

Теоретические вопросы для подготовки к экзамену по дисциплине «МАТЕМАТИКА»
(III семестр, специальности ПОИТ, ДЭВИ)

1. Элементы комбинаторики: размещения, сочетания, перестановки.

Комбинаторика изучает, сколькими различными способами можно составить множества (комбинации), удовлетворяющие определенным условиям, из элементов заданного множества.

Правило произведения: если объект типа X можно выбрать n способами и при каждом таком выборе объект типа Y можно выбрать m способами, то выбор пары (X, Y) в указанном порядке можно осуществить $n \cdot m$ способами.

Правило суммы: если объект типа X можно выбрать n способами, а объект типа $Y - m$ способами, то выбор объекта типа X или Y можно осуществить $m + n$ способами.

Число P_n всех возможных способов переставить n различных элементов – число **перестановок** (из n различных элементов) равно

$$P_n = n \cdot (n-1) \cdot (n-2) \dots 2 \cdot 1 = n!$$

Число A_n^m **размещений** (упорядоченных комбинаций) из n различных элементов по m элементам (местам), отличающихся либо самими элементами, либо их порядком, равно

$$A_n^m = n(n-1)(n-2)\dots(n-m+1), \text{ где } m \leq n.$$

Число C_n^m **сочетаний** (неупорядоченных комбинаций) из n различных элементов по m элементам (порядок выбранных элементов не учитывается) равно

$$C_n^m = \frac{n(n-1)(n-2)\dots(n-m+1)}{m!} = \frac{A_n^m}{m!} = \frac{n!}{m!(n-m)!},$$

причем $0! = 1$. Отметим, что $C_n^m = C_n^{n-m}$; $C_n^0 = C_n^n = 1$; $C_n^1 = C_n^{n-1} = n$.

Пусть имеется множество, содержащее n элементов. Каждая упорядоченная комбинация, содержащая m элементов из этих n , называется размещением из n элементов по m . Число **размещений** (упорядоченных комбинаций) из n различных элементов по m элементам (местам), отличающихся либо самими элементами, либо их порядком, называется числом размещений из n по m и обозначается A_n^m . Можно сказать, что число размещений A_{nm} – это число способов разместить m из n элементов по m местам.

Пусть имеется множество, содержащее n элементов. Неупорядоченные комбинации (порядок не имеет значения), содержащие m элементов из данных n , называются **сочетаниями** из n элементов по m . Число сочетаний из n по m обозначается C_n^m . Таким образом, число сочетаний C_n^m – это число способов выбрать m элементов из данных n элементов (порядок выбранных элементов не учитывается).

2. Пространство элементарных исходов. Классическое определение вероятности. Методы задания вероятностей.

Пусть проводится испытание с конечным числом попарно несовместных равновозможных исходов $\omega_1, \omega_2, \dots, \omega_n$, образующих полную группу событий. Такие исходы называются **элементарными исходами**, или **элементарными событиями**. При этом говорят, что

испытание сводится к схеме случаев. Множество всех элементарных исходов (*пространство элементарных исходов*) будем обозначать $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$.

1. Классическое определение вероятности: вероятность $P(A)$ случайного события A равна

$$P(A) = \frac{m}{n},$$

где $m = m_A$ – число элементарных исходов испытания, благоприятствующих появлению события A , n – общее число равновозможных элементарных исходов испытания.

Для того, чтобы можно было применить классическое определение вероятности, необходимо, чтобы случайный эксперимент сводился к схеме случаев, т. е.:

1) элементарные исходы эксперимента должны быть равновозможны;

2) элементарные исходы должны образовывать конечное (или счетное) множество.

2. Геометрическая вероятность может использоваться, если исходы случайного эксперимента равновозможны, но образуют бесконечное несчетное пространство элементарных исходов, которое можно представить в виде некоторой геометрической фигуры – области на числовой прямой, на плоскости или в пространстве.

Опр. 1. Пусть G – геометрическая фигура (область), представляющая пространство элементарных исходов данного эксперимента; g – область, представляющая все элементарные исходы, благоприятствующие событию A (рис. 1). *Геометрической вероятностью* события A называется отношение меры области g к мере области G :

$$P(A) = \frac{\mu(g)}{\mu(G)}.$$

При этом если G – отрезок или кривая, то $\mu(G)$ – длина отрезка или кривой; если G – плоская область, то $\mu(G)$ – площадь этой области; если G – пространственное тело, то $\mu(G)$ – объем этого тела.

3. Статистическая вероятность.

Классическое определение вероятности неприменимо, если исходы случайного эксперимента не равновозможны. Например, при бросании неправильной игральной кости выпадения ее различных граней не равновозможны. В таких случаях иногда используют понятие статистической вероятности.

Опр. 2. Пусть при проведении n испытаний событие A появилось в m испытаниях. Отношение $w(A) = \frac{m}{n}$ называется

относительной частотой появления события A в данной серии испытаний.

Относительная частота не является величиной постоянной. Если мы проведем еще одну серию из n или n_1 испытаний, то событие A появится m_1 раз, причем $\frac{m_1}{n_1} \neq \frac{m}{n}$, но если n и n_1 достаточно

велики и условия эксперимента достаточно стабильны, то $\frac{m_1}{n_1} \approx \frac{m}{n}$.

Опр. 3. Если относительная частота события обладает свойством статистической устойчивости, т. е. в различных сериях испытаний изменяется незначительно, в качестве *статистической вероятности* события принимают относительную частоту или ее приближенное значение.

3. Вероятностное пространство. Аксиомы теории вероятностей. Основные теоремы о вероятности.

Пусть задано некоторое множество Ω исходов эксперимента, которое мы будем называть *пространством элементарных исходов*. Пусть \mathfrak{F} – некоторый класс (система, множество) случайных событий, т. е. подмножество множества Ω .

Опр. 1. Класс событий \mathfrak{F} называется *σ -алгеброй* событий, если:

- 1) $\Omega \in \mathfrak{F}$ (достоверное событие принадлежит классу \mathfrak{F});
- 2) если $A, B \in \mathfrak{F}$, то $A + B, AB, A \setminus B \in \mathfrak{F}$ (если A и B являются событиями, то их сумма $A + B$, произведение AB и разность $A \setminus B$ также являются событиями);
- 3) если $A_1, A_2, \dots, A_n, \dots \in \mathfrak{F}$, то $A_1 + A_2 + \dots + A_n + \dots \in \mathfrak{F}$;
 $A_1 A_2 \dots A_n, \dots \in \mathfrak{F}$ (сумма и произведение счетного числа событий также являются событиями).

Опр. 2. Вероятностью (или *вероятностной мерой*) называется числовая функция $P: \mathfrak{F} \rightarrow [0; 1]$, определенная для каждого события $A \in \mathfrak{F}$ и удовлетворяющая следующим условиям (*аксиомам вероятности*):

A1. Аксиома неотрицательности: вероятность любого события неотрицательна, т. е. $P(A) \geq 0$ для любого события $A \in \mathfrak{F}$;

A2. Аксиома нормированности: вероятность достоверного события равна 1, т. е. $P(\Omega) = 1$;

A3. Аксиома аддитивности: вероятность суммы несовместных событий равна сумме их вероятностей, т. е. если события $A_1, A_2, \dots, A_n, \dots \in \mathfrak{F}$ попарно несовместны ($A_i A_k = \emptyset$ при всех $i \neq k$), то

$$P(A_1 + A_2 + \dots + A_n + \dots) = P(A_1) + P(A_2) + \dots + P(A_n) + \dots$$

Опр. 3. Тройка объектов $(\Omega, \mathfrak{F}, P)$, где Ω – некоторое множество, называемое пространством элементарных исходов; \mathfrak{F} – σ -алгебра событий (подмножество множества Ω); P – вероятность (вероятностная мера), определенная на классе событий \mathfrak{F} , называется *вероятностным пространством*.

Свойства вероятности

1. Вероятность невозможного события равна 0: $P(\emptyset) = 0$.
2. Сумма вероятностей противоположных событий равна 1:

$$P(A) + P(\bar{A}) = 1$$

для любого события A .

3. Вероятность любого события не меньше 0 и не больше 1:

$$0 \leq P(A) \leq 1$$

для любого события A .

Упражнение. Вывести свойства 1, 2, 3 из аксиом вероятности.

Теорема сложения вероятностей

Т 1 (теорема сложения вероятностей). Вероятность суммы двух событий равна сумме вероятностей этих событий за вычетом вероятности их произведения: для любых событий A и B

$$P(A + B) = P(A) + P(B) - P(AB).$$

Доказательство. Действительно, это вытекает из представления событий $A + B$ и B посредством суммы несовместных событий: $A + B = A + \bar{B}\bar{A}$, $B = \bar{B}\bar{A} + AB$ (см. рис. 9).

Следствие 1 (теорема сложения вероятностей несовместных событий). Вероятность суммы двух несовместных событий равна сумме их вероятностей: если события A и B несовместны, то

$$P(A + B) = P(A) + P(B).$$

Отметим, что это утверждение является частным случаем аксиомы А3 аддитивности вероятности.

Следствие 2 (свойство полной группы событий). Сумма вероятностей событий H_1, H_2, \dots, H_n , образующих полную группу событий, равна 1:

$$P(H_1) + P(H_2) + \dots + P(H_n) = 1.$$

Опр. 1. Условной вероятностью $P(A|B)$ события A при условии, что произошло событие B ($P(B) \neq 0$), называется отношение вероятности произведения этих событий к вероятности события B :

$$P(A|B) = \frac{P(AB)}{P(B)}.$$

Упражнение 3. Применить эту формулу для нахождения условной вероятности в примере 1.

Теорема умножения вероятностей

Из определения условной вероятности вытекает следующее утверждение.

Т 2 (теорема умножения вероятностей). Вероятность произведения двух событий равна произведению вероятности одного из них на условную вероятность другого при условии, что первое событие произошло:

$$P(AB) = P(A)P(B|A).$$

Следствие 1. $P(ABC) = P(A)P(B|A)P(C|AB)$.

Опр. 2. Событие A называется **независимым** от события B , если $P(A|B) = P(A)$.

Иными словами, событие A не зависит от события B , если вероятность его появления не зависит от того, произошло или не произошло событие B .

Упражнение 5. Показать, что если событие A не зависит от события B , то и событие B не зависит от A , т. е. если $P(A|B) = P(A)$, то $P(B|A) = P(B)$.

Таким образом, зависимость или независимость событий всегда взаимны, поэтому мы можем говорить, что события A и B независимы.

Для независимых событий теорема умножения вероятностей принимает особенно простой вид.

Следствие 2 (теорема умножения вероятностей независимых событий). Вероятность произведения независимых событий равна произведению их вероятностей: если события A и B независимы, то

$$P(AB) = P(A)P(B).$$

Замечание. При решении задач о независимости событий судят по смыслу условия задачи.

4. Сумма событий. Совместные и несовместные события. Теорема сложения вероятностей для совместных и несовместных событий.

Суммой $A+B$ событий A и B называется событие $C=A+B$, состоящее в наступлении хотя бы одного из событий A **или** B (в результате СЭ произошло или событие A , или событие B , или события A и B одновременно).

5. Произведение событий. Понятие условной вероятности. Теорема умножения вероятностей для зависимых и независимых событий.

Произведением $A \cdot B = AB$ событий A и B называется событие $C = AB$, состоящее в том, что в результате СЭ произошли **и** событие A , **и** событие B .

6. Формула полной вероятности. Формула Байеса.

Т 3 (формула полной вероятности). Если событие A может наступить при появлении одного из n попарно несовместных событий (**гипотез**) H_1, H_2, \dots, H_n , образующих полную группу событий, то вероятность события A равна сумме произведений вероятностей каждой из гипотез на соответствующую условную вероятность события A :

$$P(A) = P(H_1)P(A|H_1) + P(H_2)P(A|H_2) + \dots + P(H_n)P(A|H_n).$$

Доказательство. Гипотезы H_1, H_2, \dots, H_n образуют полную группу событий, т. е. они попарно несовместны и $H_1 + H_2 + \dots + H_n = \Omega$. Тогда событие A можно представить в виде $A = \Omega A = H_1 A + H_2 A + \dots + H_n A$, причем слагаемые попарно несовместны. Следовательно, применяя теорему сложения вероятностей несовместных событий, а затем теорему умножения вероятностей, получим

$$\begin{aligned} P(A) &= P(H_1 A) + P(H_2 A) + \dots + P(H_n A) = \\ &= P(H_1)P(A|H_1) + P(H_2)P(A|H_2) + \dots + P(H_n)P(A|H_n). \end{aligned}$$

Формула Байеса применяется, если событие A произошло и требуется *переоценить* вероятности гипотез, т. е. найти $P(H_k|A)$.

Т 4 (формула Байеса). Если события H_1, H_2, \dots, H_n образуют полную группу событий, то

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

Доказательство. Используя определение условной вероятности, теорему умножения вероятностей и формулу полной вероятности, имеем

$$P(H_k | A) = \frac{P(H_k A)}{P(A)} = \frac{P(H_k)P(A | H_k)}{P(A)} = \frac{P(H_k)P(A | H_k)}{\sum_{i=1}^n P(H_i)P(A | H_i)}. \triangleleft$$

Вероятности $P(H_k)$, известные до проведения опыта, называются *априорными* (лат. *a priori* – буквально «от предшествующего») вероятностями гипотез, вероятности $P(H_k|A)$ называются *апостериорными* (лат. *a posteriori* – от последующего).

7. Схема Бернулли. Формула Бернулли. Предельные теоремы Пуассона и Муавра-Лапласа в схеме Бернулли,

Пусть проводится n независимых в совокупности испытаний (СЭ), в каждом из которых возможно только два исхода: A – успех и \bar{A} – неуспех, причем вероятность наступления успеха в каждом испытании постоянна и равна p . Такая последовательность испытаний называется *схемой Бернулли*.

В схеме Бернулли вероятность $P_n(m)$ наступления m успехов в n независимых испытаниях – вероятность того, что в этих испытаниях событие A наступит ровно m раз, вычисляется по *формуле Бернулли*:

$$P_n(m) = C_n^m p^m q^{n-m},$$

где $C_n^m = \frac{n!}{m!(n-m)!}$, $n! = n \cdot (n-1) \cdots 2 \cdot 1$, $0! = 1$, $q = 1 - p = P(\bar{A})$ – вероятность неуспеха в одном испытании.

Вероятность того, что событие A в схеме Бернулли появится не менее m_1 раз и не более m_2 раз, равна $P_n(m_1 \leq m \leq m_2) = \sum_{k=m_1}^{m_2} C_n^k p^k q^{n-k}$.

Вероятность того, что в серии из n независимых испытаний событие A появится хотя бы один раз, равна $P_n(m \geq 1) = 1 - P_n(0) = 1 - q^n$.

8. Предельные теоремы в схеме Бернулли

При **больших** значениях n для вычисления вероятностей $P_n(m)$ используются приближенные формулы Пуассона и Муавра-Лапласа.

Если в схеме Бернулли вероятность p появления события A в каждом из n независимых испытаний **крайне мала**, а число испытаний n **достаточно велико**, то вероятность $P_n(m)$ вычисляется приближенно по **формуле Пуассона** (теорема Пуассона):

$$P_n(m) \approx \frac{a^m e^{-a}}{m!}, \quad a = np.$$

Формулу Пуассона применяют, когда событие A является **редким**, но количество испытаний n **велико** и **среднее число успехов** $a = np$ не значительно ($a \leq 10$).

Если в схеме Бернулли вероятность p появления события A близка к 1, а число испытаний n велико, для вычисления вероятности $P_n(m)$ также можно использовать формулу Пуассона (считая успехом событие \bar{A}).

Если в схеме Бернулли вероятность p появления события A в каждом из n независимых испытаний **существенно отличается от 0 и 1** (близко к $\frac{1}{2}$), а число испытаний n **достаточно велико**, то для вычисления вероятности $P_n(m)$ применяют приближенную **локальную формулу Муавра-Лапласа** (**локальная теорема Муавра-Лапласа**):

$$P_n(m) \approx \frac{1}{\sqrt{npq}} \Phi\left(\frac{m - np}{\sqrt{npq}}\right),$$

где $\Phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ – функция Гаусса, причем $\Phi(-x) = \Phi(x)$, на практике обычно полагают $\Phi(x) \approx 0$ при $x \geq 4$.

Если в схеме Бернулли вероятность p **существенно отличается от 0 и 1**, а n **достаточно велико**, то вероятность $P_n(m_1 \leq m < m_2)$, того, что в n независимых испытаниях событие A наступит не менее m_1 раз, но менее m_2 раз, вычисляется по **интегральной формуле Муавра-Лапласа** (**интегральная теорема Муавра-Лапласа**):

$$P_n(m_1 \leq m < m_2) \approx \Phi\left(\frac{m_2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{m_1 - np}{\sqrt{npq}}\right),$$

где $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt$ – функция Лапласа, причем $\Phi(-x) = -\Phi(x)$,

на практике обычно полагают $\Phi(x) \approx 0,5$ при $x \geq 5$.

Для функций $\Phi(x)$ и $\Phi(x)$ составлены таблицы значений. Формулы Муавра-Лапласа, как правило, используются, если $0,1 < p < 0,9$, и дают хорошие результаты, если $npq \geq 20$.

8. Понятие случайной величины. Способы задания случайных величин. Функция распределения и ее свойства.

Под **случайной величиной** (СВ) будем понимать величину, которая в результате случайного эксперимента принимает одно и только одно возможное значение, которое заранее неизвестно и зависит от случайных причин.

Примеры: а) число очков, выпавших при однократном бросании игральной кости, есть СВ, она может принять одно из значений: 1, 2, 3, 4, 5, 6;

б) число успехов в n испытаниях в схеме Бернулли – СВ, принимающая значения $0, 1, \dots, n$;

в) число бракованных изделий в данной партии – СВ, принимающая целые значения от 0 до n , где n – объем партии;

г) прирост веса домашнего животного за месяц есть СВ, которая может принять значение из некоторого промежутка.

Более строго, под **СВ понимают действительнозначную функцию ξ , определенную на множестве Ω элементарных событий, связанных с данным случаем экспериментом, и такую, что для любой системы B открытых интервалов, $B \subset \mathbb{R}$, существует $P(\omega \in \Omega : \xi(\omega) \in B)$ – вероятность того, что СВ ξ примет значение из множества B .**

Таким образом, для любой СВ ξ определена функция

$$F(x) = P(\xi < x), x \in \mathbb{R},$$

называемая ее **функцией распределения** и выражающая вероятность того, что СВ ξ примет значение, меньшее x . Под законом распределения СВ будем понимать любое правило, позволяющее найти функцию распределения этой СВ.

Основные свойства функции распределения СВ.

$$1. 0 \leq F(x) \leq 1, F(-\infty) = \lim_{x \rightarrow -\infty} F(x) = 0, F(+\infty) = \lim_{x \rightarrow +\infty} F(x) = 1.$$

2. $F(x)$ – неубывающая, непрерывная слева функция, т.е. $F(x_1) \leq F(x_2)$ при $x_1 < x_2$ и $F(x-0) = F(x)$, $x \in \mathbb{R}$.

$$3. P(\alpha \leq \xi < \beta) = F(\beta) - F(\alpha).$$

$$4. P(\xi = x_0) = F(x_0 + 0) - F(x_0).$$

9. Дискретные случайные величины, способы их задания. Примеры дискретных распределений.

Случайная величина называется **дискретной** (ДСВ), если множество ее возможных значений *конечно* или *счетно* (т. е. если все ее значения можно занумеровать).

Примеры. Дискретными СВ являются: число выпадений герба при n подбрасываниях монеты, число выстрелов до первого попадания в цель, число бракованных изделий в данной партии и т. д.

Для того чтобы задать ДСВ ξ , достаточно перечислить все ее возможные значения x_m , $m = 1, 2, \dots$, и указать, с какими вероятностями p_m она их принимает.

Закон распределения ДСВ ξ удобно задать в виде таблицы, называемой **рядом распределения** этой СВ:

| | | | | | |
|-------|-------|-------|-----|-------|-----|
| ξ | x_1 | x_2 | ... | x_m | ... |
| P | p_1 | p_2 | ... | p_m | ... |

(отметим, что $p_m \geq 0$, $p_1 + p_2 + \dots + p_m + \dots = 1$ – *условие контроля*).

Отсюда получаем **функцию распределения** ДСВ:

$$F(x) = P(\xi < x) = \sum_{x_i < x} p_i, \quad x \in \mathbb{R}.$$

График функции распределения ДСВ имеет ступенчатый вид, причем функция распределения терпит разрывы в точках x_m со скачками $p_m = P(\xi = x_m)$, $m = 1, 2, \dots$.

10. Непрерывные случайные величины, способы их задания. Плотность распределения непрерывной случайной величины и ее свойства.

Случайная величина называется **непрерывной** (НСВ), если ее функция распределения $F(x) = P(\xi < x)$ непрерывна на всей числовой оси. НСВ принимает все значения из некоторого интервала или системы интервалов на числовой оси. Вероятность того, что НСВ примет фиксированное значение, равна нулю, т. е. $P(\xi = x_0) = 0$.

