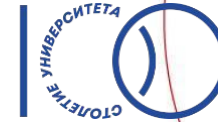




Лекции по теории вероятностей

Белоусова Вероника Игоревна
к.ф.-м.н., доцент



Полная вероятность события. Формула Байеса переоценки гипотез.

Повторные испытания Бернулли.

Вероятность повторных испытаний. Локальная и интегральная теоремы Лапласа

План лекции

- Полная вероятность события.
- Формула Байеса переоценки гипотез.
- Повторные испытания Бернулли.
- Вероятность повторных испытаний.
Локальная и интегральная теоремы
Лапласа



Полная вероятность события

Определение. Пусть событие A может произойти только совместно с одним из событий H_1, H_2, \dots, H_n , образующих полную группу несовместных событий.

Тогда события H_1, H_2, \dots, H_n называются **гипотезами.**



Например

Событие А – презентация фильма прошла
успешно

Н1 – судьи состояли из детей

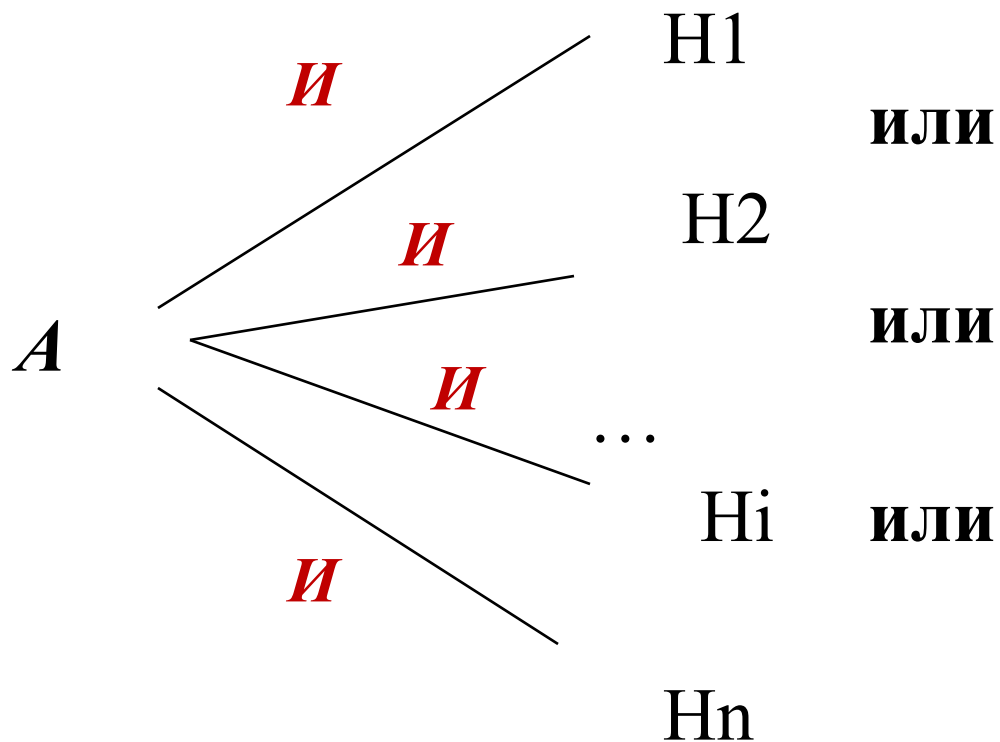
Н2 – судьи из взрослых

Н3 – смешанное судейство



Вероятность события A

Когда наступает событие A ?



Теорема.

Вероятность события A , наступающего совместно с гипотезами H_1, H_2, \dots, H_n , равна:

$$p(A) = \sum_{i=1}^n p(H_i) p(A / H_i),$$

где $p(H_i)$ – вероятность i -й гипотезы,

а $p(A/H_i)$ – вероятность события A при условии реализации этой гипотезы.

Формула называется **формулой полной вероятности**.

доказательство

Можно считать событие A суммой попарно несовместных событий AH_1, AH_2, \dots, AH_n .

Тогда из теорем сложения и умножения следует, что

$$P(A) = P(AH_1 + AH_2 + \dots + AH_n) =$$

$$P(A) = P(AH_1) + P(AH_2) + \dots + P(AH_n) =$$

$$= \sum_{i=1}^n P(H_i)P(A/H_i).$$

что и требовалось доказать.

