Воронежский государственный технический университет

Кафедра прикладной математики и механики

Некрасова Н.Н.

МЕТОДИЧЕСКИЕ УКАЗАНИЯ

для выполнения контрольных работ № 4

Воронеж 2010

I. Теория вероятностей

Формулы комбинаторики

При подсчете вероятностей часто бывают, полезны так называемые формулы комбинаторики, к описанию которых мы сейчас и перейдем.

Пусть имеется множество A, состоящее из n различных элементов, и пусть $k \le n$..

Размещением из n элементов по k называется любое подмножество множества A, содержащее k элементов, *расположенных в определенном порядке*. Таким образом, одно размещение, отличается от другого *либо составом* элементов, *либо порядком* следования элементов.

Если $k=n$, то размещение из n элементов по n называется *перестановкой.*

Сочетанием из n элементов по k называется любое подмножество множества A, содержащее k элементов. Таким образом, одно сочетание отличается от другого только составом элементов. Порядок следования элементов не важен.

Число всех возможных размещений из n элементов по k обозначается A_n^k и вычисляется по формуле

$$A_n^k = n(n-1)...(n-k+1). \qquad (1.1)$$

Число всех возможных перестановок из n элементов обозначается P_n и вычисляется по формуле

$$P_n = A_n^n = n(n-1)...1 = n!. \qquad (1.2)$$

Число всех возможных сочетаний из n элементов по k обозначается C_n^k и вычисляется по формуле

$$C_n^k = \frac{A_n^k}{P_n} = \frac{n!}{k!(n-k)!}. \qquad (1.3)$$

Основные понятия и формулы теории вероятности

Опыт – это комплекс условий, при выполнении которых может появиться то или иное событие. Событие называется детерминированным, если в результате опыта оно всегда наступает.

Событие называется случайным, если в результате опыта оно может как наступить, так и не наступить.

Теория вероятностей изучает массовые случайные события, опыт по получению которых можно повторять как угодно много раз, а частота появления события с ростом числа опытов приближается к некоторому пределу (свойство устойчивости частоты).

События обозначаются большими буквами латинского алфавита.

Пример. Опыт состоит в том, что студент, прослушав курс теории вероятностей и выполнив необходимые задания, сдает экзамен.

Возможные события: A - студент сдал экзамен на 5, B - сдал на 4, C - сдал на 3, D - не сдал экзамен, E - студент сдал экзамен (на какую оценку не важно).

Пример. Опыт состоит в том, что игральная кость бросается один раз. Возможные события: A - выпала цифра, большая 3, B - выпало четное число, C - выпало простое число, и т. д.

В теории вероятностей степень возможности наступления данного события A в результате опыта принято выражать числом от 0 до 1. Это число называют вероятностью события и обозначают P(A).

Предполагается, что для большинства недетерминированных явлений такая степень возможности их наступления объективно существует.

По определению $0 \le P(A) \le 1$. Таким образом, P(A) есть числовая функция, определенная на множестве всех событий, которые могут произойти в опыте.

События A и B называются несовместными, если они не могут произойти одновременно (в одном опыте).

События называются независимыми, если вероятность наступления одного не зависит от того, наступило или нет другое событие.

События называются равновероятными, если их вероятности равны.

События образуют полную группу, если в каждом опыте одно из них обязательно наступает.

Пусть A и B некоторые события. *Событие, которое наступает, когда наступает событие A или событие B называется суммой этих событий и обозначается A+B.*

Событие, которое наступает, когда наступают A и B одновременно, называется произведением этих событий и обозначается AB.

Событие называется противоположным, если оно наступает тогда и только тогда, когда не наступает само событие. Событие, противоположное событию A обозначается $\overline{A}$.

Событие называется элементарным, если его нельзя представить в виде суммы других более простых событий.

Опыт называется *схемой случаев*, если для него существует полная группа элементарных равновероятных несовместных событий. Такие события называются исходами опыта.

Если опыт есть схема случаев, то вероятность любого события A определяется равенством:

$$P(A) = \frac{m}{n}, \tag{1.4}$$

где n – число всех исходов опыта, а m – число исходов опыта, благоприятствующих появлению события A (то есть таких исходов, при которых событие A наступает). Это так называемая *классическая формула вычисления вероятности.*

При статистическом определении в качестве вероятности события A приближенно принимают его относительную частоту

$$P(A) \approx W(A) = \frac{m}{n}, \qquad (1.5)$$

где m – число испытаний, в которых событие A наступило, а n – общее число произведенных испытаний. Чем больше n, тем точнее формула.

Пример. Два раза бросается монета. Какова вероятность того, что хотя бы один раз выпадет “орел” (событие A)?

Решение. Равновозможными несовместными элементарными исходами являются четыре пары исходов (о,о), (о,р), (р,о) и (р,р), где буквами “о” и “р” обозначены выпадение, соответственно, орла или решки; общее число элементарных исходов $n = 4$. Событию A соответствуют первые три пары исходов, поэтому число благоприятных исходов $m = 3$. Следовательно,

$$p(A) = \frac{m}{n} = \frac{3}{4}.$$

Ответ: $p(A) = 0{,}75$.

Пример. Студент идет на экзамен, зная 16 вопросов из 25. Найти вероятность того, что из трех вопросов два будут ему известны?

Решение. Множество всех элементарных исходов равно количеству способов, сколькими можно выбрать 3 вопроса из 25, а так как порядок следования вопросов в билете не важен, то это будет число сочетаний $n = C_{25}^{3}$.

Определим число исходов, благоприятных событию A (среди трех вопросов два известны). Два известных вопроса можно выбрать C_{16}^{2} способами, а один неизвестный из оставшихся 9 невыученных (25–16=9) можно выбрать C_{9}^{1} способами. Общее число благоприятных событию A исходов равно произведению полученных вариантов: $m = C_{16}^{2} C_{9}^{1}$. Следовательно,

$$p(A) = \frac{m}{n} = \frac{C_{16}^{2} C_{9}^{1}}{C_{25}^{3}} = \left| \begin{array}{l} C_{16}^{2} C_{9}^{1} = \dfrac{16! \cdot 9}{(16-2)!2!} = \dfrac{15 \cdot 16}{2!} \cdot 9 = 1080 \\ C_{25}^{3} = \dfrac{25!}{(25-3)!3!} = \dfrac{23 \cdot 24 \cdot 25}{3!} = 2300 \end{array} \right| \approx 0{,}47.$$

Ответ: $p(A) = 0{,}47$.

Пример. Сборная конструкция состоит из четырех объемных элементов, которые завозятся на стройку по одному случайным образом. Какова вероятность того, что конструкция будет смонтирована без простоев, связанных с завозом, если 1) все элементы различны; 2) среди них два одинаковых?

Решение. Общее решение различных вариантов очередности завоза для четырех элементов равно числу перестановок $n = P_4 = 4! = 24$.