Примеры. Непрерывными СВ являются, например, время безотказной работы прибора; дальность полета снаряда; прибыль фирмы; расход электроэнергии на предприятии за месяц; вес новорожденного; ошибка измерения и т. п.

Особый интерес вызывают НСВ, имеющие плотность распределения. Закон распределения такой НСВ обычно задают функцией или плотностью распределения.

Функция $p(x)$ называется **плотностью распределения вероятностей** НСВ ξ с функцией распределения $F(x)$, если

$$F(x) = \int_{-\infty}^x p(x)dx, \text{ откуда } p(x) = F'(x), x \in \mathbb{R}.$$

Основные свойства плотности распределения НСВ.

1. $p(x) \geq 0$ при всех $x \in \mathbb{R}$.

2. $\int_{-\infty}^{+\infty} p(x)dx = 1$.

Геометрически это означает, что график плотности распределения лежит не ниже оси Ox и площадь под графиком плотности равна единице.

3. Вероятности попадания НСВ ξ в интервал, отрезок или полуинтервал с одними и теми же концами одинаковы и равны

$$\begin{aligned} P(\alpha \leq \xi \leq \beta) &= P(\alpha < \xi \leq \beta) = P(\alpha < \xi < \beta) = \\ &= P(\alpha \leq \xi < \beta) = \int_{\alpha}^{\beta} p(x)dx = F(\beta) - F(\alpha). \end{aligned}$$

11. Числовые характеристики случайных величин. Свойства математического ожидания и дисперсии.

Опр. 1. Математическим ожиданием дискретной СВ ξ называется число, равное сумме произведений всех значений СВ ξ на соответствующие им вероятности:

$$M\xi = \sum_k x_k p_k = x_1 p_1 + x_2 p_2 + \dots + x_k p_k + \dots \quad (1)$$

(предполагается, что ряд в правой части этого равенства абсолютно сходится).

Опр. 2. Математическим ожиданием непрерывной СВ ξ называется число, равное

$$M\xi = \int_{-\infty}^{+\infty} xf(x)dx, \quad (2)$$

где $f(x)$ – плотность распределения вероятностей СВ ξ , при условии, что этот несобственный интеграл сходится абсолютно, т. е.

$$\int_{-\infty}^{+\infty} |x| f(x) dx < \infty.$$

Математическое ожидание $M\xi$ характеризует среднее значение СВ ξ (с учетом ее более и менее вероятных значений).

Замечание. Существуют случайные величины, не имеющие математического ожидания, так как интеграл (2) или ряд (1) в случае дискретной СВ, имеющей бесконечное множество значений, могут быть расходящимися.

Наиболее используемыми числовыми характеристиками СВ являются:

1) математическое ожидание $M\xi$, определенное выше, которое характеризует среднее значение (центр рассеивания) СВ ξ ;

2) дисперсия $D\xi = M(\xi - M\xi)^2$, которая характеризует величину рассеивания значений СВ вокруг ее математического ожидания;

3) среднее квадратическое отклонение $\sigma_\xi = \sqrt{D\xi}$, которое (в отличие от дисперсии) имеет размерность СВ ξ , что оказывается более удобным в приложениях ТВ, например, в математической статистике.

Приведем основные свойства.

Свойства математического ожидания:

1. Математическое ожидание постоянной равно этой постоянной: $Mc = c$, если $c = \text{const}$.

2. Постоянный множитель выносится за знак математического ожидания: $M(c\xi) = cM\xi$.

3. Математическое ожидание суммы СВ равно сумме их матема-

тических ожиданий: $M(\xi + \eta) = M\xi + M\eta$.

4. Математическое ожидание произведения *независимых* СВ равно произведению их математических ожиданий: $M(\xi\eta) = M\xi \cdot M\eta$. (СВ ξ и η называются *независимыми*, если для любых $x, y \in \mathbb{R}$ события $\{\xi < x\}$ и $\{\eta < y\}$ независимы.)

Свойства дисперсии:

1. Дисперсия постоянной равна нулю: $Dc=0$, если $c=\text{const}$.
2. Дисперсия неотрицательна: $D\xi \geq 0$.
3. Постоянный множитель выносится за знак дисперсии в квадрате: $D(c\xi) = c^2 D\xi$.
4. Дисперсия суммы *независимых* СВ равна сумме их дисперсий: $D(\xi + \eta) = D\xi + D\eta$.

5. Дисперсия разности *независимых* СВ равна *сумме* их дисперсий: $D(\xi - \eta) = D\xi + D\eta$.

Из других числовых характеристик СВ отметим:

$M\xi^k$ – начальные моменты k -го порядка,

$M(\xi - M\xi)^k$ – центральные моменты k -го порядка.

Таким образом, математическое ожидание является начальным моментом первого, а дисперсия – центральным моментом второго порядков.

В заключение приведем важнейшие числовые характеристики для основных законов распределения.

Числовые характеристики основных законов распределения

| № п/п | Распределение | $M\xi$ | $D\xi$ | σ_ξ |
|----------|--|---------------------|-----------------------|-------------------------|
| 1. | Биномиальное (с параметрами n и p) | np | npq | \sqrt{npq} |
| 2. | Пуассона (с параметром a) | a | a | a |
| 3. | Равномерное на $[a; b]$ | $\frac{a+b}{2}$ | $\frac{(b-a)^2}{12}$ | $\frac{b-a}{2\sqrt{3}}$ |
| 4. | Показательное (с параметром λ) | $\frac{1}{\lambda}$ | $\frac{1}{\lambda^2}$ | $\frac{1}{\lambda}$ |
| 5. | Нормальное (Гаусса) с параметрами a и σ | a | σ^2 | σ |

12. Биномиальное распределение, его числовые характеристики.

1. СВ ξ имеет *биномиальное распределение* с параметрами n и p , если она принимает значения $0, 1, 2, \dots, n$ с вероятностями

$$P(\xi = m) = C_n^m p^m q^{n-m}, \quad m = 0, 1, 2, \dots, n, \quad \text{где } 0 < p < 1, \quad q = 1 - p.$$

Биномиальный закон распределения имеет место в том случае, когда СВ ξ выражает число появлений события A (число успехов) при n независимых испытаниях в схеме Бернулли.

Математическое ожидание и дисперсия СВ ξ , распределенной по биномиальному закону, вычисляются по формулам: $M\xi = np$, $D\xi = npq$.

Доказательство. Представим СВ ξ , имеющую биномиальное распределение с параметрами n и p , как сумму СВ:

$$\xi = \xi_1 + \xi_2 + \dots + \xi_n,$$

где

$$\xi_i = \begin{cases} 1, & \text{если событие } A \text{ в } i\text{-м испытании произошло,} \\ 0, & \text{если событие } A \text{ в } i\text{-м испытании не произошло.} \end{cases}$$

Тогда $P(\xi_i = 1) = p$; $P(\xi_i = 0) = q$, т. е. ξ_i – бернуллиевские СВ с параметром p . При этом все слагаемые $\xi_1, \xi_2, \dots, \xi_n$ попарно независимы, поэтому

$$M\xi = M(\xi_1 + \xi_2 + \dots + \xi_n) = M\xi_1 + M\xi_2 + \dots + M\xi_n = p + p + \dots + p = np;$$

$$D\xi = D(\xi_1 + \xi_2 + \dots + \xi_n) = D\xi_1 + D\xi_2 + \dots + D\xi_n = pq + pq + \dots + pq = npq.$$

13. Распределение Пуассона, его числовые характеристики.

2. Дискретная СВ ξ имеет **распределение Пуассона** с параметром a , если она принимает значения $0, 1, 2, \dots, n, \dots$ с вероятностями

$$P(\xi = m) = \frac{a^m}{m!} e^{-a}, \quad m = 0, 1, 2, \dots, n, \dots$$

Математическое ожидание и дисперсия СВ ξ , распределенной по закону Пуассона, равны $M\xi = D\xi = a$.

Доказательство.

$$M\xi = \sum_{k=0}^{\infty} kp_k = \sum_{k=0}^{\infty} k \cdot \frac{a^k e^{-a}}{k!} = \sum_{k=1}^{\infty} \frac{a^k e^{-a}}{(k-1)!} = a e^{-a} \sum_{k=1}^{\infty} \frac{a^{k-1}}{(k-1)!} =$$

$$= a e^{-a} \left(\frac{a^0}{0!} + \frac{a^1}{1!} + \frac{a^2}{2!} + \dots + \frac{a^k}{k!} + \dots \right) = a e^{-a} e^a = a;$$

$$M(\xi^2) = \sum_{k=0}^{\infty} k^2 p_k = \sum_{k=0}^{\infty} k^2 \cdot \frac{a^k e^{-a}}{k!} = e^{-a} \sum_{k=1}^{\infty} \frac{k a^k}{(k-1)!} =$$

$$= e^{-a} \sum_{k=1}^{\infty} \frac{(k-1)+1}{(k-1)!} a^k = e^{-a} \sum_{k=2}^{\infty} \frac{a^k}{(k-2)!} + e^{-a} \sum_{k=1}^{\infty} \frac{a^k}{(k-1)!} =$$

$$= a^2 e^{-a} \sum_{m=0}^{\infty} \frac{a^m}{m!} + a = a^2 + a,$$

откуда $D\xi = M(\xi^2) - (M\xi)^2 = a^2 + a - a^2 = a$. \diamond

14. Непрерывное равномерное распределение, его числовые характеристики.

1. Непрерывное равномерное распределение.

Опр. 1. Непрерывная СВ ξ *распределена равномерно* на отрезке $[a; b]$, если ее плотность распределения постоянна на этом отрезке, а вне его равна нулю.

Таким образом, плотность распределения имеет вид

$$f(x) = \begin{cases} c & \text{при } x \in [a; b], \\ 0 & \text{при } x \notin [a; b], \end{cases}$$

причем значение константы c можно определить из условия нормировки: $\int_{-\infty}^{+\infty} f(x)dx = 1$. Вычисляя

$$\int_{-\infty}^{+\infty} f(x)dx = \int_{-\infty}^a 0dx + \int_a^b cdx + \int_b^{+\infty} 0dx = 0 + c(b-a) + 0 = c(b-a) = 1,$$

получим $c = \frac{1}{b-a}$. Следовательно, плотность распределения имеет вид

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{при } x \in [a; b], \\ 0 & \text{при } x \notin [a; b]; \end{cases}$$

график плотности распределения изображен на рис. 14.

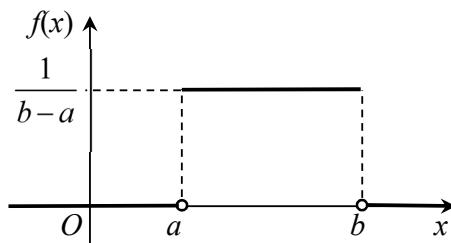


Рис. 14. Плотность равномерного на $[a; b]$ распределения

Для того чтобы СВ подчинялась закону равномерного распределения необходимо, чтобы ее значения лежали внутри некоторого определенного интервала и были равновероятны внутри этого интервала. Примером равномерно распределенной СВ может служить время ожидания пассажиром транспорта, курсирующего с определенным интервалом, или ошибка округления. Так, ошибка округления числа до ближайшего целого есть СВ, распределенная равномерно на промежутке $[-0,5; 0,5]$; если мы измеряем некоторую физическую величину, например, длину с точностью до 1 см, то ошибка округления этой величины (длины) будет распределена равномерно на $[-0,5 \text{ см}; 0,5 \text{ см}]$.

Утв. 1. Числовые характеристики равномерного распределения:

$$M\xi = \frac{a+b}{2}; \quad D\xi = \frac{(b-a)^2}{12}; \quad \sigma_\xi = \frac{b-a}{2\sqrt{3}}.$$

Доказательство.

$$\begin{aligned} M\xi &= \int_{-\infty}^{+\infty} xf(x)dx = \int_{-\infty}^a 0dx + \frac{1}{b-a} \int_a^b xdx + \int_b^{+\infty} 0dx = \\ &= 0 + \frac{1}{b-a} \left. \frac{x^2}{2} \right|_a^b + 0 = \frac{b^2 - a^2}{2(b-a)} = \frac{b+a}{2}; \\ M\xi^2 &= \int_{-\infty}^{+\infty} x^2 f(x)dx = \frac{1}{b-a} \int_a^b x^2 dx = \frac{1}{b-a} \left. \frac{x^3}{3} \right|_a^b = \frac{b^3 - a^3}{3(b-a)} = \frac{b^2 + ab + a^2}{3}, \end{aligned}$$

тогда

$$\begin{aligned} D\xi &= \frac{b^2 + ab + a^2}{3} - \left(\frac{b+a}{2} \right)^2 = \frac{4b^2 + 4ab + 4a^2 - 3b^2 - 6ab - 3a^2}{12} = \\ &= \frac{b^2 - 2ab + a^2}{12} = \frac{(b-a)^2}{12}, \end{aligned}$$

$$\text{а значит, } \sigma_\xi = \sqrt{\frac{(b-a)^2}{12}} = \frac{b-a}{2\sqrt{3}}. \quad \triangleleft$$

15. Показательное распределение, его числовые характеристики.

2. НСВ ξ имеет **показательное (экспоненциальное) распределение** с параметром $\lambda > 0$, если ее плотность распределения имеет вид

$$p(x) = \begin{cases} \lambda e^{-\lambda x} & \text{при } x \geq 0, \\ 0 & \text{при } x < 0. \end{cases}$$

Функция показательного распределения имеет вид

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{при } x \geq 0, \\ 0 & \text{при } x < 0. \end{cases}$$

Числовые характеристики показательного распределения:

$$M\xi = \frac{1}{\lambda}, \quad D\xi = \frac{1}{\lambda^2}, \quad \sigma_\xi = \frac{1}{\lambda}.$$

Показательное распределение является одним из основных в теории массового обслуживания и теории надежности. Примером СВ, имеющей показательное распределение, является время ожидания редких явлений: время между двумя вызовами на АТС, продолжительность безотказной работы приборов и т. д.

Доказательство. Найдем

$$\begin{aligned} M\xi &= \int_{-\infty}^{+\infty} xf(x) dx = \int_{-\infty}^0 0 dx + \int_0^{+\infty} x\lambda e^{-\lambda x} dx = \lim_{B \rightarrow +\infty} \int_0^B x\lambda e^{-\lambda x} dx = \\ &= \left| \begin{array}{l} u = x; \quad du = dx \\ dv = \lambda e^{-\lambda x} dx; \quad v = -e^{-\lambda x} \end{array} \right| = \lim_{B \rightarrow +\infty} \left(-xe^{-\lambda x} \Big|_0^B + \int_0^B e^{-\lambda x} dx \right) = \\ &= \lim_{B \rightarrow +\infty} \left(-B e^{-\lambda B} - \frac{1}{\lambda} e^{-\lambda x} \Big|_0^B \right) = \lim_{B \rightarrow +\infty} \left(-B e^{-\lambda B} - \frac{1}{\lambda} e^{-\lambda B} + \frac{1}{\lambda} \right) = \\ &= -\lim_{B \rightarrow +\infty} \frac{B}{e^{\lambda B}} - \lim_{B \rightarrow +\infty} \frac{1}{\lambda e^{\lambda B}} + \frac{1}{\lambda} = 0 - 0 + \frac{1}{\lambda} = \frac{1}{\lambda}. \end{aligned}$$

16. Нормальное распределение, его числовые характеристики. Правило трех сигм.

3. Распределение НСВ ξ называется **нормальным** (или *распределением Гаусса*) с параметрами a и $\sigma > 0$: $\xi \in N(a, \sigma)$, если плотность распределения вероятностей имеет вид

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}, \quad x \in (-\infty, +\infty).$$

Параметры a и σ имеют смысл математического ожидания и среднего квадратического отклонения СВ ξ : $M\xi = a$, $D\xi = \sigma^2$.

График плотности нормального распределения изображен на рис. 1 и называется кривой Гаусса.

Функция распределения СВ ξ , имеющей нормальное распределение с параметрами a и σ , выражается через функцию Лапласа

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt$$
 следующим образом:

$$F(x) = \frac{1}{2} + \Phi\left(\frac{x-a}{\sigma}\right),$$

а вероятность попадания СВ ξ на заданный интервал (α, β) вычисляется по формуле

$$P(\alpha < \xi < \beta) = \Phi\left(\frac{\beta-a}{\sigma}\right) - \Phi\left(\frac{\alpha-a}{\sigma}\right).$$

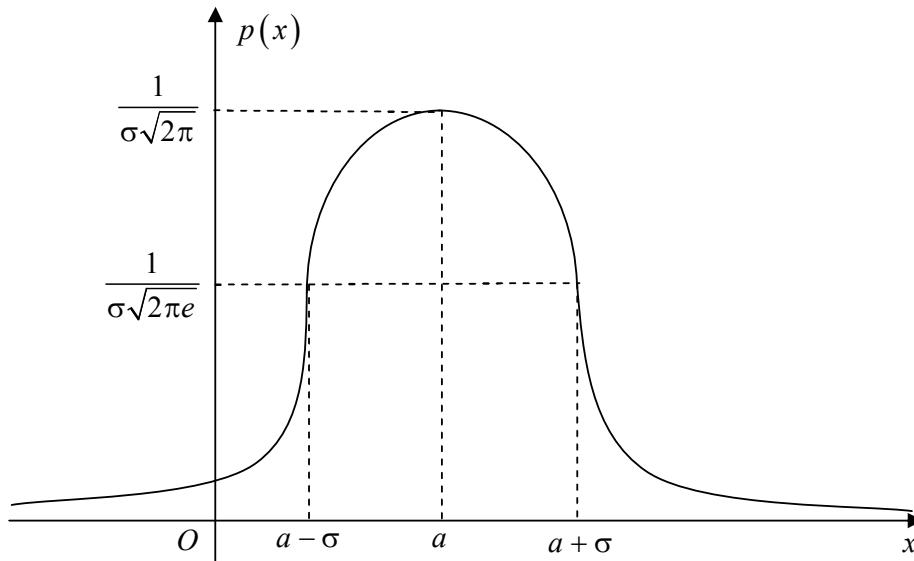


Рис. 1. График плотности нормального распределения

В силу непрерывности СВ эта формула справедлива как со строгими, так и с нестрогими знаками неравенств.

Вероятность того, что СВ ξ , распределенная нормально с параметрами a и σ , отклонится от своего математического ожидания менее, чем на δ , определяется соотношением

$$P(|\xi - a| < \delta) = 2\Phi\left(\frac{\delta}{\sigma}\right).$$

Полагая $\delta=3\sigma$, получим

$$P(|\xi - a| < 3\sigma) = 2\Phi(3) \approx 2 \cdot 0,49865 = 0,9973 \approx 1.$$

Правило «трех сигм» для нормального распределения. Если СВ ξ распределена нормально с параметрами a и σ , то попадание ее в интервал $(a - 3\sigma, a + 3\sigma)$ является практически достоверным событием и, стало быть, вероятность противоположного события ничтожно мала и на практике таким событием пренебрегают.

Нормальное распределение имеет большое теоретическое и практическое значение. В частности, считается, что погрешности измерения различных физических величин, ошибки, порожденные большим количеством случайных причин, распределены по нормальному закону. Кроме того, нормальный закон распределения является предельным законом, к которому приближаются другие законы распределения при весьма часто встречающихся типичных условиях, что делает нормальное распределение исключительным в ТВ и ее приложениях.

17. Неравенство Чебышева. Закон больших чисел и центральная предельная теорема теории вероятностей.

Т 1 (неравенство Чебышева). Для любой СВ ξ , имеющей конечные математическое ожидание и дисперсию, вероятность того, что отклонение СВ ξ от ее математического ожидания превзойдет по абсолютной величине положительное число ε , не больше дисперсии этой СВ, деленной на ε^2 :

$$P(|\xi - M\xi| > \varepsilon) \leq \frac{D\xi}{\varepsilon^2}.$$

Доказательство (для случая непрерывной СВ ξ). Пусть $f(x)$ – плотность распределения непрерывной СВ ξ . Тогда, разбивая область интегрирования на две области и отбрасывая один из интегралов как неотрицательный (за счет неотрицательности подынтегральной функции), получим следующую оценку дисперсии:

$$\begin{aligned} D\xi &= M(\xi - M\xi)^2 = \int_{-\infty}^{+\infty} (x - M\xi)^2 f(x) dx = \\ &= \int_{|x-M\xi|\leq\varepsilon} (x - M\xi)^2 f(x) dx + \int_{|x-M\xi|>\varepsilon} (x - M\xi)^2 f(x) dx \geq \\ &\geq 0 + \int_{|x-M\xi|>\varepsilon} (x - M\xi)^2 f(x) dx \geq \int_{|x-M\xi|>\varepsilon} \varepsilon^2 f(x) dx = \varepsilon^2 P(|\xi - M\xi| > \varepsilon). \end{aligned}$$

Выражая вероятность, получим требуемое неравенство.