Пример

Имеется 2 урны с шарами.

В первой из них 3 белых и 4 черных шара, во второй – 2 белых и 5 черных. Из первой урны случайно выбирают шар и перекладывают во вторую.

Найти вероятность того, что вытасченный затем наудачу шар из второй урны белый.

Решение: Сформулируем гипотезы:

H_1 : из первой урны вынут белый шар,

H_2 : из первой урны вынут черный шар.

Тогда

1 урна

3 белых 4 черных

$$\bigcirc P(H_1) = \frac{3}{7}$$

2 урна

3 белых 5 черных

$$P(A / H_1) = \frac{3}{8}$$

2 урна

2 белых 5 черных

$$\bullet P(H_2) = \frac{4}{7}$$

2 урна

2 белых 6 черных

$$P(A / H_2) = \frac{2}{8}$$

$$\begin{aligned} P(A) &= P(H_1)P(A / H_1) + P(H_2)P(A / H_2) = \\ &= \frac{9}{56} + \frac{8}{56} = \frac{17}{56} \end{aligned}$$

Переоценка гипотез

- Пусть известен результат опыта, а именно то, что произошло событие A .
- Этот факт может изменить априорные (то есть известные до опыта) вероятности гипотез.
- **Вопрос**: Чему равна вероятность того, что к возникновению события привело появление определенной гипотезы $p(H_i / A)$?

Вероятность события

- В какую формулу входит условная вероятность

$$p(H_i / A) ?$$

- Условная вероятность $p(H_i / A)$ входит в формулу теоремы умножения вероятностей:

$$p(A \cdot H_i) = p(A)p(H_i / A) = p(H_i)p(A / H_i),$$

Формула Байеса (теорема гипотез).

Из формулы $p(A \cdot H_i) = p(H_i)p(A / H_i)$ выразим $p(H_i / A)$:

Получим используемую для переоценки вероятностей гипотез при известном результате опыта **формулу Байеса**:

$$p(H_i / A) = \frac{p(H_i)p(A / H_i)}{p(A)}.$$

Пример

После двух выстрелов двух стрелков, вероятности попаданий которых равны 0,6 и 0,7, в мишени оказалась одна пробоина.

Найти вероятность того, что попал первый стрелок.

Решение:

Пусть событие A – одно попадание при двух выстрелах,

H_1 – первый попал, а второй промахнулся,

H_2 – первый промахнулся, а второй попал,

H_3 – оба попали,

H_4 – оба промахнулись.

Пример

Вероятности гипотез:

$$p(H_1) = 0,6 \cdot 0,3 = 0,18, \quad (\text{первый попал, а второй промахнулся})$$

$$p(H_2) = 0,4 \cdot 0,7 = 0,28, \quad (\text{первый промахнулся, а второй попал})$$

$$p(H_3) = 0,6 \cdot 0,7 = 0,42, \quad (\text{оба попали})$$

$$p(H_4) = 0,4 \cdot 0,3 = 0,12. \quad (\text{оба промахнулись})$$

$$\text{Тогда } p(A/H_1) = p(A/H_2) = 1, \quad p(A/H_3) = p(A/H_4) = 0.$$

$$p(A) = 0,18 \cdot 1 + 0,28 \cdot 1 + 0,42 \cdot 0 + 0,12 \cdot 0 = 0,46.$$

$$\text{Применяя формулу Байеса, получим: } p(H_1 / A) = \frac{0,18 \cdot 1}{0,46} = \frac{9}{23} \approx 0,391.$$

Схема Бернулли повторения испытаний

Рассмотрим серию из n испытаний, в каждом из которых событие A появляется с одной и той же вероятностью p , причем результат каждого испытания не зависит от результатов остальных.

Подобная постановка задачи называется **схемой повторения испытаний.**



Схема повторения испытаний. Формула Бернулли

Найдем вероятность того, что в такой серии событие A произойдет ровно k раз (неважно, в какой последовательности).

Событие A - сумма равновероятных несовместных событий, заключающихся в том, что:

1) *из n выбраны k таких испытаний, при которых:*

2) A произошло в этих k испытаниях

и

3) не произошло в остальных $(n-k)$ испытаниях.

