

1.20. Формула Бернулли

Простейшая задача, относящаяся к схеме Бернулли, состоит в определении вероятности того, что в n независимых испытаниях событие A наступит m раз ($0 \leq m \leq n$). Обозначается искомая вероятность так: $P_n(m)$ или $P_{n,m}$ или $P(\mu_n = m)$, где μ_n — число появления события A в серии из n опытов.

Например, при бросании игральной кости 3 раза $P_3(2)$ означает вероятность того, что в 3-х опытах событие A — выпадение цифры 4 — произойдет 2 раза. Очевидно,

$$P_3(2) = p^2q + p^2q + p^2q = \\ = \left[(A, A, \bar{A}); (A, \bar{A}, A); (\bar{A}, A, A) \right] = 3p^2q = 3 \cdot \left(\frac{1}{6}\right)^2 \cdot \frac{5}{6} = \frac{5}{72} = 0,069.$$

Теорема 1.4. Если производится n независимых испытаний, в каждом из которых вероятность появления события A равна p , а вероятность его не появления равна $q = 1 - p$, то вероятность того, что событие A произойдет m раз определяется *формулой Бернулли*

$$P_n(m) = C_n^m \cdot p^m \cdot q^{n-m}, \quad m = 0, 1, 2, \dots, n. \quad (1.32)$$

□ Вероятность одного сложного события, состоящего в том, что событие A в n независимых опытах появится m раз в первых m опытах и не появится ($n - m$) раз в остальных опытах (это событие $\underbrace{A \cdot A \cdot A \cdot \dots \cdot A}_{m \text{ раз}} \cdot \underbrace{\bar{A} \cdot \bar{A} \cdot \dots \cdot \bar{A}}_{(n-m) \text{ раз}}$) по теореме умножения вероятностей равна $p^m q^{n-m}$. Вероятность появления события A снова m раз, но в другом

порядке (например, $\bar{A} \cdot \underbrace{A \cdot A \cdot \dots \cdot A}_{m \text{ раз}} \cdot \bar{A} \cdot \bar{A} \cdot \dots \cdot \bar{A}$ или $\bar{A} \bar{A} \bar{A} \dots \bar{A} \bar{A}$ и т. д.) будет той же самой, т. е. $p^m q^{n-m}$.

Число таких сложных событий — в n опытах m раз встречается событие A в различном порядке — равно числу сочетаний из n по m , т. е. C_n^m . Так как все эти сложные события несовместны, то по теореме сложения вероятностей искомая вероятность равна сумме вероятностей всех возможных сложных событий, т. е.

$$P_n(m) = \underbrace{p^m q^{n-m} + \dots + p^m q^{n-m}}_{C_n^m \text{ слагаемых}} = C_n^m p^m q^{n-m}, \quad m = 0, 1, \dots, n. \quad \blacksquare$$

Можно заметить, что вероятности $P_n(m)$, $m = 0, 1, \dots, n$ являются коэффициентами при x^m в разложении $(q + px)^n$ по формуле бинома Ньютона:

$$(q + px)^n = q^n + C_n^1 q^{n-1} px + C_n^2 q^{n-2} p^2 x^2 + \dots + C_n^m q^{n-m} p^m x^m + \dots + p^n x^n.$$

Поэтому совокупность вероятностей $P_n(m)$ называют *биномиальным законом распределения вероятностей* (см. п. 2.7), а функцию $\varphi(x) = (q + px)^n$ — *производящей функцией* для последовательности независимых опытов.

Если в каждом из независимых испытаний вероятности наступления события A разные, то вероятность того, что событие A наступит m раз в n опытах, равна коэффициенту при m -й степени многочлена $\varphi_n(z) = (q_1 + p_1x)(q_2 + p_2x) \cdot \dots \cdot (q_n + p_nz)$, где $\varphi_n(z)$ — производящая функция.

Если в серии из n независимых опытов, в каждом из которых может произойти одно и только одно из k событий A_1, A_2, \dots, A_k с соответствующими вероятностями p_1, p_2, \dots, p_k , то вероятность того, что в этих опытах событие A_1 появится m_1 раз, событие A_2 — m_2 раз, ..., событие A_k — m_k раз, равна

$$P_n(m_1, m_2, \dots, m_k) = \frac{n!}{m_1! m_2! \dots m_k!} p_1^{m_1} p_2^{m_2} \dots p_k^{m_k}, \quad (1.33)$$

где $m_1 + m_2 + \dots + m_k = n$. Вероятности (1.33) называются *полиномиальным распределением*.