В первом случае благоприятным исходом m_1 будет один, т.к. задержек не будет только в том случае, если элементы конструкции завозятся в том же порядке, в котором они должны быть смонтированы. Следовательно, $p_1 = \frac{m_1}{n} = \frac{1}{24} \approx 0{,}04$.

Во втором случае число благоприятных исходов m_2 равно двум, так как одинаковые элементы можно поменять между собой. Следовательно, $p_2 = \frac{m_2}{n} = \frac{2}{24} \approx 0{,}08$.

Ответ: $p_1 \approx 0{,}04$, $p_2 \approx 0{,}08$.

Пример. Бросают две игральные кости. Какова вероятность того, что сумма выпавших очков равна 8 (событие *A*)?

Решение. Равновозможными элементарными исходами здесь являются пары (*x, y*), где *x* и *y* – *цифры, выпавшие на первой и второй кости, соответственно x, y* = 1, 2, …6. Общее число элементарных исходов *n* = 36.

Событию *A* благоприятствуют пары (6, 2), (2, 6), (5, 3), (3, 5), (4, 4), число которых *m* =5.

Следовательно, $P(A) = \frac{m}{n} = \frac{5}{36}$.

Пример. Студент пришел на зачет, зная из 30 вопросов только 24. Какова вероятность сдать зачет, если после отказа отвечать на первый вопрос преподаватель задает еще один вопрос?

Решение. Рассмотрим следующие события: A – студент сдал зачет, B_1 - студент ответил на первый вопрос, $\overline{B_1}$ - студент не ответил на первый вопрос, B_2 - студент ответил на второй вопрос, $\overline{B_2}$ - студент не ответил на второй вопрос.

Очевидно, что $A = B_1 + \overline{B_1} \cdot B_2$, причем события B_1 и $\overline{B_1} \cdot B_2$ являются несовместными. Следовательно,

$$P(A) = P(B_1) + P(\overline{B_1} \cdot B_2) = P(B_1) + P(\overline{B_1}) \cdot P_{B_1}(B_2) =$$

$$= \frac{24}{30} + \frac{6}{30} \cdot \frac{24}{29} = \frac{24}{30}\left(1 + \frac{6}{29}\right) = \frac{24 \cdot 35}{30 \cdot 29} = \frac{28}{29} \approx 0{,}966.$$

Задачу можно было решить другим способом:

$$P(A) = 1 - P(\overline{A}) = 1 - P(\overline{B_1} \cdot \overline{B_2}) = 1 - P(\overline{B_1}) \cdot P_{B_1}(\overline{B_2}) =$$

$$= 1 - \frac{6}{30} \cdot \frac{5}{29} = 1 - \frac{1}{29} = \frac{28}{29} \approx 0{,}966.$$

Теорема сложения вероятностей. Вероятность появления хотя бы одного из двух событий равна сумме их вероятностей без вероятности их совместного появления:

$$P(A+B)=P(A)+P(B)-P(A \cdot B). \qquad (1.6)$$

Если события *A* и *B* несовместны (т.е. в результате опыта они не могут появиться одновременно), то *P(AB)=0* и

$$P(A+B)=P(A)+P(B). \qquad (1.7)$$

При решении задач часто вычисляют вероятность противоположного события *A,* а затем находят вероятность самого события *A* по формуле

$$P(A)=1-P(\overline{A}). \qquad (1.8)$$

Вероятность появления события *A* при условии, что в опыте наступило событие *B* называется *условной вероятностью события A при условии B*, и обозначается $P_B(A)$.

<u>Теорема умножения вероятностей.</u> Вероятность одновременного наступления двух событий равна вероятности одного из них, умноженной на условную вероятность другого

$$P(A\cdot B)=P(A)\cdot P_A(B)=P(B)\cdot P_B(A). \qquad (1.9)$$

Если события *A* и *B* независимы (т.е. появление одного из них не влияет на вероятность появления другого), то

$$P(A\cdot B)=P(A)\cdot P(B). \qquad (1.10)$$

Замечание. *Теоремы сложения и умножения вероятностей можно обобщить на большее число слагаемых или сомножителей. Например, для трех событий*

$$P(A+B+C)=P(A)+P(B)+P(C)-P(A\cdot B)-P(A\cdot C)- \qquad (1.11)$$
$$-P(B\cdot C)+P(A\cdot B\cdot C);$$
$$P(A\cdot B\cdot C)=P(A)\cdot P_A(B)\cdot P_{A\cdot B}(C).$$

Для нахождения суммы независимых событий $A_1, A_2, ..., A_n$ выгодно перейти к противоположным событиям:

$$P(A_1+A_2+...+A_n)=1-P(\overline{A_1})\cdot P(\overline{A_2})\cdot ...\cdot P(\overline{A_n}). \qquad (1.12)$$

Если событие *A* может наступать только при появлении одного из несовместных событий (гипотез) $H_1, H_2, ..., H_n$, образующих полную группу событий, то вероятность события *A* вычисляется по формуле полной вероятности:

$$P(A)=\sum_{i=1}^{n} P_{H_i}(A)\cdot P(H_i). \qquad (1.13)$$

где $P(H_i)$- вероятность гипотезы, $H_i, P_{Hi}(A)$ – условная вероятность события *A* при этой гипотезе. Для полной группы несовместных событий $\sum_{i=1}^{n} P(H_i)=1.$

Вероятность того, что в *n* независимых испытаниях "успех" наступит ровно *k* раз, выражается формулой Бернулли

$$P(k)=C_n^k p^k q^{n-k}, \qquad (1.14)$$

где *p* – вероятность появления "успеха" в каждом испытании, *q=1-p* – вероятность "неудачи".

В случае, когда *n* велико, а *p* – мало (обычно $p\le 0{,}1$) вместо формулы Бернулли применяют приближенную формулу Пуассона:

$$P_n(k) \approx \frac{\lambda^k e^{-\lambda}}{k!}, \qquad (1.15)$$

где $\lambda = n \cdot p$.

Таблица значений функции $\frac{\lambda^k e^{-\lambda}}{k!}$ приведена в [1, 2].

Вероятность того, что в n независимых испытаниях число успехов k находится между k_1 и k_2 равна

$$P_n(k_1, k_2) = P_n(k_1) + P_n(k_1 + 1) + \ldots + P_n(k_2), \qquad (1.16)$$

где $P_n(k_1), P_n(k+1), \ldots, P_n(k_2)$ вычисляют по формуле Бернулли или по формуле Пуассона.

Пользоваться формулой (1.14) при больших $\boldsymbol{n}$ затруднительно из-за большого объема счетной работы. В этом случае применяется приближенная формула – **локальная теорема Лапласа**:

$$\boldsymbol{p_n(k) \approx \frac{1}{\sqrt{npq}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}}, \qquad (1.17)$$

где $\boldsymbol{x = \frac{k - np}{\sqrt{npq}}}$, а значения функции $\boldsymbol{\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}}$ находятся по таблице из [1, прил.1]. Очевидно, $\boldsymbol{\varphi(x)}$ четная функция.