Последовательность СВ $\xi_1, \xi_2, \dots, \xi_n, \dots$ сходится по вероятности к числу a : $\xi_n \xrightarrow{P} a$, если для любого $\varepsilon > 0$ вероятность события $\{\lvert \xi_n - a \rvert < \varepsilon\}$ при $n \rightarrow \infty$ стремится к единице, т. е.

$$\lim_{n \rightarrow \infty} P(\lvert \xi_n - a \rvert < \varepsilon) = 1.$$

Закон больших чисел в форме Я. Бернулли. Относительная частота появления события A в n независимых испытаниях, в каждом из которых это событие появляется с одной и той же вероятностью p , при неограниченном увеличении числа испытаний n сходится по вероятности к вероятности p этого события: $\frac{m}{n} \xrightarrow{P} p$ при $n \rightarrow \infty$.

Закон больших чисел в форме Бернулли является теоретическим обоснованием статистического метода задания вероятности, согласно которому вероятность события можно оценить относительной частотой $\frac{m}{n}$ появления этого события при достаточно большом числе n независимых испытаний.

Центральная предельная теорема

ЗБЧ устанавливает факт приближения среднего арифметического СВ к определенному числу. Оказывается, что при определенных условиях, а именно, если суммируемые СВ равноправны, никакая из них не является доминирующей, распределение среднего арифметического этих СВ сходится к нормальному распределению независимо от того, каков закон распределения слагаемых. В этом заключается смысл ЦПТ, и в этом – причина важности нормального распределения.

Сформулируем ЦПТ для частного случая – независимых одинаково распределенных СВ.

Пусть $\xi_1, \xi_2, \dots, \xi_n, \dots$ – взаимно независимые одинаково распределенные СВ, $M\xi_i = a$, $D\xi_i = \sigma^2$ для всех i . Найдем числовые характеристики СВ $S_n = \frac{\xi_1 + \xi_2 + \dots + \xi_n}{n}$:

$$MS_n = \frac{M\xi_1 + M\xi_2 + \dots + M\xi_n}{n} = \frac{a + a + \dots + a}{n} = a;$$

$$DS_n = \frac{D\xi_1 + D\xi_2 + \dots + D\xi_n}{n^2} = \frac{\sigma^2 + \sigma^2 + \dots + \sigma^2}{n^2} = \frac{\sigma^2}{n}.$$

Т 4 (ЦПТ для независимых одинаково распределенных СВ).

Пусть $\xi_1, \xi_2, \dots, \xi_n, \dots$ – взаимно независимые одинаково распределенные СВ, $M\xi_i = a$, $D\xi_i = \sigma^2$ для всех i . Тогда функция распределения СВ $\frac{S_n - MS_n}{\sqrt{DS_n}}$ сходится при $n \rightarrow \infty$ к функции стандартного

нормального распределения, т. е. при любом значении x

$$F_n(x) = P\left(\frac{S_n - MS_n}{\sqrt{DS_n}} < x\right) \rightarrow \frac{1}{2} + \Phi(x) \text{ при } n \rightarrow \infty.$$

Иными словами, среднее арифметическое $S_n = \frac{\xi_1 + \xi_2 + \dots + \xi_n}{n}$

независимых одинаково распределенных СВ с $M\xi_i = a$, $D\xi_i = \sigma^2$ имеет приближенно нормальное распределение с параметрами a и $\frac{\sigma}{\sqrt{n}}$.

Нормальный закон возникает во всех случаях, когда исследуемая СВ может быть представлена в виде суммы достаточно большого числа независимых (или слабо зависимых) элементарных слагаемых, каждое из которых в отдельности сравнительно мало влияет на сумму. Поэтому нормальный закон является самым распространенным из законов распределения.

Пусть производится измерение некоторой физической величины. Любое измерение дает приближенное значение, так как на результат измерения влияют очень многие независимые факторы: температура, влажность, колебания прибора и т.д. Каждый из факторов порождает ничтожно малую ошибку. Так как число факторов велико, то их совокупное действие порождает уже заметную «суммарную ошибку», которая имеет распределение близкое к нормальному распределению.

Следствиями ЦПТ являются рассмотренные ранее локальная и интегральная теоремы Муавра-Лапласа.

18. Двумерные случайные величины, способы их задания. Свойства функции распределения двумерной случайной величины. Свойства плотности распределения непрерывной двумерной случайной величины.

Опр. 1. Пусть имеется некоторое вероятностное пространство (Ω, \mathcal{F}, P) . Двумерной СВ $(\xi; \eta)$ называется совокупность двух числовых функций, заданных на одном и том же пространстве элементарных исходов Ω , если для любых действительных чисел x, y существует $P(\xi < x, \eta < y)$.

Опр. 2. Двумерная СВ $(\xi; \eta)$ называется *дискретной*, если обе ее составляющие ξ и η являются дискретными СВ.

Опр. 3. Двумерная СВ $(\xi; \eta)$ называется *непрерывной*, если обе ее составляющие ξ и η являются непрерывными СВ.

Отметим, что для наглядности значения двумерной СВ $(\xi; \eta)$ могут изображаться точками на плоскости Oxy . Дискретная двумерная СВ принимает конечное или счетное множество отдельных значений. Непрерывная СВ принимает значения из некоторой плоской области или нескольких областей.

Если одна из СВ дискретная, а другая непрерывная, то двумерная СВ относится к смешанному типу.

Совместная функция распределения СВ ξ и η

Универсальным способом задания двумерной СВ является функция распределения.

Опр. 4. Функция распределения двумерной СВ $(\xi; \eta)$ – это функция двух действительных переменных x и y , которая определяется с помощью равенства

$$F_{\xi; \eta}(x; y) = P(\xi < x; \eta < y). \quad (1)$$

Геометрически (1) означает вероятность попадания значения СВ в четверть плоскости левее и ниже точки с координатами $(x; y)$ (рис. 20).

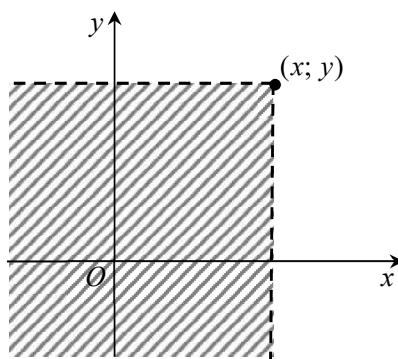


Рис. 20. К понятию функции распределения двумерной СВ

Свойства функции распределения двумерной СВ.

1. $0 \leq F(x; y) \leq 1$ при всех $(x; y)$.

2. Функция распределения является неубывающей по каждому из своих аргументов:

$$F(x_1; y) \leq F(x_2; y), \text{ если } x_1 < x_2;$$

$$F(x; y_1) \leq F(x; y_2), \text{ если } y_1 < y_2.$$

3. $F(-\infty; y) = F(x; -\infty) = F(-\infty; -\infty) = 0$.

4. $F(+\infty; +\infty) = 1$.

5. $F_{\xi; \eta}(x; +\infty) = F_\xi(x)$ – функция распределения СВ ξ ;

$F_{\xi; \eta}(+\infty; y) = F_\eta(y)$ – функция распределения СВ η .

6. Функция распределения непрерывна слева по каждому из своих аргументов.

Распределение непрерывной двумерной случайной величины может быть задано с помощью плотности распределения.

Опр. 5. Функция $f_{\xi; \eta}(x; y)$ называется **плотностью распределения** двумерной СВ $(\xi; \eta)$, если

$$F_{\xi; \eta}(x; y) = P(\xi < x; \eta < y) = \int_{-\infty}^x \int_{-\infty}^y f_{\xi; \eta}(x; y) dx dy.$$

Следовательно, плотность распределения двумерной СВ $(\xi; \eta)$ может быть найдена по формуле

$$f_{\xi; \eta}(x; y) = \frac{\partial^2 F_{\xi; \eta}(x; y)}{\partial x \partial y}. \quad (2)$$

Свойства плотности распределения.

1. $f(x; y) \geq 0$.

$$2. \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x; y) dx dy = 1.$$

3. Вероятность попадания СВ $(\xi; \eta)$ в область D равна

$$P((\xi; \eta) \in D) = \iint_D f(x; y) dx dy. \quad (3)$$

4. Плотности распределения составляющих двумерной СВ $(\xi; \eta)$:

$$f_{\xi}(x) = \int_{-\infty}^{+\infty} f_{\xi; \eta}(x; y) dy; \quad f_{\eta}(y) = \int_{-\infty}^{+\infty} f_{\xi; \eta}(x; y) dx.$$

Т 3. Если двумерная СВ $(\xi; \eta)$ имеет плотность распределения, то СВ ξ и η независимы тогда и только тогда, когда их совместная плотность распределения представима в виде произведения плотностей распределения этих СВ: $f_{\xi; \eta}(x; y) = f_{\xi}(x)f_{\eta}(y)$ для всех x и y .

19. Критерии независимости двух случайных величин.

Напомним, что две СВ ξ и η называются независимыми, если для любых числовых множеств X и Y события $\{\xi \in X\}$ и $\{\eta \in Y\}$ независимы, т. е. $P(\xi \in X, \eta \in Y) = P(\xi \in X)P(\eta \in Y)$.

Т 1. СВ ξ и η независимы тогда и только тогда, когда

$$F_{\xi; \eta}(x; y) = F_{\xi}(x)F_{\eta}(y).$$

для всех действительных x и y , т. е. их совместная функция распределения представима в виде произведения функций распределения этих СВ.

Таким образом, чтобы по таблице двумерного распределения найти законы распределения составляющих, нужно просуммировать вероятности по строкам – для одной СВ, по столбцам – для другой СВ.

Т 2. Дискретные СВ ξ и η независимы тогда и только тогда, когда $p_{ij}^* = p_i^* p_j^{**}$ для всех i, j .

20. Числовые характеристики двумерной случайной величины. Коэффициент корреляции, его свойства.

Основными *числовыми характеристиками* двумерной СВ $(\xi; \eta)$ являются математические ожидания и дисперсии ее составляющих, т. е. СВ ξ и η , а также корреляционный момент и коэффициент корреляции.

Запишем формулы для вычисления математических ожиданий и дисперсий СВ ξ и η , если известен закон распределения двумерной СВ $(\xi; \eta)$.

Для *дискретной* двумерной СВ $(\xi; \eta)$ с $p_{ij} = P(\xi = x_i; \eta = y_j)$, $1 \leq i \leq n, 1 \leq j \leq m$, математические ожидания СВ ξ и η равны соответственно

$$M\xi = \sum_{i=1}^n x_i p_i^*;$$

$$M\eta = \sum_{j=1}^m y_j p_j^{**},$$

где

$$p_i^* = P(\xi = x_i) = \sum_{j=1}^m p_{ij}; \quad p_j^{**} = P(\eta = y_j) = \sum_{i=1}^n p_{ij}.$$

Отсюда получим

$$M\xi = \sum_{i=1}^n \sum_{j=1}^m x_i p_{ij}; \quad M\eta = \sum_{i=1}^n \sum_{j=1}^m y_j p_{ij}.$$

Эти формулы можно обобщить в следующем утверждении.

Утв. 1. Для дискретной двумерной СВ $(\xi; \eta)$ с $p_{ij} = P(\xi = x_i; \eta = y_j)$, $1 \leq i \leq n, 1 \leq j \leq m$, при некоторых ограничениях на функцию $g(x; y)$ для математического ожидания от функции двух дискретных СВ имеет место формула

$$Mg(\xi; \eta) = \sum_{i=1}^n \sum_{j=1}^m g(x_i; y_j) p_{ij}.$$

Аналогично для непрерывных СВ.

Утв. 2. Для непрерывной двумерной СВ $(\xi; \eta)$ с плотностью распределения $f_{\xi; \eta}(x; y)$ при некоторых ограничениях на функцию $g(x; y)$ для математического ожидания от функции двух непрерывных СВ имеет место формула

$$Mg(\xi; \eta) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x; y) f_{\xi; \eta}(x; y) dx dy.$$

Следовательно,

$$M\xi = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x f_{\xi; \eta}(x; y) dx dy;$$

$$M\eta = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} y f_{\xi; \eta}(x; y) dx dy.$$

Дисперсии $D\xi$ и $D\eta$ можно найти по формулам $D\xi = M(\xi - M\xi)^2$ или $D\xi = M(\xi^2) - (M\xi)^2$.

Математические ожидания $M\xi$, $M\eta$ и дисперсии $D\xi$, $D\eta$ характеризуют среднее значение и рассеяние каждой из составляющих двумерной СВ.

Для характеристики степени зависимости двух СВ вводится новая числовая характеристика.

$$r_{\xi; \eta} = \frac{\text{cov}(\xi; \eta)}{\sqrt{D\xi D\eta}}.$$

Свойства коэффициента корреляции.

$$1. \boxed{-1 \leq r_{\xi; \eta} \leq 1.}$$

Доказательство. Пусть ξ и η – две СВ, не обязательно независимые. Рассмотрим при произвольном постоянном λ дисперсию СВ $\lambda\xi + \eta$:

$$\begin{aligned} D(\lambda\xi + \eta) &= M(\lambda\xi + \eta - M(\lambda\xi + \eta))^2 = \\ &= M(\lambda\xi + \eta - (\lambda M\xi + M\eta))^2 = M(\lambda(\xi - M\xi) + (\eta - M\eta))^2 = \\ &= M(\lambda^2(\xi - M\xi)^2 + 2\lambda(\xi - M\xi)(\eta - M\eta) + (\eta - M\eta)^2) = \\ &= \lambda^2 M(\xi - M\xi)^2 + 2\lambda M(\xi - M\xi)(\eta - M\eta) + M(\eta - M\eta)^2 = \\ &= \lambda^2 D\xi + 2\lambda \text{cov}(\xi; \eta) + D\eta. \end{aligned}$$

Поскольку дисперсия всегда $D(\lambda\xi + \eta) \geq 0$, то

$$\lambda^2 D\xi + 2\lambda \text{cov}(\xi; \eta) + D\eta \geq 0$$

при всех λ .

С другой стороны, выражение в левой части неравенства – это квадратный трехчлен относительно λ с положительным коэффициентом при λ^2 , поэтому для того, чтобы неравенство было верно при всех λ , дискриминант квадратного трехчлена должен быть меньше либо равен 0:

$$D = (2 \text{cov}(\xi; \eta))^2 - 4D\xi D\eta \leq 0;$$

$$(\text{cov}(\xi; \eta))^2 \leq D\xi D\eta;$$

$$r_{\xi; \eta}^2 = \frac{(\text{cov}(\xi; \eta))^2}{D\xi D\eta} \leq 1.$$

Следовательно, $|r_{\xi; \eta}| \leq 1$. \triangleleft

2. Если СВ ξ и η независимы, то $r_{\xi; \eta} = 0$.

Обратное утверждение неверно: если $r_{\xi; \eta} = 0$, то СВ ξ и η могут быть как зависимыми, так и независимыми.

3. СВ ξ и η связаны линейной зависимостью в том и только том случае, если $r_{\xi; \eta} = \pm 1$:

$$\eta = k\xi + b, k > 0 \Leftrightarrow r_{\xi; \eta} = +1;$$

$$\eta = k\xi + b, k < 0 \Leftrightarrow r_{\xi; \eta} = -1.$$

Доказательство. Докажем утверждение в одну сторону: если СВ ξ и η связаны линейной зависимостью, то $r_{\xi; \eta} = \pm 1$.

Пусть $\eta = k\xi + b$, тогда

$$\begin{aligned} \text{cov}(\xi; \eta) &= M(\xi - M\xi)(\eta - M\eta) = M(\xi - M\xi)(k\xi + b - kM\xi - b) = \\ &= M(\xi - M\xi)k(\xi - M\xi) = k \text{cov}(\xi; \xi) = kD\xi; \end{aligned}$$

$$D\eta = D(k\xi + b) = D(k\xi) = k^2 D\xi.$$

Таким образом,

$$r_{\xi; \eta} = \frac{\text{cov}(\xi; \eta)}{\sqrt{D\xi D\eta}} = \frac{kD\xi}{\sqrt{D\xi k^2 D\xi}} = \frac{kD\xi}{|k| D\xi} = \frac{k}{|k|} = \begin{cases} 1, & \text{если } k > 0, \\ -1, & \text{если } k < 0. \end{cases} \triangleleft$$

Итак, коэффициент корреляции $r_{\xi; \eta}$ показывает степень линейной зависимости между СВ ξ и η .

Особое место среди законов распределения двумерных СВ занимает двумерное нормально распределение.

21. Задачи математической статистики. Генеральная и выборочная совокупности. Вариационный ряд. Статистический ряд. Полигон и гистограмма. Эмпирическая функция распределения и ее свойства.

Теория вероятностей и математическая статистика занимаются анализом закономерностей случайных массовых явлений. В теории вероятностей определяются вероятности тех или иных событий по известным вероятностям более простых событий, числовые характеристики случайных величин или вероятности, связанные с этими величинами, по известным законам распределения этих случайных величин. На практике для нахождения законов распределения случайных величин необходимо использовать экспериментальные данные. **Основной задачей математической статистики** является разработка методов получения вероятностных характеристик случайных явлений на основе результатов эксперимента.

Исходными понятиями математической статистики являются понятия генеральной и выборочной совокупностей.

Выборка (случайная выборка, выборочная совокупность) - множество значений результатов наблюдений над одной и той же случайной величиной при одних и тех же условиях. Элементы выборки называются **выборочными значениями**. Количество проведенных наблюдений называется **объемом выборки**.

Генеральной совокупностью называется множество всех возможных наблюдений над случайной величиной при данном комплексе условий.

В большинстве случаев генеральная совокупность бесконечна (можно производить сколь угодно много наблюдений).

В задаче контроля качества данной партии товаров объем генеральной совокупности равен объему этой партии. Если обследование

всей партии невозможно, то о качестве партии судят по случайной выборке товаров из этой партии.

Назначение статистических методов в том, чтобы по выборке ограниченного объема сделать вывод о свойствах генеральной совокупности в целом.

Для того, чтобы по данным выборки можно было достаточно уверенно судить об интересующем нас признаке генеральной совокупности, необходимо, чтобы объекты выборки «правильно» его представляли, т.е. выборка должна быть *репрезентативной* (представительной). Считается, что это требование выполняется, если объем выборки достаточно велик и все объекты генеральной совокупности имеют одинаковую вероятность попасть в выборку, т.е. при отборе сохраняется принцип случайности. Такую выборку называют *случайной выборкой*.

2. Статистический ряд и его графическое изображение

Пусть имеется выборка объема n : x_1, x_2, \dots, x_n .

Вариационным рядом выборки x_1, x_2, \dots, x_n называется способ её записи, при котором её элементы упорядочены (как правило, в порядке не убывания): $x_1 \leq x_2 \leq \dots \leq x_n$. Разность ω между максимальным и минимальным элементами называется *размахом выборки*:

$$\omega = x_{\max} - x_{\min}.$$

Как правило, некоторые выборочные значения могут совпадать, поэтому часто выборку представляют в виде статистического ряда.

Пусть в выборке элемент x_i встречается n_i раз. Число n_i называется **частотой** выборочного значения x_i , а $\frac{n_i}{n}$ – **относительной частотой**. Очевидно, что $\sum_{i=1}^k n_i = n$, где k – число различных элементов выборки. Последовательность пар $(x_i^*; n_i)$, где $x_1^*, x_2^*, \dots, x_k^*$ – различные выборочные значения, а n_1, n_2, \dots, n_k – соответствующие им частоты, называется **статистическим рядом**. Обычно статистический ряд записывается в виде таблицы, первая строка которой содержит различные выборочные значения x_i^* , а вторая – их частоты n_i .

При большом объёме (больше 30) выборки её элементы объединяют в группы (разряды), представляя результаты опытов в виде **интервального (группированного) статистического ряда**. Для этого интервал, содержащий все элементы выборки, разбивают на k непере-

секающихся интервалов. Число интервалов выбирается произвольно и, как правило, $5 \div 10 \leq k \leq 20 \div 25$. Вычисления значительно упрощаются, если интервалы имеют одинаковую длину $h \approx \frac{\omega}{k}$. В дальнейшем будет рассматриваться именно этот случай. После того, как частичные интервалы выбраны, определяют частоты n_i – количество элементов выборки, попавших в i -й интервал (элемент, совпадающий с верхней границей интервала, относится к последующему интервалу) и относительные частоты $\frac{n_i}{n}$. Полученные данные сводятся в таблицу:

Интервальный статистический ряд

| | | | | |
|---|-----------------|-----------------|-----|------------------|
| Интервалы наблюдаемых значений СВ ξ | $[x_0; x_1)$ | $[x_1; x_2)$ | ... | $[x_{k-1}; x_k]$ |
| Середины интервалов | x_1^* | x_2^* | ... | x_k^* |
| Частоты | n_1 | n_2 | ... | n_k |
| Относительные частоты | $\frac{n_1}{n}$ | $\frac{n_2}{n}$ | ... | $\frac{n_k}{n}$ |

В ряде случаев для наглядного представления выборки используют полигон и гистограмму относительных частот (частот).

