3 Последовательность испытаний. Схема Бернулли.

Пусть производится **n** независимых испытаний, в каждом из которых событие А может появиться либо не появиться (то есть появится событие $\overline{A}$). При этом вероятность события А в каждом испытании одна и та же и равна **p**. Требуется вычислить вероятность того, что при **n** испытаниях событие А осуществится ровно **k** раз.

Формула Бернулли. Вероятность того, что в n независимых испытаниях, в каждом из которых вероятность появления события равна p, событие А наступит ровно k раз (безразлично в какой последовательности) равна:

$$P_n^k(k) = C_n^k \cdot p^k \cdot q^{n-k}, \text{ где } q = 1 - p.$$

Формула Пуассона. В том случае, когда вероятность появления события p мала ($p < 0{,}1$), а число независимых испытаний велико, для оценки вероятности появления события ровно k раз в n независимых испытаниях используется асимптотическая формула Пуассона:

$$P_n^k(k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad \text{где } \lambda = n \cdot p.$$

Значения $P_n^k(k)$ при фиксированных k и λ можно найти с помощью таблицы 1 приложения 2.

Локальная теорема Муавра-Лапласа. Если число независимых испытаний велико, а вероятность появления события А в каждом из них постоянна и равна p $(0 < p < 1)$, то вероятность того, что в n независимых испытаниях, событие наступит ровно k раз (безразлично в какой последовательности), приближенно равна (тем точнее, чем больше n):

$$P_n^k(k) = \frac{1}{\sqrt{npq}} \varphi(x),$$

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

Значения функции $\varphi(x)$ можно найти с помощью таблицы 2 приложения 2. Данная формула, в отличие от формулы Бернулли, используется при больших n и k.

Интегральная теорема Муавра-Лапласа. Если число независимых испытаний велико, а вероятность появления события А в каждом из них постоянна и равна p $(0 < p < 1)$, то вероятность того, что в n независимых испытаниях событие наступит не менее k_1 раз и не более k_2 раз, приближенно равна:

$$P_n^k(k_1 \le k \le k_2) \approx \Phi(x_2) - \Phi(x_1)$$

где $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{\frac{-t^2}{2}} dt$ – интегральная функция Лапласа (см. прил. 2, табл. 3);

$$x_1 = \frac{k_1 - np}{\sqrt{npq}}; \; x_2 = \frac{k_2 - np}{\sqrt{npq}}.$$

Оценка отклонения относительной частоты от постоянной вероятности.

Вероятность того, что в n независимых испытаниях, в каждом из которых вероятность появления события А постоянна и равна p $(0<p<1)$, абсолютная величина отклонения относительной частоты появления события А от его постоянной вероятности не превысит положительного числа ε, приближенно равна:

$$P\left(\left|\frac{m}{n} - p\right| \le \varepsilon\right) \approx 2\Phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right).$$

Пример 5.1. Игральная кость брошена четыре раза. Найти вероятность того, что шестерка появится не более двух раз.

Решение.

Испытанием является бросание игральной кости.

Событие А (успех) – выпадение шестерки.

Событие $\overline{A}$ (неудача) – выпадение любого числа очков, кроме шести.

Параметры **p** и **q** найдем по классическому определению вероятности:

$$P(A) = \frac{1}{6} = p, \quad P(\overline{A}) = \frac{5}{6} = q.$$

Число произведенных испытаний **n**=4.

Тогда число успехов **k** может принимать значения 0, 1, 2, 3, 4. Среди этих вариантов нам подходит, согласно условию, 0, 1 и 2.