Вероятность того, что событие $\boldsymbol{A}$ в $\boldsymbol{n}$ испытаниях произойдет от $\boldsymbol{k_1}$ до $\boldsymbol{k_2}$ раз (включительно), приближенно считается по **интегральной теореме Лапласа**:

$$\boldsymbol{p_n(k_1, k_2) \approx \Phi(x_2) - \Phi(x_1)}, \qquad (1.18)$$

где $\boldsymbol{x_1 = \frac{k_1 - np}{\sqrt{npq}}}$ и $\boldsymbol{x_2 = \frac{k_2 - np}{\sqrt{npq}}}$, а значения функции Лапласа $\boldsymbol{\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{z^2}{2}} dz}$ находятся по таблице из [1, прил. 2]. Заметим, что $\boldsymbol{\Phi(x)}$ нечетная функция.

Отметим, что очень часто найти вероятность $\boldsymbol{p}$ по классическому определению вероятности невозможно. В этом случае на практике используют схему Бернулли при достаточно большом числе независимых испытаний. При **статистическом определении** в качестве **вероятности** события $\boldsymbol{A}$ принимают его относительную частоту

$$\boldsymbol{W(A) = \frac{n_A}{n}}, \qquad (1.19)$$

где $\boldsymbol{n_A}$ - количество появлений события $\boldsymbol{A}$ в $\boldsymbol{n}$ испытаниях.

Например, нужно практически определить вероятность попадания стрелка в мишень. Будем считать, что $\boldsymbol{p(A)} = 0{,}7$, если он попал 7 раз из 10, или 70 раз из 100, или 700 раз из 1000 и т.д. При этом, чем больше произведено выстрелов, тем ответ достовернее. Чем больше число испытаний, тем численно ближе

относительная частота к неизвестной нам теоретической вероятности события $\boldsymbol{A}$: $\boldsymbol{W(A) \to p(A)}$ при $n \to \infty$.

Пример. Прибор состоит из пяти узлов. Вероятность безотказной работы в течение времени t для каждого узла равна 0,9. Узлы выходят из строя независимо друг от друга. Найти вероятность того, что за время t откажут два узла.

Решение. Пусть событие A есть выход узла из строя за время t. Число узлов $n = 5$, а число отказавших узлов $k = 2$.

Вероятность безотказной работы узла $p(A) = p = 0{,}9$, тогда $q = 1 - p = 1 - 0{,}9 = 0{,}1$. Вычислим искомую вероятность по формуле Бернулли:

$$p_5(2) = C_5^2 0{,}9^2 0{,}1^{5-2} = \frac{5!}{3!2!} 0{,}81 \cdot 0{,}001 = 0{,}0081.$$

Ответ: $p_5(2) = 0{,}0081$.

Пример. Два шахматиста, первый из которых выигрывает в два раза чаще другого, играют матч из нескольких партий. Считая все партии результативными, найти

1) вероятность выигрыша матча из трех партий первым игроком;
2) вероятность выигрыша первым игроком 13 партий из 18.

Решение.

1) Пусть событие A – матч выигран первым игроком, т. е. первый игрок выиграет две или три партии из трех. Вероятность выиграть одну партию для первого игрока, т.к. он выигрывает в два раза чаще, $p = 2/3$, тогда $q = 1 - p = 1/3$. Матч состоит из трех повторяющихся событий. Вероятность выиграть две партии из трех $p_3(2) = C_3^2 (2/3)^2 (1/3)^{3-2} = 4/9$, а три партии из трех

$p_3(3) = C_3^3 (2/3)^3 (1/3)^{3-3} = 8/27$. Итак, $p(A) = p_3(2) + p_3(3) = \frac{4}{9} + \frac{8}{27} = \frac{20}{27} \approx 0{,}74$.

2) В нашем случае $n = 18$, $k = 13$, $p = 2/3$, $q = 1/3$. Вычислим $p_{18}(13)$, т.к. $n = 18$ большое число, то воспользуемся локальной формулой Лапласа (15)

$$p_{18}(13) = \frac{1}{\sqrt{18 \cdot \frac{2}{3} \cdot \frac{1}{3}}} \cdot \varphi(x), \text{ где } x = \frac{k - np}{\sqrt{npq}} = \frac{13 - 18 \cdot 2/3}{\sqrt{18 \cdot \frac{2}{3} \cdot \frac{1}{3}}} = \frac{1}{2}$$

и найдем из [1, прил. 1] $\varphi(0{,}5) \approx 0{,}3521$, тогда $p_{18}(13) = \frac{1}{\sqrt{4}} \cdot 0{,}3521 \approx 0{,}176$.

Ответ: вероятность выиграть матч из трех партий для первого игрока 0,74; $p_{18}(13) \approx 0{,}176$.

Пример. Игральная кость бросается 100 раз, найти вероятность того, что

1) шесть очков выпадут не менее 20 и не более 30 раз;
2) шесть очков выпадут не более 20 раз.

Решение.

1) Воспользуемся интегральной теоремой Лапласа. В нашем случае $n=100$, $k_1=20$, $k_2=30$, $p=1/6$, $q=1-p=1-1/6=5/6$. По формуле (16)

$$p_{100}(20,30)\approx \Phi(x_2)-\Phi(x_1),$$

где $x_1=\dfrac{k_1-np}{\sqrt{npq}}=\dfrac{20-100\cdot 1/6}{\sqrt{100\cdot 1/6\cdot 5/6}}=0{,}89$, $x_2=\dfrac{k_2-np}{\sqrt{npq}}=\dfrac{30-100\cdot 1/6}{\sqrt{100\cdot 1/6\cdot 5/6}}=3{,}58$.

Тогда $p_{100}(20,30)\approx \Phi(3{,}58)-\Phi(0{,}89)=0{,}4998-0{,}3133=0{,}1865$. Значение функции Лапласа $\Phi(3{,}58)$ и $\Phi(0{,}89)$ найдено из таблицы [1, прил. 2].

3) Воспользуемся интегральной теоремой Лапласа. В нашем случае $n=100$, $k_1=0$, $k_2=20$, $p=1/6$, $q=1-p=1-1/6=5/6$. По формуле (1.18)

$$x_1=\frac{k_1-np}{\sqrt{npq}}=\frac{0-100\cdot 1/6}{\sqrt{100\cdot 1/6\cdot 5/6}}=-4{,}47,\ x_2=\frac{k_2-np}{\sqrt{npq}}=\frac{20-100\cdot 1/6}{\sqrt{100\cdot 1/6\cdot 5/6}}=0{,}89.$$

Тогда

$p_{100}(0,20)\approx \Phi(0{,}89)-\Phi(-4{,}47)=\Phi(0{,}89)+\Phi(4{,}47)=0{,}3133+0{,}4999=0{,}8132$.

Ответ: $p_{100}(20,30)\approx 0{,}1865$, $p_{100}(0,20)\approx 0{,}8132$.