Полигоном частот группированной выборки называется ломаная с вершинами в точках $(x_i^*; n_i)$, $i = \overline{1, k}$, а **полигон относительных частот** – ломаная линия с вершинами в точках $(x_i^*; \frac{n_i}{n})$, $i = \overline{1, k}$.

Гистограммой относительных частот (частот) группированной выборки называют ступенчатую фигуру, составленную из прямоугольников, построенных на интервалах группировки так, что площадь каждого прямоугольника равна соответствующей данному интервалу относительной частоте (частоте). Площадь гистограммы относительных частот равна 1.

При достаточно большом объеме выборки и достаточно малых интервалах группировки гистограмма относительных частот является статистическим аналогом плотности распределения наблюдаемой случайной величины. Поэтому по виду гистограммы можно выдвинуть предположение (гипотезу) о распределении изучаемой случайной величины.

Эмпирической функцией распределения называется функция $F^*(x)$, определяющая для каждого значения x относительную частоту наблюдения значений, меньших x :

$$F^*(x) = \sum_{x_i^* < x} \frac{n_i}{n}.$$

Основное значение эмпирической функции распределения в том, что она используется в качестве оценки теоретической функции распределения $F(x) = P(\xi < x)$ наблюдаемой случайной величины ξ и обладает всеми свойствами функции распределения дискретной случайной величины:

- 1) $0 \leq F^*(x) \leq 1$;
- 2) $F^*(x)$ – неубывающая непрерывная слева кусочно-постоянная функция;
- 3) если x_1 – наименьшее, а x_n – наибольшее значения статистического ряда, то $F^*(x) = 0$ при $x \leq x_1$ и $F^*(x) = 1$ при $x > x_n$.

Эмпирическая функция распределения $F^*(x)$ является случайной: для разных выборок она получается разной. Если график $F^*(x)$ строится по группированным данным, то скачки происходят в точках, соответствующих серединам интервалов группировки.

22. Точечное оценивание параметров распределения. Свойства точечных оценок. Несмешенные оценки математического ожидания и дисперсии.

Выборка представляет собой ряд наблюдений над одной и той же случайной величиной. Для содержательного статистического анализа экспериментальных данных необходимо знать распределение этой величины.

Во многих случаях можно считать, что наблюдаемая величина имеет нормальное распределение. Например, при измерениях одного и того же показателя на нескольких однотипных объектах колебания результатов будут вызваны незначительными случайными погрешностями в технологии изготовления или измерения. Если случайные колебания значений некоторой величины вызваны большим числом случайных причин, более или менее равноправных, то на основании центральной предельной теоремы теории вероятностей можно считать, что эта величина имеет нормальное распределение.

Нормальное распределение полностью определяется двумя параметрами – математическим ожиданием m и дисперсией σ^2 . Математическое ожидание и дисперсия являются основными числовыми характеристиками любой случайной величины. Математическое ожидание – это в некотором смысле (т. е. с учетом большей и меньшей вероятности различных значений) среднее значение случайной величины. Дисперсия характеризует разброс значений случайной величины относительно ее математического ожидания. Поэтому при статистическом анализе выборки в первую очередь стремятся оценить математическое ожидание и дисперсию.

Пусть имеется выборка объема n : x_1, x_2, \dots, x_n . По результатам этого ограниченного числа наблюдений невозможно вычислить числовые характеристики наблюдаемой случайной величины, а можно только оценить их.

Одна из задач математической статистики состоит в нахождении оценок неизвестных параметров по выборке. В качестве оценки параметра берут ту или иную функцию $\hat{\theta}_n = \hat{\theta}_n(x_1, x_2, \dots, x_n)$ выборки (выборочных значений), которая называется **статистикой** или **выборочной функцией**.

Точечной оценкой параметра θ называется любая статистика $\hat{\theta}_n$, предназначенная для оценки этого параметра и определяемая одним числом. Подчеркнем, что точечная оценка практически никогда не совпадает с истинным значением параметра, она может только оценивать его с большей или меньшей точностью.

Для любого параметра можно предложить разные оценки. Так, в качестве оценки для математического ожидания можно использовать первый элемент выборки, среднее арифметическое наибольшего и наименьшего элементов выборки, среднее арифметическое всех элементов выборки и т. д.

Задача статистического оценивания параметров заключается в том, чтобы из всего множества оценок выбрать в некотором смысле наилучшую. Это означает, что распределение случайной величины $\hat{\theta}_n(x_1, x_2, \dots, x_n)$ должно концентрироваться около истинного значения параметра θ .

Замечание. Если, имея выборку x_1, x_2, \dots, x_n значений некоторой случайной величины, повторно провести n независимых наблюдений над этой случайной величиной, то новая выборка x'_1, x'_2, \dots, x'_n , вообще говоря, не будет совпадать с первоначальной. Поэтому выбороч-

ные значения можно рассматривать как случайные величины. **Основное предположение математической статистики**: выборочные значения x_1, x_2, \dots, x_n являются независимыми в совокупности одинаково распределёнными случайными величинами. Следовательно, любая статистика и любая оценка $\hat{\theta}_n(x_1, x_2, \dots, x_n)$ также являются случайными величинами.

Качество точечной оценки характеризуется следующими основными свойствами.

1. Оценка $\hat{\theta}$ называется **несмешённой**, если её математическое ожидание равно оцениваемому параметру: $M[\hat{\theta}] = \theta$. Разность $M[\hat{\theta}] - \theta$ называется **смещением**.

Требование несмешенности гарантирует отсутствие систематических ошибок при оценивании. Оно особенно важно при малом числе наблюдений (в случае выборок объема не более 30).

2. Оценка $\hat{\theta}_n$ называется **состоятельной**, если при увеличении объема выборки n оценка $\hat{\theta}_n$ сходится по вероятности к θ :

$$\lim_{n \rightarrow \infty} P(|\theta - \hat{\theta}_n| < \varepsilon) = 1.$$

Это свойство означает, что при большом объеме выборки практически достоверно, что $\hat{\theta}_n \approx \theta$. Чем больше объем выборки, тем более точные оценки можно получить.

3. Пусть $\hat{\theta}_1$ и $\hat{\theta}_2$ – две различные **несмешённые** оценки параметра. Если для дисперсий $D[\hat{\theta}_1]$ и $D[\hat{\theta}_2]$ выполняется условие $D[\hat{\theta}_1] < D[\hat{\theta}_2]$, то говорят, что оценка $\hat{\theta}_1$ более эффективна, чем оценка $\hat{\theta}_2$. Оценка с наименьшей дисперсией называется **эффективной**.

Это означает, что распределение эффективной оценки наиболее тесно сконцентрировано около истинного значения параметра.

Кроме этих свойств имеются и другие. К сожалению, не всегда можно найти статистики, которые имели бы все указанные свойства.

Формулы для расчёта основных числовых характеристик выборки

| Для не группированной выборки | Для группированного статистического ряда |
|---|--|
| Выборочное среднее | |
| $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ | $\bar{x} = \frac{1}{n} \sum_{i=1}^k x_i^* n_i$ |
| Выборочная дисперсия | |
| $D_B = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$ | $D_B = \frac{1}{n} \sum_{i=1}^k (x_i^* - \bar{x})^2 n_i$ |
| $D_B = \frac{1}{n} \sum_{i=1}^n x_i^2 - (\bar{x})^2$ | $D_B = \frac{1}{n} \sum_{i=1}^k (x_i^*)^2 n_i - (\bar{x})^2$ |
| Несмешённая оценка дисперсии | |
| $s^2 = \frac{n}{n-1} D_B$ | |
| $s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$ | $s^2 = \frac{1}{n-1} \sum_{i=1}^k (x_i^* - \bar{x})^2 n_i$ |
| $s^2 = \frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 - n \bar{x}^2 \right)$ | $s^2 = \frac{1}{n-1} \sum_{i=1}^k (x_i^*)^2 n_i - \frac{n}{n-1} (\bar{x})^2$ |

Выборочное среднее $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ (среднее арифметическое элементов выборки) характеризует центр распределения (рассеивания) изучаемой случайной величины и является несмешённой и состоятельной оценкой, а в случае выборки из нормального распределения также и эффективной оценкой для математического ожидания наблюдаемой случайной величины.

Выборочная дисперсия D_B характеризует степень разброса (рассеяния) выборочных значений относительно среднего и является состоятельной, но смешённой (дает заниженное значение) оценкой дисперсии изучаемой случайной величины. В связи с этим вместо нее вводится **несмешенная оценка дисперсии** $s^2 = \frac{n}{n-1} D_B$.

23. Интервальные оценки параметров генеральной совокупности. Доверительная вероятность.

Точечные оценки не дают информации о степени близости оценки к истинному значению оцениваемого параметра. Чтобы получить информацию о точности и надежности оценки, используют интервальные оценки.

Интервальной оценкой параметра θ называется интервал, границы которого $\hat{\theta}_1 = \hat{\theta}_1(x_1, x_2, \dots, x_n)$ и $\hat{\theta}_2 = \hat{\theta}_2(x_1, x_2, \dots, x_n)$ являются функциями выборочных значений и который с заданной вероятностью γ накрывает истинное значение оцениваемого параметра θ :

$$P(\hat{\theta}_1(x_1, x_2, \dots, x_n) < \theta < \hat{\theta}_2(x_1, x_2, \dots, x_n)) = \gamma.$$

Интервал $(\hat{\theta}_1; \hat{\theta}_2)$ называется **доверительным интервалом**; число γ - **доверительной вероятностью** или **надёжностью** интервальной оценки; значение $\alpha = 1 - \gamma$ - **уровнем значимости**.

В практике важную роль играет величина (длина) доверительного интервала, поскольку чем меньше его длина, тем точнее оценка. Если длина доверительного интервала достаточно велика, то оценка малопригодна для практики.

Величина доверительного интервала существенно зависит от объема выборки (уменьшается с ростом n , т. е. чем больше объем выборки, тем более точную оценку можно получить) и от доверительной вероятности γ (величина доверительного интервала увеличивается с приближением γ к 1, т. е. чем более надежный вывод мы хотим получить, тем меньшую точность мы можем гарантировать).

Выбор доверительной вероятности определяется конкретными условиями. Обычно используются значения 0,90; 0,95; 0,99; 0,9973, т. е. такие, чтобы получить интервал, который с большой вероятностью накроет истинное значение оцениваемого параметра.

Доверительный интервал для математического ожидания m в случае выборки из нормального распределения с известной дисперсией σ^2 определяется соотношением

$$\bar{x} - u_\alpha \frac{\sigma}{\sqrt{n}} < m < \bar{x} + u_\alpha \frac{\sigma}{\sqrt{n}},$$

где n – объем выборки; \bar{x} – выборочное среднее; α – уровень значимости; u_α – квантиль нормального распределения, удовлетворяющая уравнению $\Phi(u_\alpha) = \gamma/2$ и определяемая из таблицы функции Лапласа;

$$u_\alpha \frac{\sigma}{\sqrt{n}} = \varepsilon \quad \text{точность оценки.}$$

Доверительный интервал для математического ожидания m в случае выборки из нормального распределения с неизвестной дисперсией σ^2 определяется формулой

$$\bar{x} - t_{\alpha; n-1} \frac{s}{\sqrt{n}} < m < \bar{x} + t_{\alpha; n-1} \frac{s}{\sqrt{n}},$$

где n – объем выборки; \bar{x} – выборочное среднее; s^2 – несмещенная оценка дисперсии; α – уровень значимости, $t_{\alpha; n-1}$ – квантиль распределения Стьюдента, удовлетворяющая уравнению $P(|t_{n-1}| \geq t_{\alpha; n-1}) = \alpha$ для случайной величины t_{n-1} , имеющей распределение Стьюдента с числом степеней свободы $k = n - 1$.

Доверительный интервал для дисперсии σ^2 в случае выборки из нормального распределения с неизвестным математическим ожиданием определяется соотношением

$$\frac{s^2(n-1)}{\chi_{\frac{\alpha}{2}; n-1}^2} < \sigma^2 < \frac{s^2(n-1)}{\chi_{1-\frac{\alpha}{2}; n-1}^2},$$

где n – объем выборки; s^2 – несмещенная оценка дисперсии; α – уровень значимости; $\chi_{\frac{\alpha}{2}; n-1}^2$ и $\chi_{1-\frac{\alpha}{2}; n-1}^2$ – квантили распределения χ^2 с числом степеней свободы $k = n - 1$.

24. Построение доверительного интервала для математического ожидания нормально распределенной генеральной совокупности.

Утв. 1. Пусть имеется выборка объема n из нормального распределения с математическим ожиданием a и дисперсией σ^2 , т. е. $x_1, x_2, \dots, x_n \sim \mathcal{N}(a; \sigma^2)$. Тогда статистика \bar{x} распределена по нормальному закону с параметрами a и $\frac{\sigma}{\sqrt{n}}$, а статистика $\frac{\bar{x} - a}{\sigma / \sqrt{n}}$ имеет стандартное нормальное распределение:

$$\bar{x} \sim \mathcal{N}\left(a; \frac{\sigma}{\sqrt{n}}\right); \quad \frac{\bar{x} - a}{\sigma / \sqrt{n}} \sim \mathcal{N}(0; 1).$$

Отметим, что

$$M\bar{x} = M \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{n} \sum_{i=1}^n Mx_i = \frac{1}{n} \sum_{i=1}^n a = \frac{1}{n} na = a;$$

$$D\bar{x} = D \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{n^2} \sum_{i=1}^n Dx_i = \frac{1}{n^2} \sum_{i=1}^n \sigma^2 = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}.$$

Это означает, в частности, что \bar{x} является более точной, чем одиночное наблюдение, оценкой для математического ожидания,

поскольку чем меньше дисперсия, т. е. разброс значений, тем точнее оценка.

Утв. 2. Доверительный интервал для математического ожидания a в случае выборки из нормального распределения с известной дисперсией σ^2 определяется соотношением

$$P\left(\bar{x} - u_\alpha \frac{\sigma}{\sqrt{n}} < a < \bar{x} + u_\alpha \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha, \quad (1)$$

где n – объем выборки; \bar{x} – выборочное среднее; α – уровень значимости; u_α – квантиль нормального распределения уровня α , т. е. такое число, что для СВ $\xi \sim \mathcal{N}(0; 1)$, имеющей стандартное нормальное распределение, $P(|\xi| \geq u_\alpha) = \alpha$.

Квантиль u_α определяется по таблице функции Лапласа из соотношения $\Phi(u_\alpha) = \frac{1-\alpha}{2}$.

Формула (1) означает, что при достаточно большом количестве выборок одного и того же объема n примерно в $100(1-\alpha)\%$

выборок интервал $\left(\bar{x} - u_\alpha \frac{\sigma}{\sqrt{n}}, \bar{x} + u_\alpha \frac{\sigma}{\sqrt{n}}\right)$ накрывает истинное

значение математического ожидания a .

Доказательство. Из утверждения 1 следует, что

$$P\left(\left|\frac{\bar{x} - a}{\sigma / \sqrt{n}}\right| < u_\alpha\right) = 1 - \alpha;$$

$$P\left(-u_\alpha < \frac{a - \bar{x}}{\sigma / \sqrt{n}} < u_\alpha\right) = 1 - \alpha;$$

$$P\left(-u_\alpha \frac{\sigma}{\sqrt{n}} < a - \bar{x} < u_\alpha \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha;$$

$$P\left(\bar{x} - u_\alpha \frac{\sigma}{\sqrt{n}} < a < \bar{x} + u_\alpha \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha. \triangleleft$$

25. Построение доверительного интервала для дисперсии нормально распределенной генеральной совокупности.

Утв. 3. Доверительный интервал для дисперсии σ^2 в случае выборки из нормального распределения с *неизвестным* математическим ожиданием a определяется соотношением

$$P\left(\frac{(n-1)s^2}{\chi_{\alpha/2; n-1}^2} < \sigma^2 < \frac{(n-1)s^2}{\chi_{1-\alpha/2; n-1}^2}\right) = 1 - \alpha,$$

где n – объем выборки; s^2 – несмещенная оценка дисперсии; α – уровень значимости; $\chi_{\alpha/2; n-1}^2$ и $\chi_{1-\alpha/2; n-1}^2$ – квантили распределения χ^2 с числом степеней свободы $k = n - 1$, определяемые соотношением $P(\xi \geq \chi_{\alpha; n-1}^2) = \alpha$ для СВ ξ , имеющей распределение χ^2 с числом степеней свободы $k = n - 1$.

Квантили $\chi_{\alpha/2; n-1}^2$ и $\chi_{1-\alpha/2; n-1}^2$ определяются по таблице распределения χ^2 .

Доказательство. Из утверждения 3 следует, что для выборки из нормального распределения с неизвестным математическим ожиданием

$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$. В силу несимметричности графика плотности распределения χ^2 для построения доверительного интервала будут использованы две квантили $\chi^2_{\alpha/2; n-1}$ и $\chi^2_{1-\alpha/2; n-1}$ (см. рис. 7), такие, что для СВ $\xi \sim \chi^2_{n-1}$ имеют место соотношения $P(\xi \geq \chi^2_{\alpha/2; n-1}) = \frac{\alpha}{2}$ и $P(\xi \leq \chi^2_{1-\alpha/2; n-1}) = \frac{\alpha}{2}$.

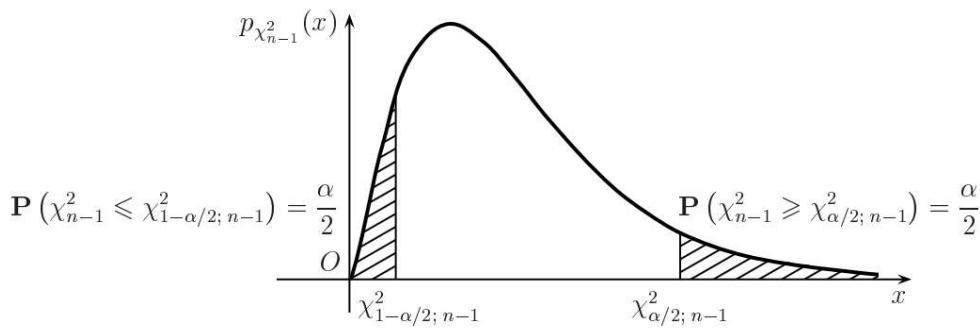


Рис. 7. К построению доверительного интервала для дисперсии в случае выборки из нормального распределения с неизвестным математическим ожиданием

Тогда

$$P\left(\chi^2_{1-\alpha/2; n-1} < \frac{(n-1)s^2}{\sigma^2} < \chi^2_{\alpha/2; n-1}\right) = 1 - \alpha;$$

$$P\left(\frac{1}{\chi^2_{\alpha/2; n-1}} < \frac{\sigma^2}{(n-1)s^2} < \frac{1}{\chi^2_{1-\alpha/2; n-1}}\right) = 1 - \alpha;$$

$$P\left(\frac{(n-1)s^2}{\chi^2_{\alpha/2; n-1}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\alpha/2; n-1}}\right) = 1 - \alpha. \triangleleft$$

26. Основные понятия теории проверки гипотез. Простая и сложная гипотезы. Нулевая и альтернативная гипотезы. Статистический критерий. Область принятия гипотезы и критическая область. Ошибки первого и второго родов. Уровень значимости и мощность критерия. Двусторонняя и односторонняя критические области.

Опр. 1. Статистической гипотезой называется любое предположение о виде (**непараметрическая гипотеза**) или параметрах (**параметрическая гипотеза**) неизвестного распределения.

Опр. 2. Статистическая гипотеза называется **простой**, если она полностью определяет функцию распределения. В противном случае гипотеза называется **сложной**.

Пример 1. Предположим, что введен новый способ производства некоторого товара. Для определения качества товара измеряется некоторая его характеристика $\xi \sim \mathcal{N}(a_0; \sigma_0)$, где a_0, σ_0 известны. Если необходимо выяснить, как новый способ производства влияет на качество товара, можно выдвинуть, например, такие гипотезы:

$H_1 : a = a_0, \sigma = \sigma_0$, т. е. распределение СВ ξ не изменилось после изменения процесса производства;

$H_2 : a > a_0, \sigma = \sigma_0$, т. е. увеличилось среднее значение показателя качества;

$H_3 : a = a_0, \sigma < \sigma_0$, т. е. разброс значений показателя качества стал меньше.

Гипотеза H_1 является простой, а гипотезы H_2 и H_3 – сложными. •

Опр. 3. Проверяемую гипотезу обычно называют **нулевой** и обозначают H_0 . Наряду с нулевой рассматривают **альтернативную**, или **конкурирующую**, гипотезу H_a (или H_1 , или \bar{H}).

Опр. 4. Правило, которое позволяет по выборке принять или отвергнуть проверяемую гипотезу, называется **критерием проверки статистической гипотезы (статистическим критерием)**.

Замечание. Статистическими методами *нельзя доказать* правильность гипотезы. Критерий проверки статистической гипотезы позволяет отбросить гипотезу как неправильную, но не позволяет доказать, что она верна, т. е. статистические критерии указывают лишь на отсутствие опровержения со стороны имеющихся экспериментальных данных. Если по результатам проверки статистическая гипотеза принимается, то говорят, что она *согласуется с выборочными данными* или что она *не противоречит результатам наблюдений*.