Формула Бернулли

1. Число таких событий равно числу сочетаний C_n^k из n по k ,

2. Вероятность каждого из испытаний равна:

$$p^k q^{n-k},$$

где $q = 1 - p$ – вероятность того, что в данном опыте A не произошло.

Применяя теорему сложения для несовместных событий, получим **формулу Бернулли**:

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



Якоб Бернулли



6 января 6 января 1655 -
16 августа 1705,
швейцарский математик, один
из создателей теории
вероятностей.

Пример 1

Для получения приза нужно собрать 5 изделий с особым знаком на этикетке. Найти вероятность того, что придется купить 10 изделий, если этикетки с этим знаком имеют 5% изделий.

Решение:

Пример 1

Для получения приза нужно собрать 5 изделий с особым знаком на этикетке. Найти вероятность того, что придется купить 10 изделий, если этикетки с этим знаком имеют 5% изделий.

Решение:

Из постановки задачи следует, что последнее купленное изделие имеет особый знак. Следовательно, из предыдущих девяти эти знаки имели 4 изделия.

Пример 1

Мы ищем вероятность события A :

Последнее изделие было с особым знаком

и

4 и 9 из предыдущих тоже были с особым знаком.

Найдем вероятность второго условия по формуле

Бернулли: $p_9(4) = C_9^4 \cdot (0,05)^4 \cdot (0,95)^5 = 0,0006092.$

Тогда $P(A) = 0,0006092 \cdot 0,05 = 0,0000304.$



Пример 2

Что вероятнее, выиграть у равносильного соперника не менее 3 партий из 4 или не менее 5 из 8?

Пример 2

Что вероятнее, выиграть у равносильного соперника не менее 3 партий из 4 или не менее 5 из 8?

Решение:

A- выиграть у равносильного соперника не менее 3 партий из 4

B- выиграть у равносильного соперника не менее 5 из 8

$$P(A) = C_4^3 \cdot (0,5)^4 + C_4^4 \cdot (0,5)^4 = (0,5)^4 \cdot 5 = \frac{5}{16}$$

$$P(B) = (0,5)^8 \left(C_8^5 + C_8^6 + C_8^7 + C_8^8 \right) =$$

$$= (0,5)^8 \cdot (56 + 28 + 8 + 1) = \frac{93}{256}$$

$$\frac{P(B)}{P(A)} = \frac{93 \cdot 16}{16 \cdot 16 \cdot 5} = \frac{93}{80} > 1$$

Пример 3

Проводится 10 испытаний по схеме Бернулли с вероятностью успеха p и неудачи q . Найти вероятность того, что успех произойдет только в двух случаях, при этом между ними 3 раза испытание будет неуспешным.

Решение:



Когда формула Бернулли неудобна?

Приближение Пуассона для схемы Бернулли

При повторении испытаний равновозможные исходы испытания будут наступать в среднем одинаково часто.

Пусть при большом числе испытаний вероятность появления A в одном опыте мала, а произведение $np = \lambda$ сохраняет постоянное значение для разных серий опытов

(то есть среднее число появлений события A в разных сериях испытаний остается неизменным).

Приближение Пуассона для схемы Бернулли

Из $np = \lambda$ следует $p = \lambda/n$. Применим формулу Бернулли:

$$p_n(k) = \frac{n(n-1)(n-2)\dots(n-k+1)}{k!} p^k (1-p)^{n-k} =$$

$$= \frac{n(n-1)\dots(n-k+1)}{k!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \quad \text{затем при } n \rightarrow \infty$$

$$\begin{aligned} p_n(k) &\approx \frac{\lambda^k}{k!} \lim_{n \rightarrow \infty} \left(1 \cdot \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{k-1}{n}\right) \left(1 - \frac{\lambda}{n}\right)^{n-k} \right) = \\ &= \frac{\lambda^k}{k!} \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k} = \frac{\lambda^k}{k!} \cdot e^{-\lambda} \cdot 1. \end{aligned}$$

Формула Пуассона

Таким образом, получена **формула Пуассона**

$$P_n(k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

которая позволяет найти вероятность k появлений события A для массовых (n велико) и редких (p мало) событий.