Пример. Каменщик за смену может уложить 1000 кирпичей. Вероятность того, что он уронит кирпич, равна 0,004. Найти вероятность того, что в течение смены произойдет падение 5 кирпичей.

Решение. Искомую вероятность найдем по формуле Пуассона (1.15).

По условию задачи $k=5$, $n=1000$, $\lambda=np=1000\cdot 0{,}004=4$. Подставим найденные значения в формулу и, воспользовавшись [1, табл. 3], получим

$$p_{1000}(5)\approx \frac{4^5e^{-4}}{5!}\approx 0{,}156.$$

Ответ: $p_{1000}(5)\approx 0{,}156$.

Пример. Каменщик за смену укладывает 1000 кирпичей. Вероятность того, что он уронит взятый кирпич, равна 0,004. Найти вероятность того, что в течение одной смены произойдет падение 5 кирпичей.

Решение. Искомую вероятность найдем по формуле Пуассона:

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

По условию задачи k = 5, λ = 1000 $\cdot$0,004= 4, тогда определяем $P_{1000}(5)\approx \dfrac{4^5e^{-4}}{5!}=0{,}156.$

Пример. Перегорела одна из трех лампочек, включенных в сеть последовательно. Наудачу выбранную лампочку заменяют годной, после чего проверяется исправность линии. Если повреждение не устранено, то заменяется другая лампочка. Найти вероятность того, что повреждение будет устранено после второй замены лампочки.

Решение. Пусть событие A – линия исправна после второй замены лампочки. Событие A является произведением события A_1 – линия неисправна после первой замены лампочки (т.е. заменена исправная лампочка) и события A_2 – линия исправна после второй замены лампочки (т.е. заменена неисправная лампочка): $A = A_1 A_2$. Поскольку события A_1 и A_2 зависимы, то $p(A_1 A_2) = p(A_1) p_{A_1}(A_2)$.

Вероятность заменить первый раз годную лампочку, которых две из трех, $p(A_1) = 2/3$. Вероятность во второй раз заменить перегоревшую лампочку, которая одна из двух оставшихся непроверенных, $p_{A_1}(A_2) = 1/2$, тогда $p(A) = \frac{2}{3} \cdot \frac{1}{2} \approx 0{,}33$.

Ответ: $p(A) \approx 0{,}33$.

Замечание. Для эффективного решения задач при переводе словесного описания событий на язык алгебры событий полезно помнить, что союзу “и” соответствует произведение событий AB, союзу “или” – сумма событий $A + B$, частице “не” – противоположное событие $\overline{A}$.

Пример. По мишени стреляют три стрелка. Вероятность попадания в мишень первым стрелком – 0,9, вторым – 0,8 и третьим – 0,5. Найти вероятность следующих событий:

1) B_1 - все три стрелка попали в мишень;
2) B_2 - никто не попал в мишень;
3) B_3 - хотя бы один стрелок попал в мишень;
4) B_4 - только один стрелок попал в мишень;
5) B_5 - только два стрелка попали в мишень.

Решение. Очевидно, каждое из событий B_k состоит из нескольких событий, перечислим их:

A_1 - первый стрелок попал в цель, $p(A_1) = 0{,}9$, $p(\overline{A}_1) = 1 - p(A_1) = 1 - 0{,}9 = 0{,}1$;

A_2 - второй стрелок попал в цель, $p(A_2) = 0{,}8$, $p(\overline{A}_2) = 1 - p(A_2) = 1 - 0{,}8 = 0{,}2$;

A_3 - третий стрелок попал в цель, $p(A_3) = 0{,}5$, $p(\overline{A}_3) = 1 - p(A_3) = 1 - 0{,}5 = 0{,}5$.

1) Событие B_1 состоит в том, что и первый стрелок попал в цель, и второй попал, и третий попал, следовательно, $B_1 = A_1 A_2 A_3$. Так как события A_1, A_2 и A_3 независимы, то по формуле, аналогичной (1.10),

$$p(B_1) = p(A_1 A_2 A_3) = p(A_1) p(A_2) p(A_3) = 0{,}9 \cdot 0{,}8 \cdot 0{,}5 = 0{,}36.$$

2) Событие B_2 состоит в том, что и первый стрелок не попал в цель, и второй не попал, и третий не попал, следовательно, $B_2 = \overline{A}_1\overline{A}_2\overline{A}_3$. Так как события $\overline{A}_1$, $\overline{A}_2$ и $\overline{A}_3$ независимы, то по формуле, аналогичной (1.10),

$$p(B_2) = p(\overline{A}_1\overline{A}_2\overline{A}_3) = p(\overline{A}_1)p(\overline{A}_2)p(\overline{A}_3) = 0{,}1 \cdot 0{,}2 \cdot 0{,}5 = 0{,}01.$$

3) Событие B_3 складывается из большого количества вариантов – или $A_1\,A_2\,A_3$, или $\overline{A}_1\,A_2\,A_3$, или $A_1\,\overline{A}_2\,A_3$ и т.д. Поэтому выгодно перейти к противоположному событию $\overline{B}_3$ – никто не попал, в наших обозначениях $\overline{B}_3 = B_2$:
$p(B_3) = 1 - p(\overline{B}_3) = 1 - p(B_2) = 1 - 0{,}01 = 0{,}99$.

4) Событие B_4 – только одно попадание – складывается из следующих вариантов: первый стрелок попал, а второй и третий не попали, или второй стрелок попал, а первый и третий не попали, или третий стрелок попал, а первый и третий не попали. Следовательно, $B_4 = A_1\overline{A}_2\overline{A}_3 + \overline{A}_1A_2\overline{A}_3 + \overline{A}_1\overline{A}_2A_3$. Отметим, что одновременно события $A_1\overline{A}_2\overline{A}_3$, $\overline{A}_1A_2\overline{A}_3$ и $\overline{A}_1\overline{A}_2A_3$ произойти не могут, т.е. являются несовместными. Поэтому

$$\begin{aligned} p(B_4) &= p(A_1\overline{A}_2\overline{A}_3) + p(\overline{A}_1A_2\overline{A}_3) + p(\overline{A}_1\overline{A}_2A_3) = \\ &= p(A_1)p(\overline{A}_2)p(\overline{A}_3) + p(\overline{A}_1)p(A_2)p(\overline{A}_3) + p(\overline{A}_1)p(\overline{A}_2)p(A_3) = \\ &= 0{,}9 \cdot 0{,}2 \cdot 0{,}5 + 0{,}1 \cdot 0{,}8 \cdot 0{,}5 + 0{,}1 \cdot 0{,}2 \cdot 0{,}5 = 0{,}14. \end{aligned}$$

5) Событие B_5 – только два попадания – как и в предыдущем пункте, состоит в переборе трех вариантов, когда один из стрелков не попал в мишень, а остальные двое попали: $B_5 = \overline{A}_1A_2A_3 + A_1\overline{A}_2A_3 + A_1A_2\overline{A}_3$;

$$\begin{aligned} p(B_5) &= p(\overline{A}_1A_2A_3) + p(A_1\overline{A}_2A_3) + p(A_1A_2\overline{A}_3) = \\ &= p(\overline{A}_1)p(A_2)p(A_3) + p(A_1)p(\overline{A}_2)p(A_3) + p(A_1)p(A_2)p(\overline{A}_3) = \\ &= 0{,}1 \cdot 0{,}8 \cdot 0{,}5 + 0{,}9 \cdot 0{,}2 \cdot 0{,}5 + 0{,}9 \cdot 0{,}8 \cdot 0{,}5 = 0{,}49. \end{aligned}$$

Ответ: $p(B_1) = 0{,}36$, $p(B_2) = 0{,}01$, $p(B_3) = 0{,}99$, $p(B_4) = 0{,}14$, $p(B_5) = 0{,}49$.