Пример 1.31. Производится 3 независимых выстрела по цели. Вероятности попадания при разных выстрелах одинаковы и равны $p = 0,9$. Какова вероятность: а) промаха; б) одного попадания; в) двух попаданий; г) трех попаданий? Решить задачу в случае, если вероятности попадания при разных выстрелах различны: $p_1 = 0,7$, $p_2 = 0,8$, $p_3 = 0,9$.

○ В данном случае $n = 3$, $p = 0,9$, $q = 0,1$. Пользуясь формулой Бернулли (1.32), находим:

а) $P_3(0) = C_3^0 \cdot 0,9^0 \cdot 0,1^3 = 0,001$ — вероятность трех промахов;

б) $P_3(1) = C_3^1 \cdot 0,9^1 \cdot 0,1^2 = 3 \cdot 0,9 \cdot 0,01 = 0,027$ — вероятность одного попадания;

в) $P_3(2) = C_3^2 \cdot 0,9^2 \cdot 0,1^1 = 3 \cdot 0,81 \cdot 0,1 = 0,243$ — вероятность двух попаданий;

г) $P_3(3) = C_3^3 \cdot 0,9^3 \cdot 0,1^0 = 0,9^3 = 0,729$ — вероятность трех попаданий. ●

Эти результаты можно изобразить графически, отложив на оси Ox значения m , на оси Oy — значения $P_n(m)$ (рис. 14).

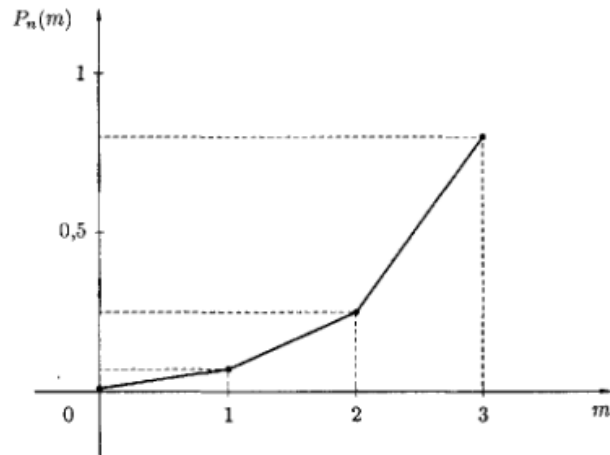


Рис. 14

Ломаная, соединяющая точки $(0; 0,001)$, $(1; 0,027)$, $(2; 0,243)$, $(3; 0,729)$, называется *многоугольником распределения вероятностей*.

Если вероятности при разных выстрелах различны, то производящая функция имеет вид $\varphi_3(z) = (0,3 + 0,7z)(0,2 + 0,8z)(0,1 + 0,9z) = 0,504z^3 + 0,398z^2 + 0,092z + 0,006$. Откуда находим вероятность трех, двух, одного попаданий, промаха соответственно: $P_3(3) = 0,504$, $P_3(2) = 0,398$, $P_3(1) = 0,092$, $P_3(0) = 0,006$. (Контроль: $0,504 + 0,398 + 0,092 + 0,006 = 1$.)

Упражнения

1. Монету подбрасывают 10 раз. Какова вероятность того, что герб выпадет (появится): а) 4 раза; б) ни разу; в) хотя бы один раз.
2. Что вероятнее выиграть у равносильного противника-шахматиста: две партии из четырех или три из шести? Ничьи во внимание не принимаются.
3. В семье трое детей. Какова вероятность того, что: а) все они мальчики; б) один мальчик и две девочки. Считать вероятность рождения мальчика 0,51, а девочки — 0,49.
4. В каждом из карманов (их 2) лежит по коробку спичек (по 10 спичек в коробке). При каждом закуривании карман выбирается наудачу. При очередном закуривании коробок оказался пустым. Найти вероятность того, что во втором коробке 6 спичек.