Достаточно условий $p < 0,1$ и $npq < 10$



Симеон Дени Пуассон

французский математик французский математик, механик французский математик, механик и физик.

(1781 — 1842 г.)



пример

Вероятность сбоя в системе при переключении ее режимов равна 0,001. Найти вероятность того, что в 5000 переключениях будет не меньше двух сбоев.

пример

Вероятность сбоя в системе при переключении ее режимов равна 0,001. Найти вероятность того, что в 5000 переключениях будет не меньше двух сбоев.

Решение

A – в 5000 переключениях будет от 2 до 5000 сбоев.

\bar{A} – в 5000 переключениях будет 0 или 1 сбой.

$$p=0,001 \quad q=0,999 \quad n=5000 \quad npq=4,995 < 10$$

$$\lambda=np=5$$

$$P_{5000}(0) \approx \frac{5^0}{0!} e^{-5} \approx 0,0067$$

$$P_{5000}(1) \approx \frac{5^1}{1!} e^{-5} \approx 5 \cdot 0,0067$$

$$P_{5000}(0) + P_{5000}(1) \approx 0,04$$

Наивероятнейшее число появлений события

Наивероятнейшим числом появления события в независимых испытаниях называется такое число , вероятность которого превышает или, по крайней мере, не меньше вероятности каждого из остальных возможных чисел появления события .

Обозначим $P_{m_0, n}$ вероятность, соответствующую наивероятнейшему числу m_0 .

Наивероятнейшее число появлений события

Используя формулу, $P_n(k) = C_n^k \cdot p^k \cdot q^{n-k}$
получим:

$$P_{m_0, n} = C_n^{m_0} \cdot p^{m_0} \cdot q^{n-m_0} = \frac{n!}{m_0!(n-m_0)!} p^{m_0} q^{n-m_0}$$

Согласно определению наивероятнейшего числа, вероятности наступления события А соответственно m_0-1 и m_0+1 раз должны, по крайней мере не превышать $P_{m_0, n}$:

$$P_{m_0+1, n} \leq P_{m_0, n} \quad P_{m_0-1, n} \leq P_{m_0, n}$$

Сначала распишем неравенство $P_{m_0+1,n} \leq P_{m_0,n}$

$$\frac{n!}{(m_0 + 1)!(n - m_0 - 1)!} p^{m_0+1} q^{n-m_0-1} \leq \frac{n!}{m_0!(n - m_0)!} p^{m_0} q^{n-m_0},$$

$$\frac{1}{(m_0 + 1)} p \leq \frac{1}{(n - m_0)} q$$

$$(n - m_0)p \leq (m_0 + 1)q \rightarrow np - q \leq m_0(p + q) = m_0.$$

Аналогично, расписывая неравенство $P_{m_0,n} \leq P_{m_0-1,n}$,

получим $np + p \geq m_0$.

Наивероятнейшее число появлений события

Для определения наивероятнейшего числа событий в повторных испытаниях используют формулу:

$$np - q \leq m_0 \leq np + p$$

Длина интервала, определяемого неравенством равна единице, т. е.

$$(np + p) - (np - q) = p + q = 1.$$

Наивероятнейшее число появлений события

$$np - q \leq m_0 \leq np + p$$

Событие может произойти в испытаниях только целое число раз, то следует иметь в виду, что:

- 1) если $np - q$ — целое число, то существуют два значения наивероятнейшего числа, а именно: $m_0 = np - q$ и $m'_0 = np - q + 1 = np + p$;
- 2) если $np - q$ — дробное число, то существует одно наивероятнейшее число, а именно: единственное целое, заключенное между дробными числами, полученными из неравенства;
- 3) если np — целое число, то существует одно наивероятнейшее число, а именно: $m_0 = np$.

Вероятность наивероятнейшего числа появлений события

При больших значениях n пользоваться формулой Бернулли для расчета вероятности, соответствующей наивероятнейшему числу, неудобно. Если в формулу Бернулли подставить формулу Стирлинга

$$n! = n^n e^{-n} \sqrt{2\pi n},$$

и принять $m_0 = np$, то получим

***формулу вероятности наступления
наивероятнейшего числа появления события A:***

$$P_{m_0, n} \approx \frac{1}{\sqrt{2\pi} \sqrt{npq}}.$$

Пример

Известно, что $1/15$ часть продукции, поставляемой заводом на торговую базу, не удовлетворяет всем требованиям стандарта. На базу была завезена партия изделий в количестве 250 шт. Найти наивероятнейшее число изделий, удовлетворяющих требованиям стандарта, и вычислить вероятность того, что в этой партии окажется наивероятнейшее число изделий.

