varphi(x) = \frac{1}{\sqrt{1000 \cdot 0,75 \cdot 0,25}} \cdot 0,1374 \approx 0,01.$$

Продолжение примера

б) По условию $n = 1000$; $k_1 = 720$; $k_2 = 770$; $p = 0,75$, $q = 0,25$.

Применим интегральную формулу Лапласа.

Найдем значения x_1 и x_2:

$$x_1 = \frac{720 - 1000 \cdot 0,75}{\sqrt{1000 \cdot 0,75 \cdot 0,25}} \approx -2,19; \quad x_2 = \frac{770 - 1000 \cdot 0,75}{\sqrt{1000 \cdot 0,75 \cdot 0,25}} \approx 1,46.$$

По таблице значений функции $\Phi(x)$ найдем $\Phi(x_2)$ и $\Phi(x_1)$:

$\Phi(x_2) = \Phi(1,46) = 0,4279$.

Так как функция $\Phi(x)$ является нечетной, то $\Phi(x_1) = \Phi(-2,19) = -\Phi(2,19) = -0,4857$.

Тогда вероятность того, что взойдет от 720 до 770 семян, будет равна $p_{1000}(720;770) = \Phi(x_2) - \Phi(x_1) = 0,4279 - (-0,4857) = 0,9136$.

Формула Пуассона и пример

Если n достаточно велико, а вероятность p появления события A в каждом испытании мала, то вероятность того, что событие A появится ровно k раз в n независимых испытаниях, находится по формуле:

$p_n(k) \approx \frac{\lambda^k}{k!} \cdot e^{-\lambda}$, где $\lambda = np$ – постоянная величина.

Эта формула называется *формулой Пуассона*.

Обычно формулу Пуассона используют, когда $n \geq 100$, а $p \leq 0{,}1$.

Пример В организации 300 человек, использующих в своей работе официальный сайт организации. Сотрудники организации работают независимо друг от друга. Для каждого сотрудника вероятность того, что он в течение часа обратится к сайту, равна 0,01. Найти вероятность того, в течение часа к официальному сайту организации обратятся: а) 2 сотрудника; б) не менее 3 сотрудников.

Продолжение примера

а) По условию, $n = 300$, $p = 0{,}01$, $k = 2$. Сотрудники обращаются к сайту независимо друг от друга, число n достаточно велико, а вероятность p мала, поэтому можно применить формулу Пуассона.

В нашем случае $\lambda = np = 300 \cdot 0{,}1 = 3$.

Искомая вероятность $p_{300}(2) = \dfrac{3^2}{2!} \cdot e^{-3} \approx \dfrac{9}{2} \cdot 0{,}0498 = 0{,}2241$.

б) События «в течение часа к сайту обратятся не менее 3 сотрудников» и «в течение часа к сайту обратятся менее 3 сотрудников» - противоположные, следовательно

$$p_{300}(3 \le k \le 300) = 1 - p_{300}(0 \le k < 3) = 1 - \left(p_{300}(0) + p_{300}(1) + p_{300}(2)\right) =$$

$$= 1 - \left(\frac{3^0}{0!} \cdot e^{-3} + \frac{3^1}{1!} \cdot e^{-3} + \frac{3^2}{2!} \cdot e^{-3}\right) \approx 1 - (0{,}0498 + 0{,}1494 + 0{,}2241) = 0{,}5767.$$