Замечание. Отметим, что сумма событий $B_1 + B_2 + B_4 + B_5$ образует достоверное событие D, т.к. B_1, B_2, B_4 и B_5 попарно несовместны и являются полной группой исходов, то должно быть выполнено условие $p(D) = p(B_1) + p(B_2) + p(B_4) + p(B_5) = 1$. Это условие можно использовать как критерий правильности вычислений: $0{,}36 + 0{,}01 + 0{,}14 + 0{,}49 = 1$. Действительно верно.

Пример. В магазине имеются электрические лампочки, изготовленные на трех заводах: 3 ящика с первого завода, 5 – со второго завода и 2 – с третьего. Процент бракованных лампочек (то есть тех, что перегорают раньше положенного срока) составляет на первом заводе – 10%, на втором – 20%, на

третьем – 25%. Найти вероятность купить бракованную лампочку при условии, что ящики заносятся в торговый зал случайным образом.

Решение. Рассмотрим события: A_1 – лампочка сделана на первом заводе, A_2 – лампочка сделана на втором заводе, A_3 – на третьем заводе. Так как в магазине всего $3+5+2=10$ ящиков лампочек, то вероятность что лампочка изготовлена первым заводом $p(A_1)=3/10$, соответственно, $p(A_2)=0{,}5$, $p(A_3)=0{,}2$.

Если лампочка произведена на первом заводе, то вероятность события B – деталь бракованная – $p_{A_1}(B)=0{,}1$, так как на первом заводе брак составляет 10%. Аналогично, $p_{A_2}(B)=0{,}2$, $p_{A_3}(B)=0{,}25$.

По формуле полной вероятности (1.13) при $n=3$:

$$p(B)=p(A_1)p_{A_1}(B)+p(A_2)p_{A_2}(B)+p(A_3)p_{A_3}(B)=$$
$$=0{,}3\cdot 0{,}1+0{,}5\cdot 0{,}2+0{,}2\cdot 0{,}25=0{,}18.$$

Ответ: $p(B)=0{,}18$.

2. Дискретная случайная величина

Числовую величину, которая в результате опыта может принять одно из своих значений, заранее не известно какое, называют *случайной величиной*.

Дискретной называют случайную величину, возможные значения которой есть отдельные изолированные числа.

Законом распределения дискретной случайной величины называют перечень ее возможных значений и соответствующих им вероятностей:

X	x_1	x_2	$\dots$	x_3	$\dots$
P	p_1	p_2	$\dots$	p_3	$\dots$

Причем всегда

$$\sum_{i=1}^{n} p_i = 1. \qquad (2.1)$$

Число возможных значений дискретной случайной величины может быть конечным или бесконечным.

Биноминальным называют закон распределения дискретной случайной величины $X_{n\,p}$, равной числу появлений события A в n независимых испытаниях, в каждом из которых вероятность появления события A равна p.

Ясно, что вероятность того, что X примет значение k (события A появится k раз) вычисляют по формуле Бернулли:

$$P(k)=C_n^k p^k q^{n-k}. \qquad (2.2)$$

Математическим ожиданием дискретной случайной величины называют сумму произведений значений случайной величины на их вероятности

$$M(x)=\sum_{i=1}^{\infty} x_1 p_1 = x_1 p_1 + x_2 p_2 + \ldots + x_n p_n + \ldots . \qquad (2.3)$$

Дисперсией случайной величины называют математическое ожидание квадрата отклонения случайной величины от ее среднего значения

$$D(X) = (M(X - M(X)))^2. \qquad (2.4)$$

Дисперсию случайной величины удобно вычислять по формуле

$$D(X) = M(X^2) - (M(X))^2 = \sum (x_1)^2 p_1 - (M(X))^2. \qquad (2.5)$$

Средним квадратическим отклонением случайной величины называют квадратный корень из дисперсии:

$$\sigma(X) = \sqrt{D(X)}. \qquad (2.6)$$

Математическое ожидание и дисперсия *биноминального закона,* соответственно, равны

$$M(X) = np \quad \text{и} \quad D(X) = npq. \qquad (2.7)$$

Функцией распределения случайной величины X называют функцию $F(x)$, которая для каждого значения x равна вероятности того, что случайная величина X примет значение, меньшее x

$$F(x) = P(X < x). \qquad (2.8)$$

Функция распределения обладает следующими свойствами:

1) $0 \le F(x) \le 1$;

2) $F(x_2) \ge F(x_1)$, если $x_2 > x_1$;

3) $\lim\limits_{x \to -\infty} F(x) = 0, \quad \lim\limits_{x \to \infty} F(x) = 1.$

Доказано, что вероятность того, что случайная величина X примет значение, заключенное в интервале $[\alpha, \beta]$, равна приращению функции распределения на этом интервале

$$P(\alpha \le X < \beta) = F(\beta) - F(\alpha). \qquad (2.9)$$

<u>Пример.</u> По мишени стреляют три стрелка. Вероятность попадания в мишень первым стрелком равна 0,9; вторым – 0,8; третьим – 0,5. Для случайной величины X, равной числу промахов, требуется:

1) составить закон распределения;
2) построить многоугольник распределения;
3) построить функцию распределения $F(x)$;
4) вычислить числовые характеристики – математическое ожидание, дисперсию и среднее квадратическое отклонение.

<u>Решение.</u> Случайная величина X – число промахов, может принимать четыре значения: $X = 0$, если все попали в цель; $X = 1$, если кто-то один из трех промахнулся и т.д.

1) Вычисление вероятности для каждого значения случайной величины было проведено в примере (см. выше): $p(X = 0) = p(B_1) = 0{,}36$; $p(X = 1) = p(B_5) = 0{,}49$; $p(X = 2) = p(B_4) = 0{,}14$; $p(X = 3) = p(B_2) = 0{,}01$.