Решение:

По условию $n=250$ $q=1/15$, значит $p=14/15$.

$$np - q \leq m_0 \leq np + p \quad 250 \cdot \frac{14}{15} - \frac{1}{15} \leq m_0 \leq 250 \cdot \frac{14}{15} + \frac{14}{15}$$

$$233,26 \leq m_0 \leq 234,26 \Rightarrow m_0 = 234.$$

$$P_{m_0, n} \approx \frac{1}{\sqrt{2\pi} \sqrt{npq}} \approx 0,254.$$

Локальная теорема Лапласа

Локальная теорема Лапласа дает асимптотическую формулу, которая позволяет приближенно найти вероятность появления событий ровно t раз в n испытаниях, если число испытаний достаточно велико, а p и q не достаточно малы.

Локальная теорема Лапласа

Теорема. Если вероятность p появления события A в каждом испытании постоянна и отлична от 0 и 1, то вероятность $P_n(m)$ того, что событие A появится в n испытаниях ровно m раз, приближенно равна (тем точнее, чем больше n) значению функции

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

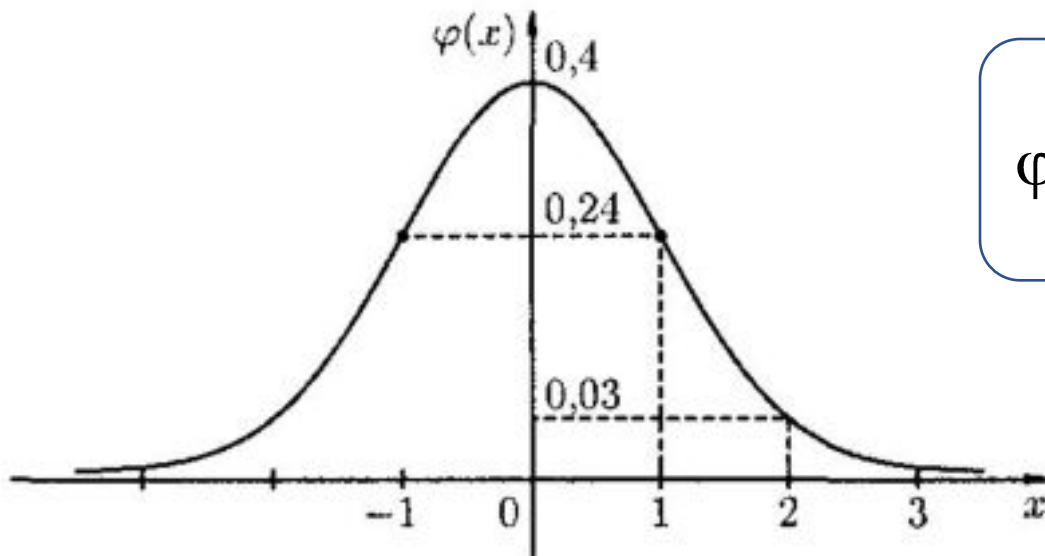
$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \quad - \text{ функция табличная}$$

Свойства функции $\varphi(x)$:

1) функция четная $\varphi(-x) = \varphi(x)$;

2) при $x \geq 4$ $\varphi(x) \approx 0$.

$\varphi(x)$ называют функцией Гаусса



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

Пример

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

Решение. По условию $n=400$, $m=80$, $p=0,2$, $q=0,8$ ($nprq=64$).
Применим асимптотическую формулу Лапласа.

$$P_{400}(80) \approx \frac{1}{\sqrt{400 \cdot 0,2 \cdot 0,8}} \cdot \varphi(x) \approx \frac{1}{8} \varphi(x).$$

Вычислим определяемое данными задачи значение x :

$$x = \frac{m - np}{\sqrt{nprq}} = \frac{80 - 400 \cdot 0,2}{8} = 0.$$

По таблице (Приложение 1) находим $\varphi(0) = 0,3989$.