1.21. Предельные теоремы в схеме Бернулли

Использование формулы Бернулли (1.32) при больших значениях n и m вызывает большие трудности, так как это связано с громоздкими вычислениями. Так, при $n = 200$, $m = 116$, $p = 0,72$ формула Бернулли принимает вид $P_{200}(116) = C_{200}^{116} \cdot (0,72)^{116} \cdot (0,28)^{84}$. Подсчитать результат практически невозможно. Вычисление $P_n(m)$ вызывает затруднения также при малых значениях p (q). Возникает необходимость в отыскании приближенных формул для вычисления $P_n(m)$, обеспечивающих необходимую точность. Такие формулы дают нам предельные теоремы; они содержат так называемые асимптотические формулы, которые при больших значениях испытаний дают сколь угодно малую относительную погрешность. Рассмотрим три предельные теоремы, содержащие асимптотические формулы для вычисления биномиальной вероятности $P_n(m)$ при $n \rightarrow \infty$.

Теорема Пуассона

Теорема 1.5. Если число испытаний неограниченно увеличивается ($n \rightarrow \infty$) и вероятность p наступления события A в каждом испытании неограниченно уменьшается ($p \rightarrow 0$), но так, что их произведение np является постоянной величиной ($np = a = \text{const}$), то вероятность $P_n(m)$ удовлетворяет предельному равенству

$$\lim_{n \rightarrow \infty} P_n(m) = \frac{a^m e^{-a}}{m!}. \quad (1.34)$$

Выражение (1.34) называется *асимптотической формулой Пуассона*.

□ Преобразуем формулу Бернулли (1.32) с учетом того, что $p = \frac{a}{n}$:

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

Переходя к пределу при $n \rightarrow \infty$, получим $\lim_{n \rightarrow \infty} P_n(m) = \frac{a^m e^{-a}}{m!}$
($\lim_{n \rightarrow \infty} \left(1 - \frac{a}{n}\right)^n = e^{-a}$ согласно второму замечательному пределу). ■

Из предельного равенства (1.34) при больших n и малых p вытекает приближенная формула Пуассона

$$P_n(m) \approx \frac{a^m e^{-a}}{m!}, \quad a = np, \quad m = 0, 1, 2, \dots \quad (1.35)$$

Формулу (1.35) применяют, когда вероятность $p = \text{const}$ успеха крайне мала, т.е. сам по себе успех (появление события A) является *редким событием* (например, выигрыш автомобиля по лотерейному билету), но количество испытаний n велико, *среднее число успехов* $np = a$ незначительно. Приближенную формулу (1.35) обычно используют, когда $n \geq 50$, а $np \leq 10$.

Формула Пуассона находит применение в теории массового обслуживания.



Пример 1.32. Завод «Золотая балка» (Крым) отправил в Москву 1500 бутылок вина «Каберне». Вероятность того, что в пути бутылка может разбиться, равна 0,002. Найти вероятность того, что в пути будет разбито не более 4-х бутылок (событие A).

○ Искомая вероятность равна

$$P_{1500}(0) + P_{1500}(1) + P_{1500}(2) + P_{1500}(3) + P_{1500}(4).$$

Так как $n = 1500$, $p = 0,002$, то $a = [np] = 3$. Вероятность события A найдем, используя формулу Пуассона (1.35):

$$P(A) = \frac{3^0 \cdot e^{-3}}{0!} + \frac{3^1 \cdot e^{-3}}{1!} + \frac{3^2 \cdot e^{-3}}{2!} + \frac{3^3 \cdot e^{-3}}{3!} + \frac{3^4 \cdot e^{-3}}{4!} \approx 0,815. \quad \bullet$$

Формулу Пуассона можно считать математической моделью простейшего потока событий.

Потоком событий называют последовательность событий, наступающих в случайные моменты времени (например, поток посетителей в парикмахерской, поток вызовов на телефонной станции, поток отказов элементов, поток обслуженных абонентов и т.п.).

Поток событий, обладающий свойствами стационарности, ординарности и отсутствия последствия называется *простейшим (пуассоновским) потоком*.

Свойство *стационарности* означает, что вероятность появления k событий на участке времени длины τ зависит только от его длины (т.е. не зависит от начала его отсчета). Следовательно, *среднее число событий, появляющихся в единицу времени*, так называемая *интенсивность* λ потока, есть величина постоянная: $\lambda(t) = \lambda$.