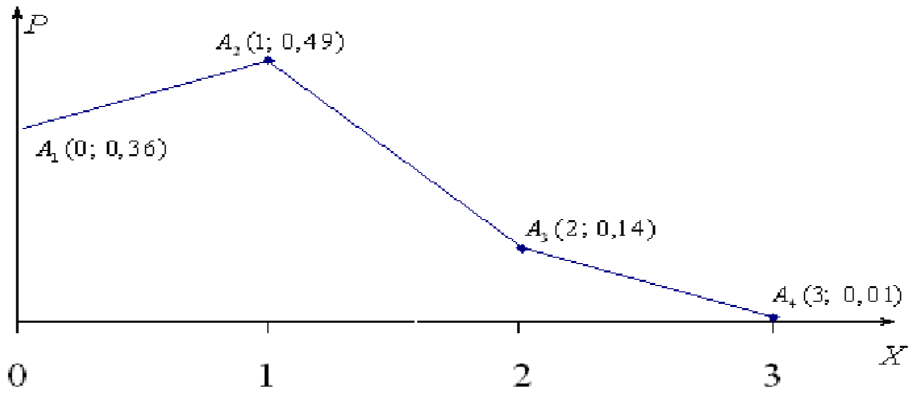
Составим закон распределения случайной величины X:

X	0	1	2	3
p	0,36	0,49	0,14	0,01

Проверка: 0,36+0,49+0,14+0,01=1.

Следовательно, мы нашли закон распределения данной случайной величины.

3) Чтобы построить многоугольник распределения, укажем на плоскости XOP 4 точки: $A_1(0;0{,}36)$, $A_2(1;0{,}49)$, $A_3(2;0{,}14)$, $A_4(3;0{,}01)$ и соединим их ломаной.



3) Для построения функции распределения $F(x)$ отметим, что ее значения могут меняться только при переходе через точки $X=0,1,2,3$, точнее, должны увеличиваться на соответствующие вероятности.

Например, $F(x=0)=0$, т.к. левее нет значений случайной величины.

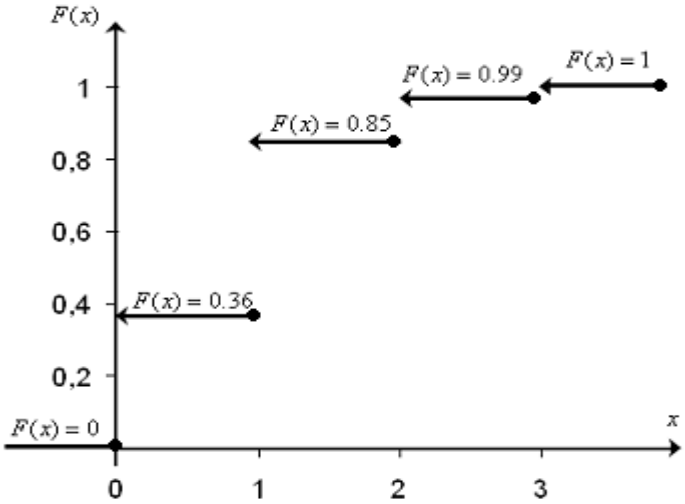
Если $0<X\le 1$, то для любого x их этого полуинтервала
$F(x)=p(X<x)=p(X=0)$=0,36, т.к. левее $X=x$ только одно значение $X=0$.

Если $1<X\le 2$, то для любого x их этого полуинтервала
$F(x)=p(X<x)=p(X=0)+p(X=1)$=0,36+0,49=0,85.

Если $2<X\le 3$, то для любого x их этого полуинтервала
$F(x)=p(X<x)=p(X=0)+p(X=1)+p(X=2)$=0,36+0,49+0,14=0,99.

Если $3<X$, то для любого x их этого бесконечного промежутка
$F(x)=p(X=0)+p(X=1)+p(X=2)+p(X=3)$=0,36+0,49+0,14+0,01=1.

Итак, функция $F(x)$ является неубывающей, непрерывной слева функцией, а ее график поднимается “лесенкой”:



4) Вычислим числовые характеристики:

- математическое ожидание дискретной случайной величины $M(x)=\sum_{k=1}^{n} x_i p_i = 0\cdot 0{,}36+1\cdot 0{,}49+2\cdot 0{,}14+3\cdot 0{,}01=0{,}8$.
- дисперсия дискретной случайной величины $D(x)=M(X^2)-\left(M(X)\right)^2$,

где $M(X^2)=\sum_{k=1}^{n} x_i^2 p_i = 0^2\cdot 0{,}36+1^2\cdot 0{,}49+2^2\cdot 0{,}14+3^2\cdot 0{,}01=1{,}14$, тогда $D(x)=1{,}14-\left(0{,}8\right)^2=0{,}5$.

- среднее квадратическое отклонение $\sigma(x)=\sqrt{D(x)}=\sqrt{0{,}5}\approx 0{,}71$.

<u>Пример.</u> Каждая из трех опор автомобильного моста через реку может быть повреждена во время ледохода. Вероятности повреждения опор независимы и соответственно равны: 0,6, 0,7, и 0,8. Составить закон распределения случайной величины X – числа поврежденных опор. Найти ее математическое ожидание, дисперсию и среднее квадратическое отклонение.

Решение. Дискретная случайная величина X – число поврежденных опор моста – имеет следующие возможные значения:

$$X_1=0,\ X_2=1,\ X_3=2,\ X_4=3.$$

Найдем вероятности этих возможных значений. Для этого рассмотрим следующие события:

A – первая опора повреждена;

B – вторая опора повреждена;

C – третья опора повреждена;

По условию задачи

$$P(A)=0{,}6,\quad P(B)=0{,}7,\quad P(C)=0{,}8.$$

Вероятности противоположных событий соответственно равны:

$$P(\overline{A})=1-P(A)=0{,}4;\quad P(\overline{B})=1-P(B)=0{,}3;\quad P(\overline{C})=1-P(C)=0{,}2.$$

Учитывая независимость событий A, B и C, получим:

$$P_1=P(X=0)=P(\overline{A}\cdot\overline{B}\cdot\overline{C})=0{,}4\cdot 0{,}3\cdot 0{,}2=0{,}024;$$

$$P_2=P(X=1)=P(A\cdot\overline{B}\cdot\overline{C}+\overline{A}\cdot B\cdot\overline{C}+\overline{A}\cdot\overline{B}\cdot C)=0{,}6\cdot 0{,}3\cdot 0{,}2+0{,}4\cdot 0{,}7\cdot 0{,}2+$$
$$+0{,}4\cdot 0{,}3\cdot 0{,}8=0{,}118;$$

$$P_3=P(X=2)=P(A\cdot B\cdot\overline{C}+A\cdot\overline{B}\cdot C+\overline{A}\cdot B\cdot C)=0{,}453;$$

$$P_4=P(X=3)=P(A\cdot B\cdot C)=0{,}6\cdot 0{,}7\cdot 0{,}8=0{,}336.$$

При этом $P_1+P_2+P_3+P_4=0{,}024+0{,}118+0{,}425+0{,}336=1$.