Так как k находится в пределах от 0 до 2, то обозначим искомую вероятность как:

$$P_4(0 \le k \le 2).$$

Так как события B_0, B_1 и B_2, соответствующие случаям $k=0$, $k=1$ и $k=2$, являются несовместными, найдем искомую вероятность как сумму вероятностей $P_4(0)$, $P_4(1)$ и $P_4(2)$. Каждое слагаемое, в свою очередь, найдем по формуле Бернулли, так как n=4 очевидно мало.

$$P_4(0) = C_4^0 p^0 q^4 = \left(\frac{1}{6}\right)^0 \left(\frac{5}{6}\right)^4 = \frac{5^4}{6^4} = \frac{625}{1296} = P(B_0).$$

Аналогично рассуждая, получим:

$$P_4(1) = C_4^1 p^1 q^3 = 4 \cdot \left(\frac{1}{6}\right)^1 \left(\frac{5}{6}\right)^3 = 4 \cdot \frac{5^3}{6^4} = \frac{500}{1296} = P(B_1),$$

$$P_4(2) = C_4^2 p^2 q^2 = 6 \cdot \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^2 = 6 \cdot \frac{5^2}{6^4} = \frac{150}{1296} = P(B_2).$$

Используя теорему сложения несовместных событий, получим:

$$P_4(0 \le k \le 2) = P(B_0 + B_1 + B_2) = P_4(0) + P_4(1) + P_4(2) ==$$

$$= \frac{625}{1296} + \frac{500}{1296} + \frac{150}{1296} = \frac{1275}{1296} \approx 0{,}98379.$$

Пример 5.2. Вероятность обрыва нити на ткацком станке в течение часа равна 0,01. Найти вероятность того, что в течение часа на станке оборвется не менее двух нитей из 100.

Решение.

Испытанием является наблюдение за нитью.

Событие A – нить оборвалась.

Событие $\overline{A}$ – нить не оборвалась.

Параметры **p** и **q** берем из условия:

$$P(A) = 0{,}01 = p;\ P(\overline{A}) = 1 - 0{,}01 = 0{,}99 = q.$$

Число произведенных испытаний **n**=100 достаточно велико. Можно считать, что оно стремится к бесконечности. В свою очередь p=0,01, достаточно мало. Можно считать, что оно стремится к нулю. Данные условия соответствуют теореме Пуассона.

Число успехов **k** может принимать значения 0, 1, 2, …, 100. Среди этих вариантов нам подходит, согласно условию, 2, 3, …, 100.

Так как k находится в пределах от 2 до 100, то обозначим искомую вероятность как:

$$P_{100}(2 \le k \le 100).$$

Так как промежуток от 2 до 100 содержит большое количество вариантов, и событие $0 \le k \le 1$ является противоположным к искомому, воспользуемся подходом «от обратного».

$$P_{100}(2 \le k \le 100) = 1 - P_{100}(0 \le k \le 1) = 1 - P_{100}(0) - P_{100}(1).$$

Для $\lambda = np = 100 \cdot 0{,}01 = 1$ по таблице 1 приложения 1 найдем:

$P_{100}(0) = 0{,}3676$; $P_{100}(1) = 0{,}3679$.

Тогда искомая вероятность примет вид:

$$P_{100}(2 \le k \le 100) = 1 - P_{100}(0) - P_{100}(1) = 1 - 0{,}3676 - 0{,}3679 = 0{,}2445.$$

Замечание:

При использовании формулы Пуассона параметр q находить не надо.

Пример 5.3. Вероятность попадания мячом в корзину для данного баскетболиста равна 0,8. Найти вероятность того, что при 100 бросках попаданий в корзину будет: а) ровно 79; б) от 75 до 84.

Решение:

Испытанием является бросок в корзину.

Событие A – попадание.

Событие $\overline{A}$ – промах.

Параметры **p** и **q** берем из условия:

$P(A) = 0{,}8 = p$; $P(\overline{A}) = 1 - 0{,}8 = 0{,}2 = q$.

Число произведенных испытаний **n**=100 достаточно велико. Можно считать, что оно стремится к бесконечности. В свою очередь p=0,8, маленьким назвать нельзя. Данные условия соответствуют как локальной так и интегральной теоремам Муавра-Лапласа.

а) Число успехов **k** может принимать значения 0, 1, 2, …, 100. Среди этих вариантов нам подходит, согласно условию пункта а) k=79.

Так как k равно конкретному числу, применим локальную теорему Муавра Лапласа для нахождения $P_{100}(79)$.

Найдем $x = \dfrac{k - np}{\sqrt{npq}} = \dfrac{79 - 100 \cdot 0{,}8}{\sqrt{100 \cdot 0{,}8 \cdot 0{,}2}} = \dfrac{-1}{4} = -0{,}25.$

По табл. 2 найдем значение $\varphi(0{,}25) = 0{,}3867$.. Так как локальная функция Лапласа – четная, то $\varphi(-0{,}25) = \varphi(0{,}25) = 0{,}3867$.