Свойство *ординарности* означает, что событие появляется не группами, а поодиночке. Другими словами, вероятность появления более одного события на малый участок времени Δt пренебрежительно мала по сравнению с вероятностью появления только одного события (например, поток катеров, подходящих к причалу, ординарен).

Свойство *отсутствия последствия* означает, что вероятность появления k событий на любом участке времени длины τ не зависит от того, сколько событий появилось на любом другом не пересекающемся с ним участком (говорят: «будущее» потока не зависит от «прошлого», например, поток людей, входящих в супермаркет).

Можно доказать, что вероятность появления m событий простейшего потока за время продолжительностью t определяется формулой

Пуассона

$$P_t(m) = p_m = \frac{(\lambda t)^m \cdot e^{-\lambda t}}{m!}.$$



Пример 1.33. Телефонная станция обслуживает 2000 абонентов. Вероятность позвонить любому абоненту в течение часа равна 0,003. Какова вероятность того, что в течение часа позвонят 5 абонентов?

○ Среднее число позвонивших в течение часа абонентов равно $2000 \cdot 0,003 = 6$ ($a = np = \lambda t$). Стало быть, $p_5 = \frac{6^5 e^{-6}}{5!} \approx 0,13$. ●

Локальная и интегральная теоремы Муавра–Лапласа

В тех случаях, когда число испытаний n велико, а вероятность p наступления события A в каждом испытании постоянна и отлична от нуля и единицы ($p \neq 0, p \neq 1$), для вычисления биномиальных вероятностей используют теоремы Муавра–Лапласа. Приведем только их формулировки в силу сложности доказательства.

Теорема 1.6 (Локальная теорема Муавра–Лапласа). Если вероятность p наступления события A в каждом испытании постоянна и отлична от нуля и единицы, а число независимых испытаний достаточно велико, то вероятность $P_n(m)$ может быть вычислена по приближенной формуле

$$P_n(m) \approx \frac{1}{\sqrt{npq}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad \text{где } x = \frac{m - np}{\sqrt{npq}}. \quad (1.36)$$

Равенство (1.36) тем точнее, чем больше n .

Выражение

$$\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} = \varphi(x) \quad (1.37)$$

называется *функцией Гаусса*, а ее график — *кривой вероятностей* (см. рис. 15).

Равенство (1.36) можно переписать в виде

$$P_n(m) \approx \frac{1}{\sqrt{npq}} \cdot \varphi(x), \quad \text{где } x = \frac{m - np}{\sqrt{npq}}. \quad (1.38)$$

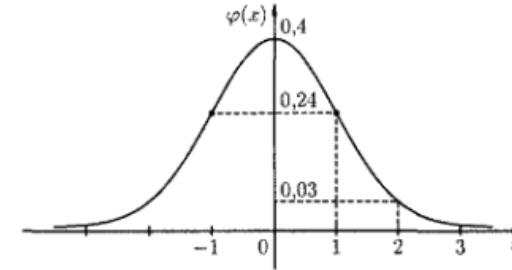


Рис. 15

Для функции $\varphi(x)$ составлены таблицы значений (они находятся, как правило, в так называемая «Приложениях» книг по теории вероятностей см. приложение 1 на с. 249). Пользуясь таблицей, следует учитывать, что:

- а) функция $\varphi(x)$ четная, т. е. $\varphi(-x) = \varphi(x)$;
- б) при $x \geq 4$ можно считать, что $\varphi(x) = 0$.

Функция Гаусса (1.37) будет подробнее рассмотрена в п. 2.7.



Пример 1.34. Вероятность попадания в мишень при одном выстреле для данного стрелка равна 0,7. Найти вероятность того, что при 200 выстрелах мишень будет поражена 160 раз.