Окончательно, искомый закон распределения имеет вид

X	0	1	2	3
P	0,024	0,188	0,452	0,336

Найдем параметры распределения:

$$M(X) = 0 \cdot 0{,}024 + 1 \cdot 0{,}188 + 2 \cdot 0{,}453 + 3 \cdot 0{,}366 = 2{,}1;$$

$$D(X) = 0 \cdot 0{,}024 + 1 \cdot 0{,}188 + 4 \cdot 0{,}452 + 9 \cdot 0{,}366 - 2{,}1^2 = 0{,}61;$$

$$\sigma(X) = \sqrt{0{,}61} = 0{,}78.$$

3. Непрерывная случайная величина

Непрерывной называют случайную величину, которая может принимать все значения из некоторого конечного или бесконечного промежутка и при этом ее функция распределения дифференцируема.

Плотностью распределения *f(x)* случайной величины *X* называют производную от функции распределения этой случайной величины

$$f(x) = F'(x). \tag{3.1}$$

Функция плотности распределения $f(x)$ обладает следующими свойствами:

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

2) $\int\limits_{-\infty}^{\infty} f(x)dx = 1.$

В частности, если все возможные значения случайной величины *X* принадлежат интервалу (*a*, *b*), то

$$\int\limits_{a}^{b} f(x)dx = 1. \tag{3.2}$$

Вероятность того, что непрерывная случайная величина *X* в результате опыта примет значение, принадлежащее интервалу (α, β), определяется равенством

$$P(\alpha < X < \beta) = \int\limits_{a}^{b} f(x)dx. \tag{3.3}$$

Зная плотность распределения $f(x)$, можно найти функцию распределения *F(x)* по формуле

$$F(x) = \int\limits_{-\infty}^{x} f(t)dt. \tag{3.4}$$

Математическое ожидание непрерывной случайной величины X вычисляют по формуле

$$M(X) = \int\limits_{-\infty}^{\infty} xf(x)dx, \tag{3.5}$$

где $f(x)$ – плотность распределения.

Дисперсию непрерывной случайной величины X вычисляют по формуле

$$D(X) = M(X^2) - M^2(X), \tag{3.6}$$

где $M(X^2) = \int_{-\infty}^{\infty} x^2 f(x)dx.$

Среднее квадратическое отклонение случайной величины *X* есть квадратный корень из дисперсии

$$\sigma(X) = \sqrt{D(X)}.$$

Случайная величина *распределена равномерно* на интервале *(a, b),* если ее плотность распределения сохраняет на этом интервале постоянное значение, равное

$$f(x) = \frac{1}{b-a} \tag{3.7}$$

Нормальным называют распределение вероятностей непрерывной случайной величины *X,* плотность которой имеет вид

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}, \tag{3.8}$$

при этом *a* – математическое ожидание, а σ - среднее квадратическое отклонение случайной величины *X*, распределенной по нормальному закону.

Вероятность того, что нормально распределенная случайная величина примет значение, принадлежащее интервалу (α, β), можно вычислить по формуле

$$P(\alpha < X < \beta) = \Phi\left(\frac{\beta - a}{\sigma}\right) - \Phi\left(\frac{\alpha - a}{\sigma}\right), \tag{3.9}$$

где *Ф(x)* – функция Лапласа, таблица значений которой приведена в [4, 5].

Вероятность того, что абсолютная величина отклонения случайной величины *X,* распределенной по нормальному закону, от ее математического ожидания будет меньше положительного числа δ, равна

$$P(|X - a| < \delta) = 2\Phi\left(\frac{\delta}{\sigma}\right). \tag{3.10}$$

Показательным называют закон распределения случайной величины, функция плотности распределения которой имеет вид

$$f(x) = \begin{cases} 0 \text{ при } x < 0, \\ \lambda e^{-\lambda x} \text{ при } x \geq 0, \end{cases} \tag{3.11}$$

где λ - положительная константа.

<u>Пример</u>. Дана функция распределения *F(X)* случайной величины. Найти плотность распределения $f(x)$. Требуется найти математическое ожидание, дисперсию и среднее квадратическое отклонение случайной величины *X.*

$$F(x)=\begin{cases} 0, & x\le 0 \\ 1-\cos x, & 0<x\le \pi/2 \\ 1, & x>\pi/2 \end{cases}.$$

Решение. Для нахождения плотности распределения $f(x)$ воспользуемся формулой $f(x)=F'(x)$. Тогда,

$$f(x)=F'(x)=\begin{cases} 0, & x\le 0 \\ \sin x, & 0<x\le \pi/2 \\ 0, & x>\pi/2 \end{cases}.$$

Найдем числовые характеристики случайной величины *X:*

$$M(x)=\int_{-\infty}^{\infty} x\,f(x)dx=\int_0^{\pi/2} x\cos x dx=\begin{pmatrix} u=x & d\nu=\sin x dx \\ du=dx & \nu=-\cos x \end{pmatrix}=$$

$$=-x\cos x\Big|_0^{\pi/2}+\int_0^{\pi/2}\cos x dx=0+\sin x\Big|_0^{\pi/2}=1.$$

$$D(x)=\int_0^{\pi/2} x^2\sin x dx-1^2=\begin{pmatrix} u=x^2 & d\nu=\sin x dx \\ du=2x dx & \nu=-\cos x \end{pmatrix}=$$

$$=-x^2\cos x\Big|_0^{\pi/2}+2\int_0^{\pi/2} x\cos x dx-1=2\int_0^{\pi/2} x\cos x dx-1=$$

$$=\begin{pmatrix} u=x & d\nu=\cos x dx \\ du=dx & \nu=\sin x \end{pmatrix}=-1+2x\sin x\Big|_0^{\pi/2}-2\int_0^{\pi/2}\sin x dx=-1+\pi-2=\pi-3$$

$$\sigma(x)=\sqrt{D(x)}=\sqrt{\pi-3}.$$

<u>Пример.</u> Задана функция распределения непрерывной случайной величины

$$F(x)=\begin{cases} 0, & x\le 0, \\ Ax^4, & 0<x\le 2, \\ 1, & 2<x. \end{cases}$$

Требуется: 1) найти параметр A;

2) построить функцию распределения $F(x)$;

3) найти функцию плотности $f(x)$ и построить ее график;

4) вычислить числовые характеристики – математическое ожидание, дисперсию и среднее квадратическое отклонение;

5) найти вероятность попадания случайной величины в интервал $(0;1)$.

Решение.

1) В силу непрерывности функции распределения $F(x)$ (слева она, очевидно, непрерывна) она должна быть непрерывна и справа при $x=2$: $Ax^4\big|_{x=2}=1$, тогда $16A=1$ и $A=1/16$.