Искомая вероятность $P_{80,100} = \frac{1}{8} \cdot 0,3989 = 0,04986$.

Интегральная теорема Лапласа

Пусть в n независимых испытаниях, в каждом из которых вероятность появления события A постоянна, равна p и отлична от 0 и 1,

необходимо вычислить вероятность $P_n(m_1, m_2)$ того, что A появится не менее m_1 и не более m_2 раз (от m_1 до m_2 раз), тогда

применяют *интегральную теорему Лапласа*.

Интегральная теорема Лапласа

Теорема. Если вероятность наступления события A постоянна и отлична от 0 и 1, то приближенно вероятность $P_n(m_1, m_2)$ появления события A от m_1 до m_2 раз равна

$$P_n(m_1, m_2) \approx \frac{1}{\sqrt{2\pi}} \int_{x'}^{x''} e^{-x^2/2} dx,$$

$$\text{где } x' = \frac{m_1 - np}{\sqrt{npq}}, \quad x'' = \frac{m_2 - np}{\sqrt{npq}}.$$

Функция Лапласа

Неопределенный интеграл $\int e^{-x^2/2} dx$

не выражается через элементарные функции.

Функция Лапласа – табличная (Приложение 2)

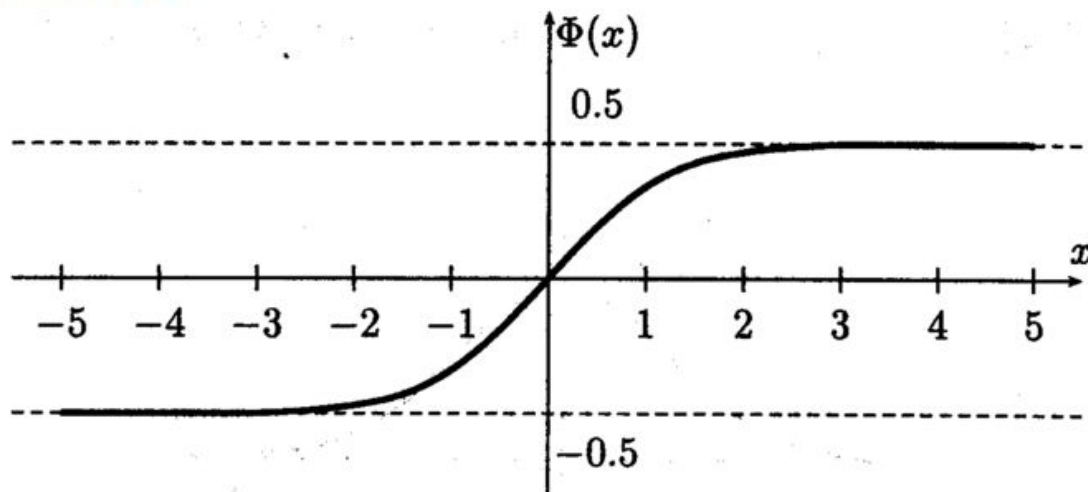
$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-z^2/2} dz$$

Свойства функции Лапласа

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{x'}^{x''} e^{-z^2/2} dz$$

1) $\Phi(x)$ нечетна, т. е. $\Phi(-x) = -\Phi(x)$.

2) для $x > 5$ можно принять $\Phi(x) = 0,5$. Таблица содержит значения функции $\Phi(x)$ лишь для $x \in [0; 5]$.



Пример

Вероятность того, что деталь изготовлена с нарушениями стандартов, $p = 0,2$. Найти вероятность того, что среди 400 случайно отобранных деталей нестандартных окажется от 70 до 100 деталей.

Решение. По условию $p = 0,2$, $q = 0,8$, $n = 400$, $m_1 = 70$, $m_2 = 100$.

Воспользуемся интегральной теоремой Лапласа:

$$P_{(70,100),400} \approx \Phi(x'') - \Phi(x').$$

Вычислим пределы интегрирования:

Нижний $x' = \frac{m_1 - np}{\sqrt{npq}} = \frac{70 - 400 \cdot 0,2}{\sqrt{400 \cdot 0,2 \cdot 0,8}} = -1,25,$

Верхний $x'' = \frac{m_2 - np}{\sqrt{npq}} = \frac{100 - 400 \cdot 0,2}{\sqrt{400 \cdot 0,2 \cdot 0,8}} = 2,5,$

Пример

Таким образом

$$P_{(70,100),400} \approx \Phi(2,5) - \Phi(-1,25) = \Phi(2,5) + \Phi(1,25).$$

По таблице прил. 2 находим $\Phi(2,5) = 0,4938$; $\Phi(1,25) = 0,3944$.