○ Здесь $n = 200, p = 0,7, q = 0,3, m = 160$. Применим формулу (1.38). Имеем: $\sqrt{npq} = \sqrt{200 \cdot 0,7 \cdot 0,3} = \sqrt{42} \approx 6,48$, следовательно, $x = \frac{160 - 200 \cdot 0,7}{6,48} = \frac{20}{6,48} \approx 3,09$. Учитывая, что $\varphi(3,09) \approx 0,0034$, получаем $P_{200}(160) \approx \frac{1}{6,48} \cdot 0,0034 \approx 0,0005$. ●

В тех случаях, когда требуется вычислить вероятность того, что в n независимых испытаниях событие A появится не менее k_1 раз, но не более k_2 раз, т. е. $P_n(k_1 \leq m \leq k_2)$ или $P_n(k_1; k_2)$, используют интегральную теорему Муавра–Лапласа (является частным случаем более общей теоремы — центральной предельной теоремы).

Теорема 1.7 (Интегральная теорема Муавра–Лапласа). Если вероятность p наступления события A в каждом испытании постоянна и отлична от нуля и единицы, то вероятность $P_n(k_1 \leq m \leq k_2)$ может быть найдена по приближенной формуле

$$P_n(k_1 \leq m \leq k_2) \approx \frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{x^2}{2}} dx, \quad \text{где } x_1 = \frac{k_1 - np}{\sqrt{npq}}, \quad x_2 = \frac{k_2 - np}{\sqrt{npq}}. \quad (1.39)$$

Равенство (1.39) тем точнее, чем больше n .

Используя функцию Гаусса (1.37), равенство (1.39) можно записать в виде

$$P_n(k_1, k_2) \approx \int_{x_1}^{x_2} \varphi(x) dx.$$

Однако для упрощения вычислений, при использовании формулы (1.39), вводят специальную функцию

$$\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt, \quad (1.40)$$

называемую *нормированной функцией Лапласа*.

Функция (1.40) нечетна ($\Phi_0(-x) = \frac{1}{\sqrt{2\pi}} \int_0^{-x} e^{-\frac{t^2}{2}} dt = [t = -z] = -\frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{z^2}{2}} dz = -\Phi_0(x)$); при $x \geq 5$ можно считать, что $\Phi_0(x) = 0,5$; график функции $\Phi_0(x)$ приведен на рис. 16.

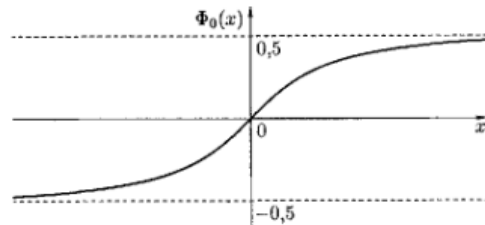


Рис. 16

Выразим правую часть равенства (1.39) через функцию Лапласа (1.40):

$$\begin{aligned} \frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{x^2}{2}} dx &= \frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{t^2}{2}} dt = \\ &= \frac{1}{\sqrt{2\pi}} \int_{x_1}^0 e^{-\frac{t^2}{2}} dt + \frac{1}{\sqrt{2\pi}} \int_0^{x_2} e^{-\frac{t^2}{2}} dt = \Phi_0(x_2) - \Phi_0(x_1). \end{aligned}$$

Равенство (1.39) принимает вид

$$P_n(k_1 \leq m \leq k_2) = \Phi_0(x_2) - \Phi_0(x_1), \quad \text{где } x_1 = \frac{k_1 - np}{\sqrt{npq}}, \quad x_2 = \frac{k_2 - np}{\sqrt{npq}}. \quad (1.41)$$

Эту формулу обычно используют на практике.

Наряду с нормированной функцией Лапласа (1.40) используют функцию

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt, \quad (1.42)$$

называемую также *функцией Лапласа*. Для нее справедливо равенство $\Phi(-x) + \Phi(x) = 1$; она связана с функцией $\Phi_0(x)$ формулой

$$\Phi(x) = 0,5 + \Phi_0(x). \quad (1.43)$$

Имеются таблицы приближенных значений функций $\Phi_0(x)$ и $\Phi(x)$ (интеграл не берется в элементарных функциях), которые приводятся в большинстве учебников по теории вероятностей (см. также приложение 2 на с. 250).