Статистический критерий обычно основывается на некоторой статистике $\hat{\theta}_n$, для которой известно ее точное или приближенное распределение. Множество всех возможных значений этой статистики разбивается на два непересекающихся подмножества: S – **область принятия нулевой гипотезы** и W – область отклонения нулевой гипотезы. W называется **критической областью**.

В задачах проверки гипотез возможны следующие четыре ситуации.

| Проверяемая гипотеза H_0 : | H_0 принимается – | H_0 отвергается – |
|------------------------------|-------------------------|-------------------------|
| объективно верна | правильное решение | ошибка 1-го рода |
| объективно неверна | ошибка 2-го рода | правильное решение |

Опр. 5. Вероятность ошибки 1-го рода, т. е. вероятность отвергнуть нулевую гипотезу, когда она верна, называется *уровнем значимости* статистического критерия и обозначается α :

$$P(H_0 \text{ отвергается} | H_0 \text{ верна}) = P(\hat{\theta}_n \in W | H_0 \text{ верна}) = \alpha.$$

Вероятность ошибки 2-го рода, т. е. вероятность ошибочно принять нулевую гипотезу, обозначается β :

$$P(H_0 \text{ принимается} | H_0 \text{ не верна}) = P(\hat{\theta}_n \in S | H_0 \text{ не верна}) = \beta.$$

Пользуясь терминологией статистического контроля качества продукции, можно сказать, что α – это риск поставщика (забраковка партии, удовлетворяющей стандарту), а β – риск потребителя (принятие партии, не удовлетворяющей стандарту).

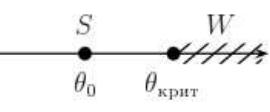
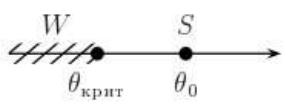
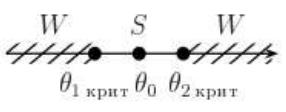
Опр. 6. Мощностью критерия называется вероятность отклонить проверяемую гипотезу H_0 , когда она неверна. Эта вероятность равна

$$P(H_0 \text{ отвергается} | H_0 \text{ не верна}) = 1 - \beta.$$

Двусторонние и односторонние критические области

Иногда возникает необходимость сравнения гипотезы $H_0 : \theta = \theta_0$ с *односторонней* альтернативой $\bar{H}_1 : \theta > \theta_0$ или $\bar{H}_2 : \theta < \theta_0$. Например, если известно, что неравенство $\theta < \theta_0$ невозможно, то в качестве альтернативной рассматривается гипотеза $\bar{H} : \theta > \theta_0$.

Вид критической области W и области S принятия гипотезы зависит от вида альтернативной гипотезы.

| | | |
|--|---|---|
| $H_0 : \theta = \theta_0$ $\bar{H} : \theta > \theta_0$ $W = \{\hat{\theta}_n > \theta_{\text{крит}}\}$ | $H_0 : \theta = \theta_0$ $\bar{H} : \theta < \theta_0$ $W = \{\hat{\theta}_n < \theta_{\text{крит}}\}$ | $H_0 : \theta = \theta_0$ $\bar{H} : \theta \neq \theta_0$ $W = \left\{ \begin{array}{l} \hat{\theta}_n < \theta_{1 \text{ крит}} \\ \text{или} \\ \hat{\theta}_n > \theta_{2 \text{ крит}} \end{array} \right\}$ |
|  $P(\hat{\theta}_n > \theta_{\text{крит}} H_0) = \alpha$ правосторонняя критическая область |  $P(\hat{\theta}_n < \theta_{\text{крит}} H_0) = \alpha$ левосторонняя критическая область |  $\begin{aligned} P(\hat{\theta}_n < \theta_{1 \text{ крит}} H_0) &= \\ &= P(\hat{\theta}_n > \theta_{2 \text{ крит}} H_0) = \frac{\alpha}{2} \end{aligned}$ двусторонняя критическая область |

Таким образом, в зависимости от вида альтернативной гипотезы \bar{H} выбирают *правостороннюю*, *левостороннюю* или *двустороннюю* критическую область.

27. Проверка гипотезы о виде закона распределения. Критерий согласия χ^2 Пирсона.

Пусть имеется выборка объема n и сгруппированный статистический ряд, в котором k групп. Например, в случае непрерывной СВ это будут k интервалов $[x_{i-1}; x_i]$.

Группы должны выбираться так, чтобы охватывать весь диапазон значений предполагаемой СВ. Если диапазон значений СВ не ограничен (к примеру, нормальная СВ принимает любые значения из $(-\infty; +\infty)$), то крайние интервалы должны быть расширены до $-\infty$ и $+\infty$ соответственно.

Кроме того, интервалы (группы) должны быть не очень маленькими, чтобы в каждый из них входило не менее 5 наблюдений. Группы с малым количеством наблюдений объединяют с соседними.

Проверяемая гипотеза представляет собой предположение о распределении наблюдаемой СВ и является простой (конкретно указывает предполагаемое распределение):

H_0 : функция распределения наблюдаемой СВ совпадает с $F(x)$;

\bar{H} : функция распределения наблюдаемой СВ не совпадает с $F(x)$.

Критерий согласия χ^2 Пирсона основан на сравнении эмпирических и теоретических частот попадания СВ в рассматриваемые группы (интервалы):

n_i – эмпирическая частота наблюдения значений из интервала $[x_{i-1}; x_i]$;

$np_i = n P(\xi \in [x_{i-1}; x_i]) = n(F(x_i) - F(x_{i-1}))$ – теоретическое значение соответствующей частоты.

Рассмотрим статистику

$$\chi^2_{\text{расч}} = \sum_{i=1}^k \frac{(n_i - np_i)^2}{np_i}.$$

Упражнение. Показать, что контроль вычислений можно осуществить по формуле $\chi^2_{\text{расч}} = \sum_{i=1}^k \frac{n_i^2}{np_i} - n$.

Для вычисления статистики $\chi^2_{\text{расч}}$ нужно знать сгруппированный статистический ряд и теоретическую функцию распределения $F(x)$ для расчета вероятностей p_i .

При этом теоретическое распределение $F(x)$ может зависеть от одного или нескольких параметров. Пусть r – число неизвестных параметров теоретического распределения. В этом случае вместо значений параметров используются их оценки.

Замечание. Оценки параметров рассчитываются по сгруппированному статистическому ряду до объединения групп.

Таким образом, **критерий согласия χ^2 Пирсона** заключается в следующем: если $\chi^2_{\text{расч}} < \chi^2_{\alpha; k-r-1}$, где $\chi^2_{\alpha; k-r-1}$ определяется по таблице квантилей распределения χ^2 , то гипотеза H_0 принимается (признается непротиворечащей экспериментальным данным; нет оснований отвергнуть гипотезу H_0) на уровне значимости α , а если $\chi^2_{\text{расч}} \geq \chi^2_{\alpha; k-r-1}$, то гипотеза H_0 отвергается (не согласуется с данными эксперимента).

Основное достоинство критерия согласия χ^2 Пирсона – его универсальность, т. е. применимость для любого закона распределения, в том числе с неизвестными параметрами. Основной недостаток – необходимость большого объема выборки (не менее 60–100 наблюдений) и произвольность группировки, влияющая на величину $\chi^2_{\text{расч}}$.

28. Критерии значимости. Проверка гипотез о математических ожиданиях одной и двух независимых нормальных выборок.

Напомним, что статистические критерии, с помощью которых проверяются гипотезы о значениях параметров распределения или о соотношениях между ними в предположении, что тип распределения известен, называются **критериями значимости** или **статистическими критериями**.

Пусть по выборке объема n получена некоторая оценка $\hat{\theta}$ для параметра θ теоретического распределения и есть основания полагать, что истинное значение параметра θ есть θ_0 . Тогда проверяется нулевая гипотеза $H_0 : \theta = \theta_0$ в сравнении с альтернативой $\bar{H} : \theta \neq \theta_0$.

Выборочное среднее является оценкой для среднего значения измеряемой величины и может служить оценкой того или иного показателя качества. Дисперсия характеризует разброс экспериментальных значений, а следовательно, служит мерой точности. Например, если произведено несколько измерений одной и той же величины, то дисперсия может характеризовать точность прибора, метода измерения и т. д.

1. Проверка гипотезы о равенстве математического ожидания нормального распределения заданному значению.

Нулевая гипотеза $H_0 : a = a_0$.

Альтернативная гипотеза $\bar{H} : a \neq a_0$.

Требуется по выборке объема n проверить гипотезу H_0 при заданном уровне значимости α . При этом предполагается, что выборка взята из нормально распределенной генеральной совокупности.

Если дисперсия σ^2 известна, то утверждение 1 §3 гласит, что при справедливости гипотезы H_0 имеет место $\frac{\bar{x} - a_0}{\sigma / \sqrt{n}} \sim \mathcal{N}(0; 1)$. Следовательно, критерий принятия гипотезы может быть выбран из условия

$$P\left(\left|\frac{\bar{x} - a}{\sigma / \sqrt{n}}\right| < u_\alpha\right) = 1 - \alpha.$$

Таким образом, если дисперсия σ^2 известна, то гипотеза H_0 принимается (т. е. согласуется с результатами наблюдений) при условии, что

$$u_{\text{расч}} = \frac{|\bar{x} - a_0|}{\sqrt{\sigma^2 / n}} < u_{\text{табл}} = u_\alpha, \quad (1)$$

где квантиль u_α удовлетворяет соотношению $\Phi(u_\alpha) = \frac{1-\alpha}{2}$.

Если дисперсия σ^2 неизвестна, то гипотеза H_0 принимается при

$$t_{\text{расч}} = \frac{|\bar{x} - a_0|}{\sqrt{s^2 / n}} < t_{\text{табл}} = t_{\alpha; n-1}, \quad (2)$$

где квантиль $t_{\alpha; n-1}$ определяется по таблице распределения Стьюдента.

29. Критерии значимости. Проверка гипотез о дисперсиях одной и двух независимых нормальных выборок.

2. Проверка гипотезы о равенстве заданному значению дисперсии нормального распределения.

Нулевая гипотеза $H_0 : \sigma^2 = \sigma_0^2$.

Альтернативная гипотеза $\bar{H} : \sigma^2 \neq \sigma_0^2$.

Гипотеза H_0 при заданном уровне значимости α принимается, если

$$\chi_{1-\alpha/2; n-1}^2 < \chi_{\text{расч}}^2 = \frac{(n-1)s^2}{\sigma_0^2} < \chi_{\alpha/2; n-1}^2, \quad (3)$$

где квантили $\chi_{1-\alpha/2; n-1}^2$ и $\chi_{\alpha/2; n-1}^2$ определяются по таблице распределения χ^2 .

3. Сравнение двух дисперсий нормально распределенных признаков. Такая задача возникает, если требуется сравнить точность приборов, инструментов, методов измерения. Лучшим будет тот прибор, инструмент, метод, который дает меньший разброс результатов, т. е. меньшую дисперсию.

Нулевая гипотеза $H_0 : \sigma_1^2 = \sigma_2^2$.

Альтернативная гипотеза $\bar{H} : \sigma_1^2 \neq \sigma_2^2$.

Пусть для первой дисперсии по выборке объема n_1 найдена несмешенная оценка s_1^2 , для второй – по выборке объема n_2 оценка s_2^2 .

В случае двух независимых выборок из нормального распределения, согласно утверждению 3 §3 и определению F -распределения Фишера, отношение $\frac{s_1^2}{s_2^2}$ имеет распределение Фишера с числами степеней свободы

$f_1 = n_1 - 1$ и $f_2 = n_2 - 1$. Следовательно, критерий принятия гипотезы может быть выбран из условия

$$P\left(F_{1-\alpha/2; f_1; f_2} < \frac{s_1^2}{s_2^2} < F_{\alpha/2; f_1; f_2}\right) = 1 - \alpha.$$

Для квантилей распределения Фишера имеет место соотношение

$$F_{1-\alpha; f_1; f_2} = \frac{1}{F_{\alpha; f_2; f_1}}.$$

Поэтому

$$F_{1-\alpha/2; f_1; f_2} < \frac{s_1^2}{s_2^2} \Leftrightarrow \frac{1}{F_{\alpha/2; f_2; f_1}} < \frac{s_1^2}{s_2^2} \Leftrightarrow \frac{s_2^2}{s_1^2} < F_{\alpha/2; f_2; f_1}.$$

Это позволяет сформулировать критерий проверки гипотезы H_0 следующим образом.

Гипотеза H_0 при заданном уровне значимости α принимается, если

$$F_{\text{расч}} = \frac{s_{\max}^2}{s_{\min}^2} < F_{\text{табл}} = F_{\alpha/2; f_1; f_2}. \quad (4)$$

Здесь $F_{\text{расч}}$ равно отношению *большей* несмешенной оценки дисперсии к *меньшей*, квантиль $F_{\alpha/2; f_1; f_2}$ определяется по таблице распределения Фишера, причем f_1 и f_2 – числа степеней свободы соответственно числителя и знаменателя, т. е. *большой* и *меньшей* оценок дисперсий.

30. Критерии значимости. Проверка гипотез о математических ожиданиях двух зависимых и независимых нормальных выборок.

4. Сравнение двух средних в случае независимых нормально распределенных признаков.

Нулевая гипотеза $H_0: \mu_1 = \mu_2$.

Альтернативная гипотеза $\bar{H}: \mu_1 \neq \mu_2$.

Требуется по выборкам объемов n_1 и n_2 проверить гипотезу H_0 при заданном уровне значимости α .

1 случай. Если дисперсии σ_1^2 и σ_2^2 известны, то гипотеза H_0 принимается при условии, что

$$u_{\text{расч}} = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} < u_{\text{табл}} = u_\alpha, \quad (5)$$

где квантиль u_α удовлетворяет соотношению $\Phi(u_\alpha) = \frac{1-\alpha}{2}$.

2 случай. Если дисперсии σ_1^2 и σ_2^2 не известны, но на основании проверки соответствующей гипотезы по критерию Фишера признаны однородными, то гипотеза H_0 принимается при

$$t_{\text{расч}} = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{s^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} < t_{\text{табл}} = t_{\alpha; f}, \quad (6)$$

где общая средневзвешенная дисперсия s^2 вычисляется по формуле

$$s^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2}$$

и имеет число степеней свободы $f = n_1 + n_2 - 2$, значение $t_{\alpha; f}$ определяется по таблице квантилей распределения Стьюдента.

3 случай. Если дисперсии σ_1^2 и σ_2^2 не известны и на основании проверки по критерию Фишера признаны неоднородными, то проверка также проводится по критерию Стьюдента, однако этот критерий является приближенным. В этом случае гипотеза H_0 принимается, если

$$t_{\text{расч}} = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} < t_{\text{табл}} = t_{\alpha; f}, \quad (7)$$

где квантиль $t_{\alpha; f}$ определяется по таблице распределения Стьюдента при

$$f \approx \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \right)^2}{\frac{\left(\frac{s_1^2}{n_1} \right)^2}{n_1-1} + \frac{\left(\frac{s_2^2}{n_2} \right)^2}{n_2-1}}.$$

6. Сравнение двух средних в случае зависимых нормально распределенных признаков. Такая задача возникает, если две выборки взаимосвязаны. Например, проводятся измерения одних и тех же величин на одних и тех же объектах двумя разными методами и требуется определить, одинаковы ли результаты использования двух методов измерения. Либо если проводятся измерения какой-то характеристики для одних и тех же объектов до и после некоторого воздействия и требуется определить, влияет ли это воздействие на значение характеристики.

В этом случае имеются две выборки одинакового объема n :

$$x_{11}, \quad x_{12}, \quad \dots, \quad x_{1n};$$

$$x_{21}, \quad x_{22}, \quad \dots, \quad x_{2n}.$$

Поскольку значения в каждой паре x_{1i}, x_{2i} связаны (например, измерены на одном и том же объекте), то получим новую выборку с элементами $\Delta x_i = x_{1i} - x_{2i}$.

Задача сводится к проверке гипотезы о равенстве нулю среднего значения новой выборки, т. е. $H_0: a_{\Delta x} = 0$. Эта проверка проводится по критерию (2).

31. Использование распределения Стьюдента при построении доверительных интервалов и проверке статистических гипотез.

Утв. 5. Доверительный интервал для математического ожидания a в случае выборки из нормального распределения с неизвестной дисперсией σ^2 определяется соотношением

$$P\left(\bar{x} - t_{\alpha; n-1} \frac{s}{\sqrt{n}} < a < \bar{x} + t_{\alpha; n-1} \frac{s}{\sqrt{n}}\right) = 1 - \alpha,$$

где n – объем выборки; \bar{x} – выборочное среднее; s^2 – несмещенная оценка дисперсии; α – уровень значимости; $t_{\alpha; n-1}$ – квантиль уровня α распределения Стьюдента с числом степеней свободы $k = n - 1$, т. е. такое число, что для СВ ξ , имеющей распределение Стьюдента с числом степеней свободы $k = n - 1$, имеет место $P(|\xi| \geq t_{\alpha; n-1}) = \alpha$.

Упражнение 3. Доказать аналогично доказательству утверждения 2.

Квантиль $t_{\alpha; n-1}$ определяется по таблице распределения Стьюдента.

При малых выборках ($n < 30$) распределение Стьюдента дает не вполне определенные результаты (широкий доверительный интервал). Это объясняется тем, что малая выборка содержит малую информацию об интересующем нас признаке. С возрастанием числа степеней свободы распределение Стьюдента быстро приближается к нормальному.

32. Использование нормального распределения при построении доверительных интервалов и проверке статистических гипотез.

Смотреть вопрос 24?... Я не понимаю, что тут нужно...

33. Использование χ^2 -распределения при построении доверительных интервалов и проверке статистических гипотез.

Смотреть вопрос 25?... Я не понимаю, что тут нужно...

34. Виды зависимостей между случайными величинами. Основные задачи корреляционного и регрессионного анализа.

Пусть на основании экспериментальных данных (по выборке объема n связанных пар наблюдений (x_i, y_i)) изучается связь между двумя величинами. Две случайные величины могут быть: 1) независимыми; 2) связаны функциональной зависимостью, когда каждому значению одной из них соответствует строго определенное значение другой; 3) связаны *статистической* зависимостью, при которой каждому значению одной из них соответствует множество возможных значений другой, т.е. изменение одной из величин влечет изменение *распределения* другой, в частности, может изменяться *среднее значение* другой.

Пример. Статистической является зависимость урожайности некоторой культуры от количества вносимых удобрений или количества осадков; зависимость спроса на товар от его цены; надежности автомобиля от его возраста и т. д.

Статистическая зависимость возникает из-за того, что на зависимую переменную влияют какие-то неучтенные или неконтролируемые факторы.

При изучении статистической зависимости обычно ограничиваются исследованием усредненной зависимости: как в среднем будет изменяться значение одной величины при изменении другой. Такая зависимость называется *регрессионной*. Более строго, регрессионная зависимость между двумя случайными величинами – это функциональная зависимость между значениями одной из них и условным математическим ожиданием другой.

Основным методом исследования статистических зависимостей является *корреляционно-регрессионный* анализ.

Корреляционный анализ состоит в определении *степени связи* между случайными величинами.

Целью *регрессионного анализа* является установление *формы зависимости* между наблюдаемыми величинами и определение по экспериментальным данным уравнения зависимости, которое называют *выборочным (эмпирическим) уравнением регрессии*, а также прогнозирование с помощью уравнения регрессии среднего значения зависимой переменной при заданном значении независимой переменной.

Вид эмпирической функции регрессии определяют исходя из: 1) соображений о физической сущности исследуемой зависимости; 2) опыта предыдущих исследований; 3) характера расположения точек на *корреляционном поле*, которое получается, если отметить на плоскости все точки с координатами (x_i, y_i) , соответствующие наблюдениям.

Наибольший интерес представляет линейное эмпирическое уравнение регрессии $y = ax + b$, т. к. 1) это наиболее простой случай для расчетов и анализа; 2) при нормальном распределении функция регрессии является линейной.

35. Выборочный коэффициент корреляции и его свойства.

Количественной мерой линейной связи между двумя наблюдаемыми величинами служит **выборочный коэффициент корреляции**.

$$r_{xy} = \frac{\bar{xy} - \bar{x} \cdot \bar{y}}{\sigma_x \cdot \sigma_y}$$

где $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$, $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$, $\bar{xy} = \frac{1}{n} \sum_{i=1}^n x_i \cdot y_i$ - выборочные средние x , y , xy

соответственно, $\sigma_x = \sqrt{D_{ex}} = \sqrt{\bar{x^2} - (\bar{x})^2}$, $\sigma_y = \sqrt{D_{ey}} = \sqrt{\bar{y^2} - (\bar{y})^2}$, $\bar{x^2} = \frac{1}{n} \sum_{i=1}^n x_i^2$,

$$\bar{y^2} = \frac{1}{n} \sum_{i=1}^n y_i^2.$$

Свойства выборочного коэффициента корреляции.