Искомая вероятность

$$P_{(70,100),400} = 0,4938 + 0,3944 = 0,8882.$$

Применение интегральной теоремы Лапласа

Поставим задачу найти вероятность того, что отклонение относительной частоты $\frac{m}{n}$ от постоянной вероятности p по абсолютной величине не превышает заданного числа $\varepsilon > 0$.

Другими словами, найдем вероятность осуществления неравенства

$$\left| \frac{m}{n} - p \right| \leq \varepsilon, \text{ что то же самое, } -\varepsilon \leq \frac{m}{n} - p \leq \varepsilon.$$

Эту вероятность будем обозначать так: $P \left\{ \left| \frac{m}{n} - p \right| \leq \varepsilon \right\}$.

Применение интегральной теоремы Лапласа

$$-\varepsilon \leq \frac{m}{n} - p \leq \varepsilon.$$

Умножим неравенство на множитель $\sqrt{\frac{n}{pq}}$.

$$-\varepsilon \sqrt{\frac{n}{pq}} \leq \frac{m - np}{\sqrt{npq}} \leq \varepsilon \sqrt{\frac{n}{pq}}.$$

По теореме Лапласа $P \left\{ x' \leq \frac{m - np}{\sqrt{npq}} \leq x'' \right\} = \Phi(x'') - \Phi(x')$.

Тогда получаем $\Phi(x'') - \Phi(x') = \Phi\left(\varepsilon \sqrt{\frac{n}{pq}}\right) - \Phi\left(-\varepsilon \sqrt{\frac{n}{pq}}\right)$.

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

Пример

Вероятность того, что деталь нестандартна, $p = 0,1$. Найти вероятность того, что среди случайно отобранных 400 деталей относительная частота появления нестандартных деталей отклонится от вероятности $p = 0,1$ по абсолютной величине не более чем на 0,03.

Решение. По условию $n = 400$, $p = 0,1$, $q = 0,9$, $\varepsilon = 0,03$.

Требуется

найти вероятность $P \left\{ \left| \frac{m}{400} - 0,1 \right| \leq 0,03 \right\}$.

Используя формулу
получаем

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

$$P \left\{ \left| \frac{m}{400} - 0,1 \right| \leq 0,03 \right\} \approx 2\Phi \left(0,03 \sqrt{\frac{400}{0,1 \cdot 0,9}} \right) = 2\Phi(2)$$

Пример

$$P \left\{ \left| \frac{m}{400} - 0,1 \right| \leq 0,03 \right\} \approx 2\Phi \left(0,03 \sqrt{\frac{400}{0,1 \cdot 0,9}} \right) = 2\Phi(2)$$

По таблице прил. 2 находим $\Phi(2) = 0,4772$, следовательно, $2\Phi(2) = 0,9544$.

Итак, искомая вероятность приближенно равна 0,9544.

Смысл полученного результата таков: если взять достаточно большое число проб по 400 деталей в каждой, то примерно в 95,44% этих проб отклонение относительной частоты от постоянной вероятности $p = 0,1$ по абсолютной величине не превысит 0,03.



**На сегодня всё.
Расходимся.**

$$\sum_{n=1}^{\infty} (2n + 1)$$

$$\sum_{n=1}^{\infty} \frac{1}{n}$$

$$\sum_{n=1}^{\infty} \frac{1}{\sqrt{n}}$$

$$\sum_{n=1}^{\infty} \frac{n}{7n + 3}$$

$$\sum_{n=1}^{\infty} \frac{n-1}{\sqrt{2n^4 - n + 5}}$$

$$\sum_{n=2}^{\infty} \frac{1}{\ln n}$$

$$\sum_{n=1}^{\infty} \frac{n^3 - 3n + 1}{n^2 + 4}$$

$$\sum_{n=1}^{\infty} \frac{1}{\sin n}$$