Приближенную формулу для вычисления вероятности $P_n(k_1 \leq m \leq k_2)$ (1.39) можно записать в виде

$$P_n(k_1 \leq m \leq k_2) = \Phi(x_2) - \Phi(x_1) = \Phi_0(x_2) - \Phi_0(x_1), \quad \text{где } x_1 = \frac{k_1 - np}{\sqrt{npq}}, \quad x_2 = \frac{k_2 - np}{\sqrt{npq}}. \quad (1.44)$$



Пример 1.35. Проверкой установлено, что цех в среднем выпускает 96% продукции высшего сорта. На базе приемщик проверяет 200 изделий этого цеха. Если среди них окажется более 10 изделий не высшего сорта, то вся партия изделий бракуется, т. е. возвращается в цех. Какова вероятность того, что партия будет принята?

○ Здесь $n = 200$, $p = 0,04$ (вероятность негодного изделия), $q = 0,96$. Вероятность принятия всей партии, т.е. $P_{200}(0 \leq m \leq 10)$, можно найти по формуле (1.44); здесь $k_1 = 0$, $k_2 = 10$. Находим, что $x_1 = \frac{0 - 200 \cdot 0,04}{\sqrt{200 \cdot 0,04 \cdot 0,96}} \approx -2,89$, $x_2 = \frac{10 - 200 \cdot 0,04}{\sqrt{200 \cdot 0,04 \cdot 0,96}} \approx 0,72$, $P_{200}(0 \leq m \leq 10) = \Phi_0(0,72) - \Phi_0(-2,89) = 0,26424 + 0,49807 = 0,7623$. Заметим, что $\Phi(0,72) - \Phi(-2,89) = 0,7642 - (1 - \Phi(2,89)) = 0,7642 - (1 - 0,998074) = 0,7623$. ●

С помощью функции Лапласа можно найти вероятность отклонения относительной частоты $\frac{n_A}{n}$ от вероятности p в n независимых испытаниях. Имеет место формула

$$P\left\{\left|\frac{n_A}{n} - p\right| \leq \varepsilon\right\} = 2\Phi_0\left(\varepsilon \cdot \sqrt{\frac{n}{pq}}\right),$$

где $\varepsilon > 0$ — некоторое число.

□ Из $\left|\frac{n_A}{n} - p\right| \leq \varepsilon$ следует: $-\varepsilon \leq \frac{n_A}{n} - p \leq \varepsilon$, $np - n\varepsilon \leq n_A \leq np + n\varepsilon$. По формуле (1.28) получаем:

$$\begin{aligned} P_n\{np - n\varepsilon \leq n_A \leq np + n\varepsilon\} &\approx \frac{1}{\sqrt{2\pi}} \int_{-\varepsilon \cdot \sqrt{\frac{n}{pq}}}^{\varepsilon \cdot \sqrt{\frac{n}{pq}}} e^{-\frac{t^2}{2}} dt = \\ &= \frac{2}{\sqrt{2\pi}} \int_0^{\varepsilon \cdot \sqrt{\frac{n}{pq}}} e^{-\frac{t^2}{2}} dt = 2\Phi_0\left(\varepsilon \cdot \sqrt{\frac{n}{pq}}\right), \end{aligned}$$

т.е. $P\left\{\left|\frac{n_A}{n} - p\right| \leq \varepsilon\right\} = 2\Phi_0\left(\varepsilon \cdot \sqrt{\frac{n}{pq}}\right)$. ■



Пример 1.36. Вероятность попадания в цель при одном выстреле равна 0,6. Найти вероятность того, что при $n = 1200$ независимых выстрелах отклонение «частоты» от вероятности по модулю не превышает $\varepsilon = 0,05$.

○ $P_{1200}\left\{\left|\frac{n_A}{n} - 0,6\right| \leq 0,05\right\} = 2\Phi_0\left(0,05 \cdot \sqrt{\frac{1200}{0,6 \cdot 0,4}}\right) = 2\Phi_0(3,54) = 0,9996$. ●

Упражнения

1. На лекции по теории вероятностей присутствуют 84 студента. Какова вероятность того, что среди них есть 2 студента, у которых сегодня день рождения?
2. Вероятность брака при изготовлении некоторого изделия равна 0,02. Найти вероятность того, что среди 200 произведенных изделий не более одного бракованного.
3. Найти вероятность того, что при подбрасывании монеты 100 раз событие A — появление герба — наступит ровно 60 раз.
4. Найти такое число m , чтобы с вероятностью 0,95 можно было бы утверждать, что среди 800 новорожденных более m девочек. Считать, что вероятность рождения девочки равна 0,485.