1. $-1 \leq r_{xy} \leq 1$
2. Если $|r_{xy}| \approx 1$, то значения x и y связаны линейной зависимостью;
3. Если значения x и y независимы, то $|r_{xy}| \approx 0$.
4. Если $r_{xy} > 0$, то с ростом одной величины увеличивается другая, если $r_{xy} < 0$, то, наоборот, уменьшается.

Проверка значимости коэффициента корреляции - это проверка того, что коэффициент корреляции значимо отличается от нуля. Т. к. выборка произведена случайно, нельзя утверждать, что если выборочный коэффициент корреляции $r_{xy} \neq 0$, то и коэффициент корреляции генеральной совокупности $\rho \neq 0$. Это зависит от соотношения объема выборки и значения r_{xy} . Если выборка из нормального распределения, то проверка производится по **критерию Стьюдента**: если

$$t_{расч} = |r_{xy}| \sqrt{\frac{n-2}{1-r_{xy}^2}} > t_{рабл} = t_{\alpha, n-2},$$

где $t_{\alpha, n-2}$ - квантиль t -распределения Стьюдента (определяется по таблице), то при заданном уровне значимости α (допускается, что вывод может быть ошибочным с небольшой вероятностью α) коэффициент корреляции считается значимо отличающимся от нуля, а следовательно, связь между величинами x, y признается статистически значимой.

Подчеркнем, что коэффициент корреляции является мерой именно *линейной* зависимости. В случае нелинейной зависимости связь между величиной коэффициента корреляции и близостью точек корреляционного поля к некоторой линии не прослеживается. Поэтому в практических задачах при выборе вида эмпирической функции регрессии обязательно учитывают характер расположения точек на корреляционном поле.

36. Эмпирическое линейное уравнение регрессии. Метод наименьших квадратов.

Основным методом исследования статистических зависимостей является **корреляционно-регрессионный анализ**.

Корреляционный анализ состоит в определении *степени связи* между случайными величинами.

Целью *регрессионного анализа* является установление **формы зависимости** между наблюдаемыми величинами и определение по экспериментальным данным уравнения зависимости, которое называют **выборочным (эмпирическим) уравнением регрессии**, а также прогнозирование с помощью уравнения регрессии среднего значения зависимой переменной при заданном значении независимой переменной.

Наиболее распространенным методом нахождения коэффициентов эмпирического уравнения регрессии $\hat{y} = ax + b$ по выборке (x_i, y_i) , $i = 1, \dots, n$ является **метод наименьших квадратов (МНК)**. Суть этого метода в том, что коэффициенты a и b выбирают так, чтобы сумма квадратов отклонений наблюдаемых значений y_i от предсказываемых по уравнению $\hat{y}_i = ax_i + b$ была минимальной. Таким образом, минимизируется функция

$$Q(a, b) = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - b - ax_i)^2 \rightarrow \min_{a, b} .$$

Необходимым условием существования минимума данной функции двух переменных является равенство нулю ее частных производных по неизвестным параметрам b, a :

$$\begin{cases} \frac{\partial Q}{\partial b} = -2 \sum_{i=1}^n (y_i - b - ax_i) = 0, \\ \frac{\partial Q}{\partial a} = -2 \sum_{i=1}^n (y_i - b - ax_i) x_i = 0. \end{cases}$$

Отсюда получаем **систему нормальных уравнений**:

$$\begin{cases} nb + a \sum x_i = \sum y_i, \\ b \sum x_i + a \sum x_i^2 = \sum x_i y_i. \end{cases}$$

Метод наименьших квадратов широко применяется при статистической обработке результатов измерений.

37. Простые и составные числа. Бесконечность множества простых чисел. Простые числа Мерсенна. Задача факторизации целых чисел.

Множество целых чисел обозначают

$$\mathbb{Z} = \{0; \pm 1; \pm 2; \dots; \pm n; \dots\}.$$

Натуральными называются числа, которые используются при счете; **множество натуральных чисел** обозначают

$$\mathbb{N} = \{1; 2; \dots; n; \dots\}.$$

На множестве целых чисел определены операции сложения и умножения, а также операция вычитания как обратная к операции сложения. Сумма, разность и произведение двух целых чисел все-

где являются целым числом. Результат деления двух целых чисел не всегда является целым числом.

Опр. 1. Если для целых чисел a и b , где $b \neq 0$, существуют целые числа q и r такие, что

$$a = b \cdot q + r, \text{ где } 0 \leq r < |b|,$$

то r называют **остатком**, а q – **частным (неполным частным)** при $r \neq 0$ от деления a на b .

Т 1 (о делении с остатком). Для любых целых чисел a и b , где $b \neq 0$, существуют единственныe целые числа q и r такие, что

$$a = b \cdot q + r, \text{ где } 0 \leq r < |b|.$$

Пример 1.

1) $a = 37, b = 15$. Поскольку $37 = 15 \cdot 2 + 7$, то $q = 2, r = 7$.

2) $a = -26, b = 4$. Поскольку $-26 = 4 \cdot (-7) + 2$, то $q = -7, r = 2$.

3) $a = 22, b = -5$. Поскольку $22 = -5 \cdot (-4) + 2$, то $q = -4, r = 2$.

4) $a = -15, b = -6$. Поскольку $-15 = -6 \cdot 3 + 3$, то $q = 3, r = 3$. •

Опр. 2. Если остаток от деления a на b равен 0 ($r = 0$), т. е. $a = b \cdot q$, то говорят,

- что a **делится на b и на q** (и пишут $a:b, a:q$);
- что a является **кратным** чисел b и q ;
- что b и q **делят a** (и пишут $b|a, q|a$);

- что b и q являются **делителями** (или **множителями**) числа a .

Будем обозначать $b \nmid a$, если b не делит a .

Число 0 делится на любое целое число $b \neq 0$.

Любое целое число $a \neq 0$ делится на $1; -1; a; -a$. В дальнейшем будем говорить только о **целых положительных**, т. е. **натуральных делителях**.

Простые и составные числа

Опр. 3. Натуральное число $n > 1$ называется **простым**, если оно делится только на 1 и на само себя, в противном случае n называется **составным**.

Для целей криптографии (как для практической реализации и обоснования стойкости криптографических средств, так и для разработки методов их вскрытия) необходимо разрабатывать эффективные методы и алгоритмы:

- проверки простоты целых чисел;
- поиска больших простых чисел (в криптографии используются большие простые числа длиной более 80–90 десятичных знаков);
- факторизации целых чисел.

Факторизацией натурального числа называется разложение этого числа в произведение простых сомножителей. Эта задача имеет большую вычислительную сложность. Один из самых популярных методов криптографии с открытым ключом, метод RSA, основан на трудоемкости задачи факторизации длинных целых чисел.

Числа Мерсенна

Числа Мерсенна – это числа вида $2^p - 1$. Числа Мерсенна были открыты в результате поиска совершенных чисел (**совершенными** называются натуральные числа, которые равны сумме всех своих делителей, меньших данного числа, например, $6 = 1 + 2 + 3$; первые четыре совершенных числа – это 6; 28; 496; 8128).

Одним из свойств чисел Мерсенна является то, что числа такого вида могут быть простыми только тогда, когда p – простое число. Однако не для любого простого p число $2^p - 1$ является простым, например, $2^{11} - 1 = 2047 = 23 \cdot 89$.

38.НОД и НОК целых чисел. Основная теорема арифметики. Методы нахождения НОД.

Опр. 1. Максимальный из общих делителей целых чисел a_1, a_2, \dots, a_n называется их **наибольшим общим делителем (НОД)** и обозначается: $\text{НОД}(a_1, a_2, \dots, a_n)$ или (a_1, a_2, \dots, a_n) .

Опр. 2. Минимальное натуральное из общих кратных целых чисел a_1, a_2, \dots, a_n называется их **наименьшим общим кратным (НОК)** и обозначается: $\text{НОК}(a_1, a_2, \dots, a_n)$ или $[a_1, a_2, \dots, a_n]$.

Утв. 1. Если канонические разложения двух чисел имеют вид

$$a = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_s^{\alpha_s}, \quad b = p_1^{\beta_1} p_2^{\beta_2} \dots p_s^{\beta_s}, \quad \text{где } \alpha_i, \beta_i \geq 0,$$

то

$$(a, b) = p_1^{\gamma_1} p_2^{\gamma_2} \dots p_s^{\gamma_s}, \quad \text{где } \gamma_i = \min\{\alpha_i, \beta_i\};$$

$$[a, b] = p_1^{\delta_1} p_2^{\delta_2} \dots p_s^{\delta_s}, \quad \text{где } \delta_i = \max\{\alpha_i, \beta_i\}.$$

Пример 1. Найдем $(168, 180)$ и $[168, 180]$.

Решение. Поскольку

$$168 = 2 \cdot 84 = 2 \cdot 2 \cdot 42 = 2 \cdot 2 \cdot 2 \cdot 21 = 2 \cdot 2 \cdot 2 \cdot 3 \cdot 7 = 2^3 \cdot 3 \cdot 7;$$

$$180 = 2 \cdot 90 = 2 \cdot 2 \cdot 45 = 2 \cdot 2 \cdot 3 \cdot 15 = 2 \cdot 2 \cdot 3 \cdot 3 \cdot 5 = 2^2 \cdot 3^2 \cdot 5,$$

то

$$(168, 180) = 2^2 \cdot 3^1 \cdot 5^0 \cdot 7^0 = 12;$$

$$[168, 180] = 2^3 \cdot 3^2 \cdot 5^1 \cdot 7^1 = 2520. \bullet$$

Т 1. НОК и НОД двух целых чисел связаны соотношением:

$$a,b=ab.$$

Доказательство. Если $d = (a,b)$, то $a = a_1d$, $b = b_1d$, где $(a_1, b_1) = 1$. Тогда

$$[a,b] = a_1b_1d = \frac{ab}{d} = \frac{ab}{(a,b)},$$

что и доказывает теорему. \triangleleft

Опр. 3. Целые числа a и b называются *взаимно простыми*, если $(a,b) = 1$. (Другими словами, это числа, не имеющие общих простых делителей.)

Пример 2. Числа 6 и 35 взаимно просты, так как $(6,35) = 1$, но 6 и 27 не являются взаимно простыми, так как $(6,27) = 3$. •

Замечание. Число 1 взаимно просто с любым целым числом; число 0 взаимно просто только с 1 и -1 .

Т 2. Если $a = bq + r$, то $(a,b) = (b,r)$.

Доказательство. Пусть $(a,b) = d$, $(b,r) = k$.

По свойству делимости, если $d|a$ и $d|b$, то $d|r$. Следовательно, $d|k$.

С другой стороны, если $k|b$ и $k|r$, то $k|a$. Следовательно, $k|d$.

Поскольку $d|k$ и $k|d$, причем $d, k > 0$, то $d = k$ и теорема доказана. \triangleleft

На этой теореме основывается алгоритм Евклида.

Т 3 (Основная теорема арифметики). Всякое натуральное число $n > 1$ однозначно раскладывается в произведение простых чисел с точностью до порядка следования множителей:

$$n = p_1p_2 \dots p_s.$$

Если в разложении натурального числа на простые множители собрать одинаковые множители, то получим *каноническое разложение* натурального числа:

$$n = p_1^{r_1} p_2^{r_2} \dots p_t^{r_t}.$$

Каноническим разложением целого отрицательного числа $z = -n$ считается, соответственно, его представление в виде $z = -p_1^{r_1} \dots p_t^{r_t}$.

39. Алгоритм Евклида. Расширенный алгоритм Евклида. Соотношение Безу.

Алгоритм Евклида – алгоритм для определения НОД двух чисел путем последовательного применения теоремы о делении с остатком.

Т 3. Наибольший общий делитель целых чисел a и b (где $a \nmid b$, $b \nmid a$, $|a| > |b|$) равен последнему отличному от нуля остатку от деления в цепочке равенств:

$$a = bq_1 + r_1;$$

$$b = r_1q_2 + r_2, \text{ если } r_1 \neq 0;$$

...

$$r_{n-2} = r_{n-1}q_n + r_n, \text{ если } r_{n-1} \neq 0; \\ r_{n-1} = r_nq_{n+1}, \text{ если } r_n \neq 0,$$

т. е. $r_n = (a, b)$.

Доказательство. Согласно теореме 2 и поскольку $r_n > 0$ как остаток от деления, имеем

$$(a, b) = (b, r_1) = (r_1, r_2) = \dots = (r_{n-1}, r_n) = r_n.$$

Процесс получения (a, b) конечен, поскольку мы оперируем только с целыми числами и, начиная с деления r_1 на r_2 , — с целыми положительными числами. При этом идет постоянное уменьшение остатков $r_i : 0 \leq r_i < r_{i-1}$, поэтому за конечное число шагов будет достигнут остаток $r_{n+1} = 0$. \triangleleft

Опр. 4. Полученное равенство называют **линейным разложением**, или **соотношением Безу** для наибольшего общего делителя целых чисел a и b , а числа u и v — **коэффициентами Безу**.

Пример 4. Найдем соотношение Безу для $(72, 26)$.

Решение. Из примера 3 следует, что

$$\begin{aligned} 2 &= 20 + 6 \cdot (-3) = \\ &= 20 + (26 + 20 \cdot (-1)) \cdot (-3) = \\ &= 20 \cdot 4 + 26 \cdot (-3) = \\ &= (72 + 26 \cdot (-2)) \cdot 4 + 26 \cdot (-3) = \\ &= 72 \cdot 4 + 26 \cdot (-11). \end{aligned}$$

Таким образом, $2 = 72u + 26v$, где $u = 4$, $v = -11$. \bullet

Замечание. Числа u и v не являются единственной парой с таким условием. Это следует из теории диофантовых линейных уравнений, которая будет рассмотрена ниже. Так, в примере 4 числа $u = -9$ и $v = 25$ также удовлетворяют соотношению Безу.

Следствие теоремы 4 (критерий взаимной простоты). Целые числа a и b взаимно просты тогда и только тогда, когда существуют такие целые u и v , что

$$au + bv = 1.$$

Доказательство. $\Rightarrow)$ Если $(a, b) = 1$, то из соотношения Безу следует, что существуют такие целые числа u и v , что $au + bv = 1$.

$\Leftarrow)$ Обратно, пусть существуют такие целые u и v , что $au + bv = 1$. Если $(a, b) = d > 1$, то $d \mid 1$ (по свойству делимости), а значит, $(a, b) = 1$. \triangleleft

Расширенный алгоритм Евклида

При нахождении соотношения Безу более удобен **расширенный (обобщенный) алгоритм Евклида**, позволяющий вычислять коэффициенты Безу параллельно с нахождением НОД.

Пусть $|a| > |b|$. Положим

$$\begin{aligned} u_0 &= 1, & v_0 &= 0, & r_0 &= a; \\ u_1 &= 0, & v_1 &= 1, & r_1 &= b. \end{aligned}$$

Далее последовательно вычисляем:

$$u_{i+1} = u_{i-1} - q_i u_i, \quad v_{i+1} = v_{i-1} - q_i v_i, \quad r_{i+1} = r_{i-1} - q_i r_i,$$

где q_i – неполное частное от деления r_{i-1} на r_i .

Работа алгоритма заканчивается, если на некотором шаге $r_{i+1} = 0$. При этом на предыдущем шаге найдены НОД и коэффициенты Безу:

$$(a, b) = r_i, \quad u = u_i, \quad v = v_i.$$

Обоснованием алгоритма служит следующая теорема.

Т 5. При всех i выполняется равенство: $u_i a + v_i b = r_i$.

Доказательство проводится индукцией по i .

1) *База индукции*: если $i = 0$ или $i = 1$, то, очевидно, равенство выполняется:

$$\begin{aligned} 1 \cdot a + 0 \cdot b &= a; \\ 0 \cdot a + 1 \cdot b &= b. \end{aligned}$$

2) *Шаг индукции*. Предположим, что утверждение верно для всех $k \leq i$. Тогда для следующего номера $i + 1$ получим

$$\begin{aligned} u_{i+1} \cdot a + v_{i+1} \cdot b &= (u_{i-1} - q_i u_i) \cdot a + (v_{i-1} - q_i v_i) \cdot b = \\ &= (u_{i-1} \cdot a + v_{i-1} \cdot b) - q_i \cdot (u_i \cdot a + v_i \cdot b) = r_{i-1} - q_i \cdot r_i = r_{i+1}, \end{aligned}$$

что и требовалось доказать. \triangleleft

40. Диофантовы линейные уравнения.

Опр. 5. *Диофантовым линейным уравнением* с двумя неизвестными называется уравнение вида

$$ax + by = c, \quad (1)$$

где $a, b, c \in \mathbb{Z}$, $a, b \neq 0$, решения $(x; y)$ ищутся в целых числах.

Иными словами, все коэффициенты и неизвестные – целые числа.

Т 6. Уравнение (1) разрешимо в целых числах тогда и только тогда, когда $(a, b) | c$.

Доказательство. 1) *Необходимость*. Пусть $(x_0; y_0)$ – целочисленное решение уравнения (1). Тогда справедливо равенство $ax_0 + by_0 = c$, а значит, по свойству делимости целых чисел $(a, b) | c$.

2) *Достаточность*. Пусть $d = (a, b)$ и $d | c$, тогда $c = dt$ для некоторого $t \in \mathbb{Z}$. Запишем соотношение Безу для $d = (a, b)$: существуют такие числа $u, v \in \mathbb{Z}$, что $au + bv = d$. Умножая последнее равенство на t , получаем $aut + bvt = c$, т. е. $(tu; tv)$ – целочисленное решение уравнения (1). \triangleleft

Т 7. Уравнение (1) либо не имеет решений в целых числах, либо имеет бесконечно много решений в целых числах.

Доказательство. Если $(a, b) \nmid c$, то в силу теоремы 6 уравнение (1) не имеет решений в целых числах.

Если $(a, b) | c$, то уравнение (1) разрешимо. При этом если $(x_0; y_0)$ – одно из решений уравнения (1), т. е. $ax_0 + by_0 = c$, то при произвольном целом t справедливо

$$a(x_0 + bt) + b(y_0 - at) = c,$$

а значит, $(x_0 + bt; y_0 - at)$ также является решением уравнения (1) при всех $t \in \mathbb{Z}$. \triangleleft

Можно показать, что имея одно решение $(x_0; y_0)$ уравнения (1), можно получить все его решения по формуле $\left(x_0 + \frac{b}{d}t; y_0 - \frac{a}{d}t \right)$, где $d = (a, b)$, $t \in \mathbb{Z}$. При этом частное решение

$(x_0; y_0)$ может быть найдено с помощью соотношения Безу для (a, b) .

Алгоритм решения диофантова линейного уравнения (1).

1. Если $(a, b) \nmid c$, то уравнение (1) не имеет решений в целых числах.

2. Если $(a, b) = d$, $d \mid c$, то получаем (например, с помощью расширенного алгоритма Евклида) соотношение Безу для (a, b) и находим такие числа $u_0, v_0 \in \mathbb{Z}$, что

$$au_0 + bv_0 = d. \quad (2)$$

3. Умножив обе части равенства (2) на $\frac{c}{d}$, получим

$$a\frac{c}{d}u_0 + b\frac{c}{d}v_0 = c,$$

а значит, $x_0 = \frac{c}{d}u_0$, $y_0 = \frac{c}{d}v_0$ – частное решение уравнения (1).

4. Множество целочисленных решений уравнения (1) задается формулой

$$\left\{ \left(x_0 + \frac{b}{d}t; y_0 - \frac{a}{d}t \right) \middle| t \in \mathbb{Z} \right\}. \quad (3)$$

41. Понятие сравнимости по модулю m . Арифметические свойства сравнений. Множество классов вычетов.

Т 1. Пусть m – натуральное число. Для любых целых чисел a и b следующие условия равносильны:

- 1) a и b имеют одинаковые остатки от деления на m ;
- 2) $a - b$ делится на m , т. е. $a - b = mq$ для подходящего целого q ;

3) $a = b + mq$ для некоторого целого q .

Доказательство проводится по схеме $1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 1$.

$1 \Rightarrow 2$). Пусть $a = q_1m + r$, $b = q_2m + r$, где $0 \leq r < m$. Тогда $a - b = q_1m - q_2m = (q_1 - q_2)m$, т. е. $a - b$ делится на m .

$2 \Rightarrow 3$). Если $a - b$ делится на m , т. е. $a - b = mq$ для некоторого $q \in \mathbb{Z}$, то $a = b + mq$.

$3 \Rightarrow 1$). Пусть $a = q_1m + r_1$, $b = q_2m + r_2$, $0 \leq r_1, r_2 < m$. Докажем, что если $a = b + mq$, то $r_1 = r_2$.

Подставляя в это соотношение выражения для a и b , получим $q_1m + r_1 = q_2m + r_2 + mq$, откуда $r_1 - r_2 = (q_2 + q - q_1)m$, т. е. $m|r_1 - r_2$. Но поскольку $0 \leq r_1, r_2 < m$, то $r_1 = r_2$. \triangleleft

Опр. 1. Целые числа a и b называются *сравнимыми по модулю m* , если они удовлетворяют одному из условий теоремы 1. Этот факт обозначают формулой $a \equiv b \pmod{m}$.