Тогда искомая вероятность примет вид:

$$P_{100}(79)=\frac{\varphi(-0,25)}{\sqrt{npq}}=\frac{0,3867}{4}=0,0967.$$

б) Число попаданий от 75 до 84 означает, что надо найти $P_{100}(75\le k\le 84)$. Воспользуемся интегральной теоремой Муавра-Лапласа. Найдем значения аргументов:

$$x_1=\frac{k_1-np}{\sqrt{npq}}=\frac{75-80}{4}=\frac{-5}{4}=-1,25;$$

$$x_2=\frac{k_2-np}{\sqrt{npq}}=\frac{84-80}{4}=\frac{6}{4}=1,5$$

По табл. 3 найдем значения интегральной функции Лапласа. $\Phi(-1,25)=-\Phi(1,25)=-0,3944$ и $\Phi(1,5)=0,4332$. Тогда искомая вероятность примет вид:

$$P_{100}(75\le k\le 84)=\Phi(x_2)-\Phi(x_1)=\Phi(1,5)-\Phi(-1,25)=$$
$$=0,4332+0,3944=0,8278.$$

Пример 5.4. Вероятность появления события в каждом из 625 независимых испытаний равна 0,8. Найти вероятность того, что относительная частота появления события отклонится от его вероятности по абсолютной величине не более, чем на 0,04.

Решение. По условию задачи задано, что: n=625; p=0,8; ε=0,04.

Отсюда q =1– p = 0,2. Требуется найти вероятность:

$$P\left(\left|\frac{m}{625}-0,8\right|\le 0,04\right)=?$$

Для решения указанной задачи воспользуемся формулой, определяющей оценку отклонения относительной частоты от постоянной вероятности, т. е.:

$$P\left(\left|\frac{m}{n}-p\right|\le \varepsilon\right)=2\Phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right),$$

Φ(x) – интегральная функция Лапласа. Найдем аргумент функции Лапласа:

$$x=\varepsilon\sqrt{\frac{n}{pq}}=0,04\sqrt{\frac{625}{0,8\cdot 0,2}}=2,5$$

По таблице 3 приложения для функции Лапласа найдем, что Φ(2,5)=0,4983, т.е. 2Φ(x)=0,9876. Итак, искомая вероятность:

$$P\left(\left|\frac{m}{625}-0,8\right|\le 0,04\right)\approx 0,9876.$$

4. Предельные теоремы в схеме Бернулли

Если число испытаний *П* Велико, то вычисления по формуле Бернулли становятся затруднительными. В то же время большие значения *П* Позволяют заменять эту формулу приближенными асимптотическими формулами. Рассмотрим три такие формулы.

Теорема 2.5. (*Формула Пуассона*) Если $n \to \infty,\ p \to 0$ так, что $np \to \lambda$, то

$$P_n(m) \cong P_m = \frac{\lambda^m}{m!} \cdot e^{-\lambda}. \quad (2.5)$$

Формула (2.5) дает хорошие результаты, если *Npq*<9. Если же *Npq*>9, то для вычисления вероятности $P_n(m)$ можно воспользоваться локальной теоремой Лапласа.

Теорема 2.6. (*Локальная теорема Муавра–Лапласа*). Вероятность появления события *Т* раз в *П* независимых испытаниях при больших значениях *П* приближенно определяется по формуле

$$P_n(m) \cong \frac{\varphi(x)}{\sqrt{npq}}, \quad (2.6)$$

Где $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad x = \frac{m - np}{\sqrt{npq}}.$

Теорема 2.7. (*Интегральная теорема Муавра–Лапласа*). Вероятность того, что число появлений события в *П* независимых испытаниях находится в пределах *Т*1£ *m* £ *m*2 и при больших значениях *П* приближенно определяется по формуле

$$P_n(m_1, m_2) = \Phi(x_2) - \Phi(x_1), \quad (2.7)$$

Где

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt, \quad x_1 = \frac{m_1 - np}{\sqrt{npq}}, \quad x_2 = \frac{m_2 - np}{\sqrt{npq}}.$$