Итак,

$$a \equiv b \pmod{m} \Leftrightarrow \boxed{\begin{array}{l} a \text{ и } b \text{ имеют} \\ \text{одинаковые} \\ \text{остатки от де-} \\ \text{ления на } m \end{array}} \Leftrightarrow a - b \vdots m \Leftrightarrow \boxed{\begin{array}{l} a = b + mq \\ \text{для неко-} \\ \text{торого} \\ q \in \mathbb{Z} \end{array}}$$

Арифметические свойства сравнений

1. В сравнении можно отбрасывать или добавлять слагаемые, делящиеся на модуль: если $a \equiv b \pmod{m}$, то для всякого $k \in \mathbb{Z}$

$$a \equiv (b \pm km) \pmod{m}.$$

2. Сравнения можно почленно складывать, вычитать, умножать, возводить в натуральную степень:

если $a \equiv b \pmod{m}$, $c \equiv d \pmod{m}$, то

$$(a \pm c) \equiv (b \pm d) \pmod{m};$$

$$ac \equiv bd \pmod{m};$$

$$a^n \equiv b^n \pmod{m} \text{ при любом } n \in \mathbb{N}.$$

Доказательство. Докажем второе соотношение (сравнения можно почленно умножать). Если $a \equiv b \pmod{m}$, $c \equiv d \pmod{m}$, то $a = b + mq_1$, $c = d + mq_2$. Тогда

$$ac = (b + mq_1)(d + mq_2) = bd + mq_1d + mq_2b + m^2q_1q_2 = bd + mt, t \in \mathbb{Z}.$$

Следовательно, $ac \equiv bd \pmod{m}$. \triangleleft

Упражнение. Доказать первое соотношение.

3. К обеим частям сравнения можно прибавить или вычесть одно и то же число; обе части сравнения можно умножить на одно и то же число:

если $a \equiv b \pmod{m}$, то для всякого $c \in \mathbb{Z}$

$$(a \pm c) \equiv (b \pm c) \pmod{m};$$

$$ac \equiv bc \pmod{m}.$$

4. Сравнение можно сократить на общий множитель, взаимно простой с модулем: пусть $a = a_1d$, $b = b_1d$, $(d, m) = 1$, тогда

если $a_1d \equiv b_1d \pmod{m}$, то $a_1 \equiv b_1 \pmod{m}$.

Доказательство. Действительно,

$$a_1d \equiv b_1d \pmod{m} \Leftrightarrow (a_1d - b_1d) : m \Leftrightarrow d(a_1 - b_1) : m,$$

откуда, поскольку d и m взаимно просты, следует $(a_1 - b_1) : m$, т. е. $a_1 \equiv b_1 \pmod{m}$. \triangleleft

5. Если в сравнении $a \equiv b \pmod{m}$ числа a , b , m имеют общий множитель d , то на него сравнение можно сократить:

$$\frac{a}{d} \equiv \frac{b}{d} \pmod{\frac{m}{d}}.$$

Доказательство. Сравнение $a \equiv b \pmod{m}$ равносильно $a = b + mq$ при некотором $q \in \mathbb{Z}$. Тогда, так как числа a , b и m делятся на d , то $\frac{a}{d} = \frac{b}{d} + \frac{m}{d}q \Leftrightarrow \frac{a}{d} \equiv \frac{b}{d} \pmod{\frac{m}{d}}$. \triangleleft

$$6. \begin{cases} a \equiv b \pmod{m_1}, \\ \dots, \\ a \equiv b \pmod{m_k} \end{cases} \Leftrightarrow a \equiv b \pmod{[m_1, \dots, m_k]}.$$

Доказательство. Сравнение $a \equiv b \pmod{m_i}$ означает, что $(a - b) : m_i$; указанная система сравнений означает, что число $a - b$ делится на каждое m_i , $1 \leq i \leq k$, а значит $a - b$ делится на НОК чисел m_i , $1 \leq i \leq k$, т. е. $(a - b) : [m_1, \dots, m_k] \Leftrightarrow a \equiv b \pmod{[m_1, \dots, m_k]}$. \triangleleft

7. Если $a \equiv b \pmod{m}$, то $(a, m) = (b, m)$.

Доказательство. Сравнение $a \equiv b \pmod{m}$ равносильно равенству $a = b + mq$ при некотором $q \in \mathbb{Z}$.

Тогда если $d | a$, $d | m$, то $d | b$; если $d | b$, $d | m$, то $d | a$, т. е. всякий делитель чисел a и m является делителем числа b , и всякий делитель чисел b и m является делителем числа a , а следовательно, $(a, m) = (b, m)$. \triangleleft

Иногда полезно иметь в виду следующее утверждение, обобщающее и уточняющее свойства 3, 4 и 5.

Утв. 1.

1) При любом натуральном $c \neq 0$

$$a \equiv b \pmod{m} \Leftrightarrow ac \equiv bc \pmod{mc}.$$

2) Если $(c, m) = 1$, то

$$a \equiv b \pmod{m} \Leftrightarrow ac \equiv bc \pmod{m}.$$

Опр. 2. Множество всех чисел, сравнимых с a по модулю m , называется **классом вычетов по модулю m** (по-латински «residua» – «остаток, оставшаяся часть») и обозначается \bar{a} , т. е.

$$\bar{a} = \{b \in \mathbb{Z} : b \equiv a \pmod{m}\},$$

или

$$\bar{a} = \{\dots; a - 2m; a - m; a; a + m; a + 2m; \dots\}.$$

Любое число из класса вычетов называют **вычетом**. При обозначении класса вычетов можно использовать любой элемент класса, поскольку каждый представитель класса однозначно определяет свой класс, т. е. для любого числа $b \in \bar{a}$ класс $\bar{b} = \bar{a}$.

Утв. 2. $a \equiv b \pmod{m} \Leftrightarrow \bar{a} = \bar{b}$.

Утв. 3. Различные классы вычетов не имеют общих элементов.

Множество классов вычетов

При делении целых чисел на натуральное число m существует ровно m различных остатков: $0, 1, \dots, m-1$. Соответственно этим остаткам множество целых чисел \mathbb{Z} разбивается на m непересекающихся классов вычетов по модулю m . В соответствии с остатком от деления на m эти классы обозначаются $\bar{0}, \bar{1}, \dots, \bar{m-1}$.

Опр. 3. Множество всех классов сравнимых друг с другом чисел по модулю m называют **множеством классов вычетов по модулю m** и обозначают через $\mathbb{Z}/m\mathbb{Z}$ или \mathbb{Z}_m :

$$\mathbb{Z}_m = \mathbb{Z}/m\mathbb{Z} = \{\bar{0}; \bar{1}; \dots; \bar{m-1}\}.$$

Таким образом, \mathbb{Z}_m – множество из m элементов.

Пример 2. $\mathbb{Z}_7 = \mathbb{Z}/7\mathbb{Z} = \{\bar{0}; \bar{1}; \bar{2}; \bar{3}; \bar{4}; \bar{5}; \bar{6}\}$. •

На множестве \mathbb{Z}_m классов вычетов по заданному модулю можно ввести арифметические операции сложения, вычитания и умножения.

Опр. 4. *Суммой* классов вычетов $\bar{a}, \bar{b} \in \mathbb{Z}_m$ называется класс $\bar{a} + \bar{b} = \bar{a+b}$; *разностью* классов вычетов $\bar{a} \in \mathbb{Z}_m$ и $\bar{b} \in \mathbb{Z}_m$ называется класс $\bar{a} - \bar{b} = \bar{a-b}$; *произведением* классов вычетов $\bar{a}, \bar{b} \in \mathbb{Z}_m$ называется класс $\bar{a} \cdot \bar{b} = \bar{a \cdot b}$.

42. Функция Эйлера. Теорема Эйлера. Малая теорема Ферма.

Опр. 1. Функция Эйлера $\varphi(m)$ ставит в соответствие каждому натуральному $m > 1$ количество натуральных чисел, не превосходящих m и взаимно простых с m .

По определению полагается $\varphi(1) = 1$.

Пример 1. $\varphi(2) = 1$; $\varphi(3) = 2$; $\varphi(4) = 2$; $\varphi(5) = 4$; $\varphi(6) = 2$; $\varphi(7) = 6$.

Т 1 (о вычислении значений функции Эйлера).

- 1) $\varphi(p) = p - 1$ для каждого простого числа p ;
- 2) $\varphi(p^s) = p^{s-1}(p - 1)$, если p – простое число;
- 3) если $(m, n) = 1$, то $\varphi(mn) = \varphi(m)\varphi(n)$;
- 4) если $m = p_1^{s_1} p_2^{s_2} \dots p_t^{s_t}$ – каноническое разложение числа m ,

то

$$\boxed{\varphi(m) = p_1^{s_1-1}(p_1 - 1)p_2^{s_2-1}(p_2 - 1)\dots p_t^{s_t-1}(p_t - 1).}$$

Т 2 [Эйлер, 1760]. Для $m \in \mathbb{N}, m > 1, a \in \mathbb{Z}$

$$(a, m) = 1 \Leftrightarrow a^{\phi(m)} \equiv 1 \pmod{m}.$$

Доказательство. $\Rightarrow)$ Пусть $x_1, x_2, \dots, x_{\phi(m)}$ – приведенная система вычетов по модулю m . В силу утверждения 2 при $(a, m) = 1$ числа $ax_1, ax_2, \dots, ax_{\phi(m)}$ также образуют приведенную систему вычетов по модулю m . Установим взаимно однозначное соответствие между этими двумя системами, поставив каждому из чисел $ax_1, ax_2, \dots, ax_{\phi(m)}$ сравнимое с ним число из системы $x_1, x_2, \dots, x_{\phi(m)}$ так, что

$$\begin{aligned} ax_1 &\equiv x_\alpha \pmod{m}, \\ ax_2 &\equiv x_\beta \pmod{m}, \\ &\dots, \\ ax_{\phi(m)} &\equiv x_v \pmod{m}, \end{aligned}$$

где $x_\alpha, x_\beta, \dots, x_v$ – это некоторым образом переставленные числа $x_1, x_2, \dots, x_{\phi(m)}$.

Перемножив все эти сравнения, получим

$$a^{\phi(m)} x_1 x_2 \dots x_{\phi(m)} \equiv x_\alpha x_\beta \dots x_v \pmod{m},$$

причем $x_\alpha x_\beta \dots x_v = x_1 x_2 \dots x_{\phi(m)}$, поскольку это те же числа, некоторым образом переставленные. Поскольку каждое из чисел $x_1, x_2, \dots, x_{\phi(m)}$ взаимно просто с m , то и их произведение взаимно просто с m . Поэтому полученное сравнение можно сократить на произведение $x_1 x_2 \dots x_{\phi(m)} = x_\alpha x_\beta \dots x_v$, что приводит к сравнению

$$a^{\phi(m)} \equiv 1 \pmod{m}.$$

$\Leftarrow)$ Поскольку $\phi(m) \geq 1$, то из $a^{\phi(m)} \equiv 1 \pmod{m}$ следует, что $\bar{a} \cdot \overline{a^{\phi(m)-1}} = \bar{1}$, а значит, \bar{a} – обратимый элемент в \mathbb{Z}_m . Отсюда в силу теоремы 2 §3 заключаем, что $(a, m) = 1$. \triangleleft

Следствие (малая теорема Ферма). Если p – простое число, a – целое число, то

$$(a, p) = 1 \Leftrightarrow a^{p-1} \equiv 1 \pmod{p}.$$

43. Линейные сравнения. Методы решения линейных сравнений.

Опр. 1. Решением сравнения (1) называется всякое целое число x_0 , которое удовлетворяет этому сравнению.

Легко понять, что в этом случае вместе с числом $\underline{x_0}$ сравнению удовлетворяют и все числа класса вычетов $\underline{x_0}$ по модулю m . Поэтому *класс вычетов по модулю m , числа которого удовлетворяют сравнению (1), считается за одно решение этого сравнения*. При таком соглашении сравнение (1) будет иметь столько решений, сколько классов вычетов по модулю m ему удовлетворяют. Поскольку полная система вычетов по модулю m состоит из m вычетов, то сравнение (1) может иметь только конечное количество решений или может не иметь их совсем.

Т 1. 1) Если $(a, m) = 1$, то сравнение (1) имеет единственное решение;

2) если $(a, m) = d > 1$ и $d \nmid b$, то сравнение (1) не имеет решений;

3) если $(a, m) = d > 1$ и $d \mid b$, то сравнение (1) имеет d решений.

Для доказательства первого утверждения теоремы отметим, что сравнение (1) равносильно диофантову уравнению

$$ax + my = b,$$

которое, согласно теореме 6 § 2, имеет решения тогда и только тогда, когда $d \mid b$, где $d = (a, m)$.

При этом множество всех решений диофантова уравнения описывается формулой $\left(x_0 + \frac{m}{d}t; y_0 - \frac{a}{d}t \right)$, где $(x_0; y_0)$ – частное решение этого уравнения, $d = (a, m)$, $t \in \mathbb{Z}$. Следовательно, решением сравнения (1) будет

$$x \equiv x_0 \left(\mod \frac{m}{d} \right),$$

что равносильно совокупности d сравнений по модулю m :

$$\begin{cases} x \equiv x_0 \pmod{m}, \\ x \equiv x_0 + \frac{m}{d} \pmod{m}, \\ \dots, \\ x \equiv x_0 + (d-1)\frac{m}{d} \pmod{m}. \end{cases}$$

Методы решения линейных сравнений

1. Метод перебора (подбора). При небольшом значении m сравнение $ax \equiv b \pmod{m}$ решается подбором.

При этом сравнение сокращают на $d = (a, m)$, а затем перебирают все классы вычетов по модулю m , подставляя их в сравнение.

2. Метод преобразования правой части сравнения путем добавления модуля. Сравнение сокращают на $d = (a, m)$, а для решения сравнения вида (1) с $d = (a, m) = 1$ рассматривают серию равносильных сравнений $ax \equiv b \pmod{m}$, $ax \equiv b + m \pmod{m}$, $ax \equiv b + 2m \pmod{m}$, ..., $ax \equiv b + km \pmod{m}$, ..., с целью получения в правой части числа $b + km$, делящегося на a , и сокращают.

3. Использование расширенного алгоритма Евклида.

Пусть $(a, m) = d$ и $d \mid b$. Тогда, в силу свойств сравнений,

$$ax \equiv b \pmod{m} \Leftrightarrow \frac{a}{d}x \equiv \frac{b}{d} \left(\bmod \frac{m}{d} \right),$$

причем, как следует из теоремы 1, последнее сравнение имеет единственное решение по модулю $\frac{m}{d}$.

С помощью расширенного алгоритма Евклида можно получить соотношение Безу для $(a, m) = d$ или $\left(\frac{a}{d}, \frac{m}{d} \right) = 1$ (коэффициенты Безу совпадают):

$$au_0 + mv_0 = d \Leftrightarrow \frac{a}{d}u_0 + \frac{m}{d}v_0 = 1.$$

Умножая последнее соотношение на b , получим

$$\frac{a}{d}bu_0 + \frac{m}{d}bv_0 = b \Leftrightarrow a\frac{b}{d}u_0 \equiv b \left(\bmod \frac{m}{d} \right),$$

а значит, решением преобразованного, а следовательно, и исходного сравнения является

$$x \equiv \frac{b}{d}u_0 \left(\bmod \frac{m}{d} \right).$$

Замечание. Сравнение всегда можно упростить, разделив обе части сравнения и модуль на их общий делитель:

$$ax \equiv b \pmod{m} \Leftrightarrow \frac{a}{d}x \equiv \frac{b}{d} \left(\bmod \frac{m}{d} \right).$$

4. Применение теоремы Эйлера. Пусть задано сравнение

$$ax \equiv b \pmod{m}, \text{ где } (a, m) = 1.$$

По теореме Эйлера $a^{\varphi(m)} \equiv 1 \pmod{m}$, откуда $a^{\varphi(m)}b \equiv b \pmod{m}$, или $a \cdot a^{\varphi(m)-1}b \equiv b \pmod{m}$. Следовательно, решением исходного сравнения будет

44. Китайская теорема об остатках.

Т 2 (китайская теорема об остатках). Пусть m_1, m_2, \dots, m_k – попарно взаимно простые натуральные числа, а c_1, c_2, \dots, c_k – целые числа. Тогда множество решений системы сравнений

$$\begin{cases} x \equiv c_1 \pmod{m_1}, \\ \dots, \\ x \equiv c_k \pmod{m_k} \end{cases}$$

имеет вид

$$x \equiv c_1 x_1 \frac{m}{m_1} + \dots + c_k x_k \frac{m}{m_k} \pmod{m},$$

где $m = [m_1, m_2, \dots, m_k]$, x_i – произвольное целое число, удовлетворяющее сравнению $x_i \frac{m}{m_i} \equiv 1 \pmod{m_i}$.

По существу, эта теорема утверждает, что можно восстановить целое число по множеству его остатков от деления на числа из некоторого набора попарно взаимно простых чисел.

На практике китайская теорема об остатках позволяет работать не с длинными числами, а с наборами их коротких по длине остатков, поскольку устанавливает взаимно однозначное соответствие между числом и множеством его остатков, определяемым набором взаимно простых чисел. Если в качестве базиса взять, к примеру, первые 500 простых чисел, длина каждого из которых не превосходит 12 бит, то этого хватит для представления десятичных чисел длиной до 1500 знаков.

Кроме того, вычисления по каждому из модулей можно выполнять параллельно.

45. Алгоритмы Диффи-Хеллмана и RSA.

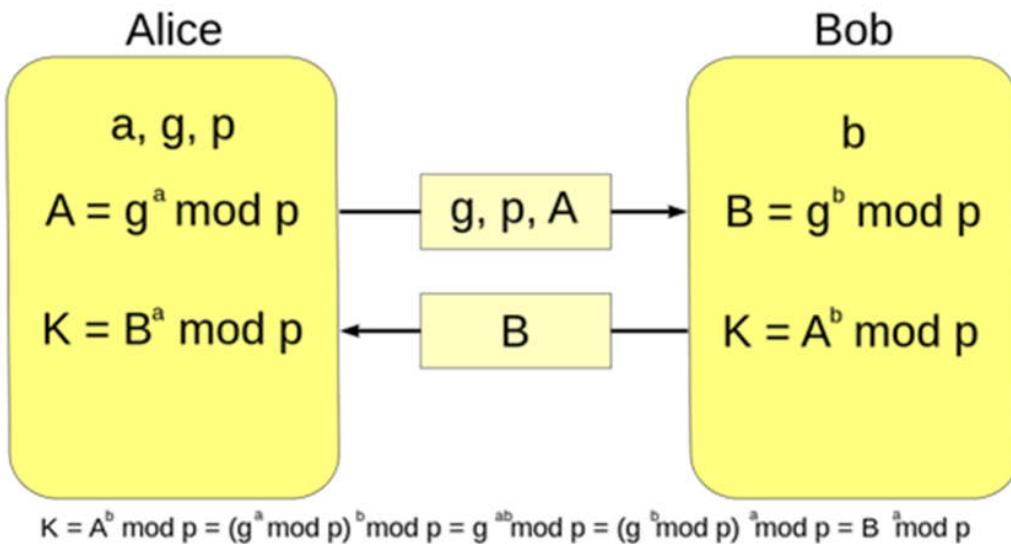
Алгоритм Диффи – Хеллмана создания секретного ключа заключается в следующем.

Пусть два пользователя (Алиса и Боб) хотят согласовать ключ – класс вычетов по модулю $m = p$ (p – простое число), с помощью которого они планируют шифровать свою переписку. При этом значение p общеизвестно; фиксируется и также не является секретным число g (в идеале, хотя и не обязательно, g – *первообразный корень* по модулю p).

На первом этапе каждая сторона выбирает некоторое число между 1 и $p - 1$, возводит g в выбранную степень (по модулю p) и посыпает результат партнеру: Алиса выбирает свой закрытый ключ – число a , $1 < a < p - 1$, и посыпает Бобу свой открытый ключ $A \equiv g^a \pmod{p}$; Боб выбирает свой закрытый ключ – число

$b, 1 < b < p - 1$, и отправляет Алисе свой открытый ключ $B \equiv g^b \pmod{p}$.

На втором этапе каждый из участников возводит полученный открытый ключ партнера в степень, равную своему закрытому ключу, и получает общий секретный ключ, поскольку $K \equiv B^a \equiv g^{ab} \pmod{p}$ и $K \equiv A^b \equiv g^{ab} \pmod{p}$ представляют один и тот же класс вычетов из множества \mathbb{Z}_p .



Таким образом, алгоритм Диффи – Хеллмана позволяет двум сторонам получить общий секретный ключ, используя незащищенный от прослушивания, но защищенный от подмены канал связи. Поскольку криптосистемы с открытым ключом значительно медленнее классических криптосистем, то они используются для генерации общего секретного ключа, который затем используется при обмене сообщениями с помощью классических симметричных криптосистем.