Функция *Ф*(*X*) называется *Функцией Лапласа*. Для функций $\varphi(x),\ \Phi(x)$ имеются таблицы ее значений. Функция $\varphi(x)$ является четной, а функция Ф(*X*) – нечетной, т. е. $\varphi(-x) = \varphi(x)$; Ф(– *X*)= – Ф(*X*);

Из интегральной теоремы Лапласа можно вывести формулу для вероятности отклонения относительной частоты *Т/n* События в серии испытаний от постоянной вероятности *Р* этого события в одном испытании:

$$\mathbf{P}\left\{\left|\frac{m}{n}-p\right|\le\varepsilon\right\}=2\Phi\left(\frac{\varepsilon\sqrt{n}}{\sqrt{pq}}\right) \quad (2.8)$$

Пример 2.21. Завод отправил на базу 500 изделий. Вероятность повреждения изделия в пути равна 0.002. Найти вероятность того, что в пути будут повреждены три изделия.

Решение. Можно считать, что имеем дело со схемой Бернулли, в которой испытания проводятся 500 раз. Так как число *П*=500 достаточно велико, а вероятность *P*=0.002 мала (причем *Npq*=500×0.002×0.998»2<9), то воспользуемся приближенной формулой (2.5), где *L*=*Np* =500×0.002=1:

$$P_{500}(3)\cong\frac{1^3}{3!}\cdot e^{-1}=0.0613.$$

Пример 2.22. Найти вероятность того, что событие происходит 80 раз в 400–х испытаниях, если вероятность этого события в каждом испытании равна 0.2.

Решение. Здесь *П*=400 достаточно велико, но величина *Npq* также велика (*Npq*=400×0.2×0.8=64>9), поэтому воспользуемся формулой (2.6). Вычисляем

$$x=\frac{m-np}{\sqrt{npq}}=\frac{80-400\cdot0.2}{\sqrt{400\cdot0.2\cdot0.8}}=0.$$

По таблице функции $\varphi(x)$ находим φ(0)=0.3989. Окончательно получаем:

$$P_{400}(80)\cong\frac{\varphi(0)}{\sqrt{400\cdot0.2\cdot0.8}}=\frac{0.3989}{8}=0.0499.$$

Пример 2.23. Найти вероятность того, что в 400–х испытаниях событие произойдет не более 70–ти раз, если вероятность появления этого события в каждом испытании равна 0.2.

Решение. Воспользуемся интегральной теоремой Лапласа для вычисления вероятности $P_{400}(0;\ 70)$:

$$x_1=\frac{0-400\cdot0.2}{\sqrt{400\cdot0.2\cdot0.8}}=-10,\quad x_2=\frac{70-400\cdot0.2}{\sqrt{400\cdot0.2\cdot0.8}}=-1.25.$$

$$P_{400}(0;\ 70)\cong\Phi(-1.25)-\Phi(-10)=\Phi(10)-\Phi(1.25)=$$
$$=0.500-0.394=0.106.$$

Пример 2.24. Определим, сколько надо провести испытаний, чтобы с вероятностью 0.95 относительная частота выпадения «орла» отличалась от вероятности *P*=0.5 этого события не более чем на 5%.

Решение. Воспользуемся формулой (2.8). В нашем случае *P*=0.5, *Q*=0.5, е=0.5 0.05=0.025. По условию задачи

$$P\left\{\left|\frac{m}{n}-p\right|\le\varepsilon\right\}=2\Phi\left(\frac{\varepsilon\sqrt{n}}{\sqrt{pq}}\right)=0.95$$

Или $\Phi\left(\varepsilon\sqrt{n}/\sqrt{pq}\right)=0.475.$ Пользуясь таблицей функции Лапласа, по значению функции находим значение аргумента:

$\left(\varepsilon\sqrt{n}/\sqrt{pq}\right)=1.96,$ т. е. $0.025\cdot\sqrt{n}/\sqrt{0.5\cdot0.5}=1.96.$

Отсюда находим, что *П*=1536.64. Таким образом, надо провести не менее чем 1537 испытаний.