Безопасность криптосистемы Диффи – Хеллмана обеспечивается трудноразрешимостью задачи дискретного логарифмирования – задачи восстановления показателя степени в классах вычетов по известному результату при данных основании степени и модуле. В настоящее время не существует алгоритма, решающего эту задачу с полиномиальной сложностью.

Криптосистема RSA

В 1977 г. был изобретен первый алгоритм асимметричного шифрования RSA, который позволил решить проблему общения через незащищенный канал. Метод RSA назван по первым буквам фамилий его создателей Р. Ривеста, А. Шамира, Л. Адлемана и основан на трудноразрешимости задачи факторизации больших целых чисел, т. е. на различии в том, насколько легко находить большие простые числа и насколько сложно раскладывать на множители произведение двух больших простых чисел. Сложность наиболее быстрых алгоритмов факторизации целых чисел сравнима со сложностью решения задачи дискретного логарифмирования.

Алгоритм создания открытого и секретного ключей в крипто-системе RSA состоит из следующих шагов.

1. Выбираются два различных случайных простых числа p и q .
2. Вычисляется модуль $m = pq$.
3. Вычисляется $\varphi(m) = (p-1)(q-1)$.
4. Выбирается целое число $e, 1 < e < \varphi(m)$, взаимно простое с $\varphi(m)$.
5. Вычисляется число d , удовлетворяющее $de \equiv 1 \pmod{\varphi(m)}$.

Открытый ключ RSA – пара (e, m) ; закрытый ключ – (d, m) .

При шифровании допустимыми сообщениями являются числа N , $N < m$, $(N, m) = 1$. Зашифрованное сообщение вычисляется по формуле: $C \equiv N^e \pmod{m}$.

При расшифровывании сообщение N вычисляется по зашифрованному сообщению C по формуле: $N \equiv C^d \pmod{m}$.

Для взлома крипtosистемы RSA необходимо определить закрытый ключ по открытому ключу, т. е. решить задачу факторизации целого числа m , представляющего собой произведение двух больших простых чисел p и q .

46. Группа, подгруппа. Порядок группы. Теорема Лагранжа. Абелева группа.

Опр. 5. Группой называется непустое множество G с определенной на нем бинарной алгебраической операцией $*$, которая обладает свойствами:

- 1) ассоциативность $(a * b) * c = a * (b * c)$ для любых $a, b, c \in G$;
- 2) существует *нейтральный (единичный)* элемент, т. е. такой элемент $e \in G$, что $e * a = a * e = a$ для каждого $a \in G$;
- 3) каждый элемент $a \in G$ имеет *обратный*, т. е. такой элемент $b \in G$, что $a * b = b * a = e$.

Опр. 6. Группа с коммутативной операцией называется *коммутативной* или *абелевой группой*.

По количеству элементов группы делятся на *конечные* и *бесконечные*.

Опр. 7. Число элементов конечной группы G называется *порядком группы* и обозначается $|G|$.

Опр. 8. Непустое подмножество H группы G называется *подгруппой* этой группы, если H само является группой относительно той же бинарной алгебраической операции.

Пример 7. 1) $(\mathbb{Z}, +) \subset (\mathbb{Q}, +) \subset (\mathbb{R}, +) \subset (\mathbb{C}, +)$;

2) $(\mathbb{Q}^*, \times) \subset (\mathbb{R}^*, \times) \subset (\mathbb{C}^*, \times)$;

3) $(\mathbb{Z}, +) \supset (2\mathbb{Z}, +) \supset (4\mathbb{Z}, +) \supset (8\mathbb{Z}, +) \supset \dots \supset (2^k \mathbb{Z}, +) \supset \dots$, где $n\mathbb{Z} = \{nq : q \in \mathbb{Z}\}$, $k \in \mathbb{N}$. •

Т 1 (теорема Лагранжа). Порядок конечной группы делится на порядок любой ее подгруппы.

47. Порядок элемента группы. Циклические группы. Циклические группы

Т 2. Пусть a – фиксированный элемент группы G . Тогда множество всевозможных целых степеней элемента a

$$\langle a \rangle = \{a^0 = e; a; a^2; \dots; a^{-1}; a^{-2}; \dots\} = \{a^k, k \in \mathbb{Z}\}$$

является подгруппой группы G , причем эта подгруппа абелева.

Замечание. Под k -й степенью элемента группы понимается k -кратное применение бинарной операции группы к этому элементу, если $k > 0$, и к его обратному элементу, если $k < 0$:

$$a^k = \underbrace{a * a * \dots * a}_{k \text{ множителей}}; \quad a^{-k} = \underbrace{a^{-1} * a^{-1} * \dots * a^{-1}}_{k \text{ множителей}}, \quad k \in \mathbb{N}; \quad a^0 = e.$$

При этом в силу ассоциативности бинарной операции группы

$$a^k * a^n = a^{k+n}, \quad (a^k)^n = a^{kn}, \quad k, n \in \mathbb{Z}.$$

Опр. 9. Подгруппа $\langle a \rangle$ называется *циклической подгруппой, порожденной элементом a* .

Т 3. Всякая циклическая группа абелева.

Доказательство. Для произвольных элементов группы $\langle a \rangle$ в силу ассоциативности операции в группе имеем $a^k * a^n = a^{k+n} = a^{n+k} = a^n * a^k$. \square

Т 4. Всякая подгруппа циклической группы является циклической.

Доказательство. Пусть $G = \langle a \rangle$ и H – подгруппа этой группы, отличная от G и $\{e\}$. Тогда найдется натуральное k такое, что $a^k \in H$. Возьмем такое наименьшее натуральное k , что $a^k \in H$, и покажем, что $H = \langle a^k \rangle$.

Для произвольного элемента $h \in H$, поскольку он также является элементом циклической группы $G = \langle a \rangle$, найдется такое целое s , что $h = a^s$. По теореме о делении с остатком $s = kq + r$, где q, r – целые числа, $0 \leq r < k$. Тогда $h = (a^k)^q * a^r$. Следовательно, $a^r = h * (a^k)^{-q} \in H$, что, в силу минимальности k возможно только при $r = 0$. Значит, всякий элемент $h \in H$ представим в виде $h = (a^k)^q$ и $H = \langle a^k \rangle$. \square

Порядок элемента группы

Опр. 11. Натуральное число n называется *порядком элемента $a \in G$* , если $a^n = e$ и $a^k \neq e$ для всех натуральных $k, 1 \leq k < n$. Если $a^k \neq e$ при всех натуральных k , то элемент $a \in G$ называется *элементом бесконечного порядка*.

Пример 11. 1) Элемент 1 в группе $(\mathbb{Z}, +)$ имеет бесконечной порядок;

2) элемент $\bar{1}$ в группе $(\mathbb{Z}_m, +)$ имеет порядок m . \bullet

Т 5. Если $a \in G$ имеет порядок n , то циклическая подгруппа $\langle a \rangle$ имеет порядок n и

$$\boxed{\langle a \rangle = \{a; a^2; \dots; a^n = e\}.}$$

Доказательство. Для любого целого k , разделив его с остатком на n , получим $k = nq + r$, где – целые числа, $0 \leq r < n$, откуда

$$a^k = (a^n)^q * a^r = e^q * a^r = a^r,$$

т. е. любой элемент циклической подгруппы $\langle a \rangle$ представим в виде a^r , где $0 \leq r < n$. \square

48. Кольцо. Коммутативное кольцо. Кольцо с единицей. Делители нуля.

Мультиликативная группа кольца.

Опр. 1. *Кольцо* – непустое множество K с двумя бинарными алгебраическими операциями $+$ (сложение) и \cdot (умножение), такими, что:

1) K является *абелевой группой* относительно операции сложения $+$;

2) операция умножения ассоциативна, т. е.

$$(a \cdot b) \cdot c = a \cdot (b \cdot c) \text{ для всех } a, b, c \in K;$$

3) операции умножения и сложения связаны законами *дистрибутивности* (умножение дистрибутивно по сложению): для произвольных $a, b, c \in K$

$$(a + b) \cdot c = a \cdot c + b \cdot c;$$

$$c \cdot (a + b) = c \cdot a + c \cdot b.$$

Условимся нейтральный элемент аддитивной группы кольца называть *нулем* и обозначать символом 0 ; противоположный к элементу a элемент обычно обозначают через $-a$; вместо $a + (-b)$ пишут $a - b$. Знак \cdot операции умножения при записи произведений элементов кольца будем, как правило, опускать.

Опр. 2. Кольцо K называется *коммутативным*, если операция умножения в нем коммутативна, т. е. $ab = ba$ для всех $a, b \in K$.

Опр. 3. Кольцо K называется *кольцом с единицей*, если оно имеет мультиликативную единицу, т. е. такой элемент e , что $ea = ae = a$ для каждого $a \in K$.

Пример 2. Все кольца из примеров 1.1)–4) являются коммутативными кольцами с единицей.

Множество квадратных матриц данного порядка (пример 1.5)) представляет собой пример некоммутативного кольца с единицей.

Примером коммутативного кольца без единицы является при $m > 1$ кольцо $m\mathbb{Z} = \{ma : a \in \mathbb{Z}\}$ – множество целых чисел, кратных m . •

Утв. 1. Если K – кольцо с единицей, содержащее более одного элемента, то в нем $e \neq 0$ (единичный элемент не равен нулю).

Доказательство. Так как K содержит более одного элемента и нулевой элемент $0 \in K$, то найдется еще один элемент кольца $a \neq 0$.

Если допустить, что $e = 0$, то из $ae = a \cdot 0 = 0$ следует, что $a = 0$, т. е. приходим к противоречию. Значит, предположение неверно и $e \neq 0$. □

Опр. 4. Если в кольце K найдутся *ненулевые* элементы a и b такие, что $ab = 0$, то их называют *делителями нуля*.

Т 1. Если K – кольцо с единицей, то множество K^* обратимых относительно умножения элементов кольца K есть группа относительно умножения.

Доказательство. 1) Проверим аксиомы группы.

1. Умножение в K^* ассоциативно, так как K – кольцо, и умножение в нем ассоциативно, а K^* – подмножество множества K .

2. Если e – нейтральный элемент кольца K относительно умножения, то $e \cdot e = e$, т. е. обратным к e элементом является $e \Rightarrow e \in K^*$.

3. Для любого $a \in K^*$ обратный элемент $a^{-1} \in K^*$, так как $a^{-1}a = aa^{-1} = e \in K^*$, а значит, $(a^{-1})^{-1} = a$, элемент a^{-1} обратим, поэтому $a^{-1} \in K^*$.

2) Проверим, что множество K^* замкнуто относительно умножения.

Пусть $a, b \in K^*$ и a^{-1}, b^{-1} – обратные элементы к a и b соответственно, причем $a^{-1}, b^{-1} \in K^*$. Покажем, что $ab \in K^*$, т. е. что элемент ab имеет обратный. Поскольку

$$(ab)(b^{-1}a^{-1}) = a(bb^{-1})a^{-1} = aea^{-1} = e;$$

$$(b^{-1}a^{-1})(ab) = b^{-1}(a^{-1}a)b = b^{-1}eb^{-1} = e,$$

то обратным к элементу $ab \in K$ является элемент $b^{-1}a^{-1} \in K$, а значит, $ab \in K^*$. \triangleleft

Опр. 5. Множество K^* обратимых относительно умножения элементов кольца K называют **мультипликативной группой кольца K** .

49. Поле. Свойства полей. Подполе, простое подполе. Изоморфизм полей. Теорема об изоморфизме простых подполей.

Опр. 6. Поле – коммутативное кольцо с единицей, в котором каждый ненулевой элемент обратим.

Пример 5. 1) \mathbb{Q} – поле рациональных чисел, \mathbb{R} – поле вещественных чисел, \mathbb{C} – поле комплексных чисел с естественными операциями сложения и умножения.

2) Множество классов вычетов \mathbb{Z}_m является полем тогда и только тогда, когда $m = p$ – простое число. При этом \mathbb{Z}_p – конечное поле из p элементов. •

Свойства полей.

1. В поле нет делителей нуля.

Доказательство. Допустим, в поле P существуют делители нуля $a, b \in P$, т. е. произведение $ab = 0$ и $a \neq 0, b \neq 0$. Поскольку в поле каждый ненулевой элемент обратим, то существует $a^{-1} \in P^*$. Тогда, с одной стороны, $a^{-1}(ab) = a^{-1} \cdot 0 = 0$, а с другой, $a^{-1}(ab) = (a^{-1}a)b = eb = b$, откуда получаем, что $b = 0$, что противоречит предположению. Следовательно, в поле нет делителей нуля. \triangleleft

2. Мультипликативная группа поля содержит все его ненулевые элементы: $P^* = P \setminus \{0\}$.

3. Если $(P, +, \cdot)$ – поле, то $(P, +)$ – аддитивная абелева группа, (P^*, \cdot) – мультипликативная абелева группа. (Здесь $P^* = P \setminus \{0\}$.)

Опр. 8. Подмножество F поля P называется **подполем** поля P , если оно замкнуто относительно имеющихся операций сложения и умножения и само является полем относительно этих операций. При этом поле P называют **расширением** поля F .

Подполе F называется **собственным подполем** поля P , если $F \neq P$.

Пример 7. Поле рациональных чисел \mathbb{Q} является собственным подполем поля вещественных чисел \mathbb{R} , которое в свою очередь будет собственным подполем поля комплексных чисел \mathbb{C} . •

Опр. 10. Поля P_1 и P_2 называются **изоморфными** ($P_1 \cong P_2$), если они изоморфны как кольца.

Утв. 4. Пересечение любого количества подполей данного поля P также является подполем P .

Опр. 11. Поле, не содержащее собственных подполей, называется **простым** или **минимальным**.

Т 2. В каждом поле P содержится одно и только одно простое подполе F . Это поле F изоморфно либо полю \mathbb{Q} , либо полю \mathbb{Z}_p при некотором простом p .

50. Многочлены с коэффициентами из произвольного поля. Кольцо многочленов. Построение конечных полей порядка p^n .

Построение конечного поля как множества классов вычетов по модулю неприводимого многочлена с коэффициентами из

$$\mathbb{Z}_p$$

Рассмотрим множество многочленов с коэффициентами из \mathbb{Z}_p :

$$\mathbb{Z}_p[x] = \left\{ a(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n : n \in \mathbb{N}, a_0, a_1, a_2, \dots, a_n \in \mathbb{Z}_p \right\}.$$

Операции сложения и умножения элементов $\mathbb{Z}_p[x]$ определяются обычным образом с учетом того, что все действия над коэффициентами осуществляются в \mathbb{Z}_p (по модулю p).

Утв. 5. $\mathbb{Z}_p[x]$ (p – простое) является коммутативным кольцом с единицей и без делителей нуля, но не является полем.

Элементы поля \mathbb{Z}_p также являются элементами кольца $\mathbb{Z}_p[x]$, поскольку могут рассматриваться как многочлены нулевой степени. Отметим, что нулевым и единичным элементами кольца $\mathbb{Z}_p[x]$ являются соответственно нулевой $\bar{0}$ и единичный $\bar{1}$ элементы поля \mathbb{Z}_p , а обратными элементами кольца $\mathbb{Z}_p[x]$ являются только ненулевые элементы поля \mathbb{Z}_p .

Т 3 (о делении с остатком). Для любых $a(x), b(x) \in \mathbb{Z}_p[x]$, где $b(x) \neq \bar{0}$, существуют и единственны $q(x), r(x) \in \mathbb{Z}_p[x]$, такие, что

$$a(x) = b(x)q(x) + r(x), \text{ где } 0 \leq \deg r(x) < \deg b(x) \text{ или } r(x) = \bar{0}.$$

В этом случае многочлен $r(x)$ называется *остатком от деления* $a(x)$ *на* $b(x)$.

Пусть $m(x) \in \mathbb{Z}_p[x]$ – фиксированный многочлен степени $n \geq 1$. Рассматривая различные остатки от деления на $m(x)$ в $\mathbb{Z}_p[x]$, можно разбить множество $\mathbb{Z}_p[x]$ на классы эквивалентности – классы вычетов по модулю $m(x)$.

Опр. 13. Классом вычетов по модулю $m(x)$ называется множество всех многочленов из $\mathbb{Z}_p[x]$, имеющих один и тот же остаток от деления на $m(x)$, т. е.

$$\overline{r(x)} = \left\{ a(x) \in \mathbb{Z}_p[x] : a(x) = m(x)q(x) + r(x), \deg r(x) < \deg m(x) \text{ или } r(x) = \bar{0} \right\}.$$

Опр. 14. Множество всех классов сравнимых друг с другом по модулю $m(x)$ многочленов из $\mathbb{Z}_p[x]$ называют *множеством классов вычетов по модулю* $m(x)$ и обозначают через $\mathbb{Z}_p[x]/(m(x))$.

Отметим, что множество $\mathbb{Z}_p[x]/(m(x))$ содержит конечное число элементов. Действительно, если $\deg m(x) = n$, то остатками от деления на $m(x)$ могут быть только многочлены степени меньше n , причем каждый коэффициент – это элемент \mathbb{Z}_p , т. е. выби-

51. Поля Галуа. Характеристика поля. Теорема о существовании и единственности конечного поля порядка p^n . Свойства конечных полей.

Конечные поля, или поля Галуа

Опр. 12. Поле P называется **конечным**, если число его элементов конечно. Число элементов в поле называется его **порядком**.

Пример 8. 1) Поле \mathbb{Q} рациональных чисел, поле \mathbb{R} вещественных чисел, поле \mathbb{C} комплексных чисел – бесконечные поля, причем $\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$.

2) Поле классов вычетов \mathbb{Z}_p – конечное поле из p элементов, если p – простое число. •

Конечное поле порядка q обозначается \mathbb{F}_q или $GF(q)$ (сокращение от *Galois Field*) и называется полем Галуа; понятие конечного поля в его общем значении (когда имеются в виду не только поля, изоморфные \mathbb{Z}_p) впервые появилось в 1830 г. в статье Э. Галуа.

Характеристика поля

Опр. 16. Если для поля P существует такое натуральное n , что сумма n единиц поля (n раз складывается с самим собой нейтральный относительно умножения элемент поля) равна 0 (нейтральному элементу относительно сложения), то наименьшее n с таким свойством называется **характеристикой поля** P и обозначается $\text{char } P$. Если в поле P любая конечная сумма единиц отлична от нуля, то говорят, что характеристика поля равна 0.

Пример 11. 1) $\text{char } \mathbb{Q} = \text{char } \mathbb{R} = \text{char } \mathbb{C}$.

2) $\text{char } \mathbb{Z}_p = p$.

3) Если $m(x)$ – неприводимый многочлен степени n над полем \mathbb{Z}_p , то $F = \mathbb{Z}_p[x]/(m(x))$ – конечное поле порядка p^n , однако $\text{char } F = p$, поскольку единица и ноль этого поля представляют собой классы многочленов, имеющих при делении на $m(x)$ остатки, равные соответственно единице и нулю поля \mathbb{Z}_p , поэтому сумма p единиц равна 0. •

Пример 12. Примером бесконечного по количеству элементов поля конечной характеристики является поле рациональных функций над \mathbb{Z}_p :

$$\mathbb{Z}_p(x) = \left\{ \frac{a(x)}{b(x)} : a(x), b(x) \in \mathbb{Z}_p[x], b(x) \neq 0 \right\}. •$$

Утв. 6. Если характеристика поля отлична от 0, то она является простым числом.

Доказательство следует из того, что если бы характеристика поля была составным числом, то в поле были бы делители нуля. ◁

Утв. 7. Если подполе поля P имеет характеристику p , то и поле P имеет ту же характеристику, и все подполя поля P имеют ту же характеристику.

Доказательство следует из единственности нейтрального элемента в группе и, следовательно, из единственности единицы в любом поле. ◁

Т 5 (о существовании и единственности конечного поля)
[Мур, 1893]. Для каждого простого числа p и любого натурального числа n существует конечное поле \mathbb{F}_q из $q = p^n$ элементов. Поле \mathbb{F}_q единственно с точностью до изоморфизма.

Замечание. Поле \mathbb{F}_q с $q = p^n$ обозначают также \mathbb{F}_p^n .

Следствие. Поле \mathbb{F}_p^n изоморфно полю $\mathbb{Z}_p[x]/(m(x))$ для любого неприводимого полинома $m(x)$ степени n из кольца $\mathbb{Z}_p[x]$.

Свойства конечных полей

Всякое конечное поле F :

- имеет простую характеристику $p > 1$;
- содержит простое (т. е. не содержащее нетривиальных подполей) подполе \mathbb{F}_p из p элементов, изоморфное полю \mathbb{Z}_p ;
- содержит $q = p^n$ элементов для некоторого натурального n ;
- изоморфно полю $\mathbb{Z}_p[x]/(m(x))$ для любого неприводимого над \mathbb{Z}_p полинома $m(x)$ степени n .

Т 6. Мультипликативная группа конечного поля – циклическая.

52. Цикличность мультипликативной группы конечного поля.