2) Поскольку $A=1/16$, то функция распределения имеет вид

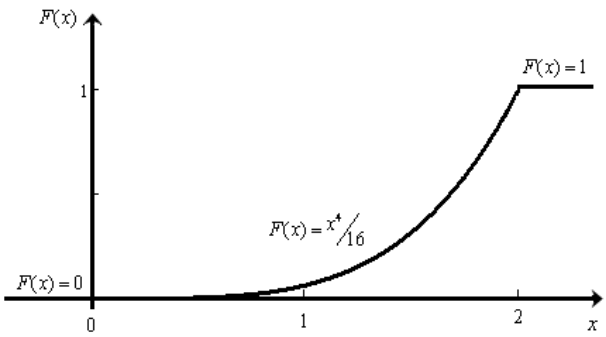
$$F(x)=\begin{cases} 0, & x\le 0,\\ x^4/16, & 0<x\le 2,\\ 1, & 2<x.\end{cases}$$

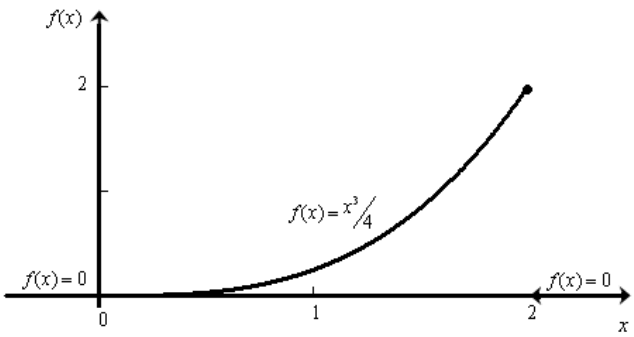
График этой функции укажем ниже.

3) Дифференцируя $F(x)$, получим функцию плотности распределения $f(x)$:

$$f(x)=\begin{cases} 0, & x\le 0,\\ x^3/4, & 0<x\le 2,\\ 0, & 2<x.\end{cases}$$

Изобразим графики функций $F(x)$ и $f(x)$:





4) Вычислим числовые характеристики по формулам, указанным в табл. 1: математическое ожидание

$$M(x)=\int_{-\infty}^{\infty} xf(x)dx=\int_{-\infty}^{0} x\cdot 0dx+\int_{0}^{2} x\frac{x^3}{4}dx+\int_{2}^{\infty} x\cdot 0dx=\int_{0}^{2}\frac{x^4}{4}dx=\frac{x^5}{20}\Big|_0^2=\frac{8}{5};$$

дисперсия $D(x)=M(X^2)-\left(M(X)\right)^2$, где

$$M(X^2)=\int_{-\infty}^{\infty} x^2 f(x)dx=\int_{-\infty}^{0} x^2\cdot 0dx+\int_{0}^{2} x^2\frac{x^3}{4}dx+\int_{2}^{\infty} x^2\cdot 0dx=\int_{0}^{2}\frac{x^5}{4}dx=\frac{x^6}{24}\Big|_0^2=\frac{64}{24}=\frac{8}{3},$$

тогда $D(x)=\frac{8}{3}-\left(\frac{8}{5}\right)^2=\frac{8}{75}$,

среднее квадратическое отклонение $\sigma(x)=\sqrt{D(x)}=\sqrt{\frac{8}{75}}\approx 0{,}33$.

5) Вероятность попадания случайной величины в интервал $(0;1)$ определяется формулой $p(0<X<1)=\int_0^1 f(x)dx=\int_0^1\frac{x^3}{4}dx=\frac{x^4}{16}\Big|_0^1=\frac{1}{16}$.

Ответ: $A=1/16$, $M(x)=8/5$, $D(x)=8/75$, $\sigma(x)\approx 0{,}33$, $p(0<X<1)=1/16$.

<u>Пример.</u> Дана функция плотности распределения $f(x)$ случайной величины X:

$$f(x)=\begin{cases} 0, & x\le 1,\\ (x-1)/2, & 1<x\le 3,\\ 0, & 3<x.\end{cases}$$

Требуется: 1) найти функцию распределения $F(x)$;

2) построить графики функций плотности $f(x)$ и распределения $F(x)$;

3) найти вероятность того, что случайная величина X примет значение, принадлежащее интервалу $(2;4)$.

<u>Решение.</u>

1) Найдем функцию распределения $F(x)$:

если $x\le 1$, то $F(x)=\int\limits_{-\infty}^{x} f(z)dz=\int\limits_{-\infty}^{x} 0dz=0$;

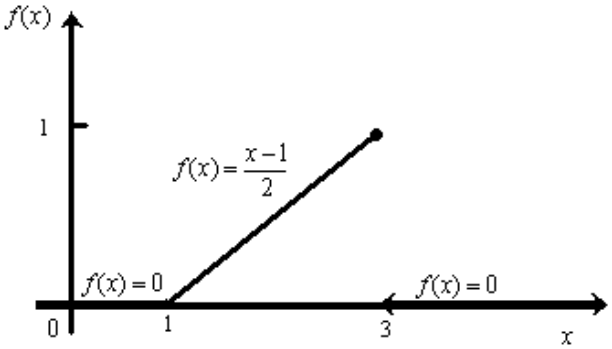
если $1<x\le 3$, то $F(x)=\int\limits_{-\infty}^{x} f(z)dz=\int\limits_{-\infty}^{1} 0dz+\int\limits_{1}^{x}\frac{z-1}{2}dz=0+\int\limits_{1}^{x}\left(\frac{z}{2}-\frac{1}{2}\right)dz=\left(\frac{z^2}{4}-\frac{z}{2}\right)\Bigg|_{1}^{x}=\frac{x^2}{4}-\frac{x}{2}+\frac{1}{4}$;

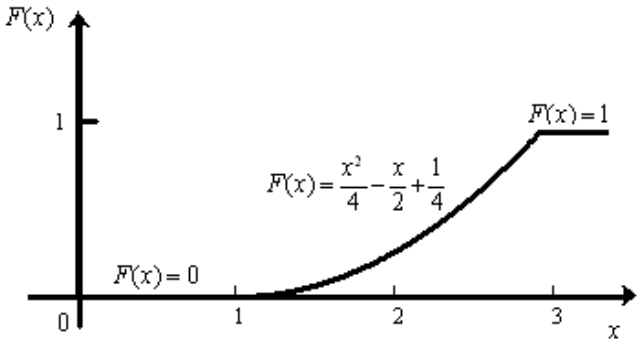
если $3<x$, то $F(x)=\int\limits_{-\infty}^{x} f(z)dz=\int\limits_{-\infty}^{1} 0dz+\int\limits_{1}^{3}\frac{z-1}{2}dz+\int\limits_{3}^{+\infty} 0dz=0+\int\limits_{1}^{3}\left(\frac{z}{2}-\frac{1}{2}\right)dz=\left(\frac{z^2}{4}-\frac{z}{2}\right)\Bigg|_{1}^{3}=1$.

Итак, искомая функция распределения имеет вид

$$F(x)=\begin{cases} 0, & x\le 1,\\ \frac{x^2}{4}-\frac{x}{2}+\frac{1}{4}, & 1<x\le 3,\\ 1, & 3<x.\end{cases}$$

2) Графики функций $f(x)$ и $F(x)$:





3) Вероятность того, что случайная величина X примет значение, принадлежащее интервалу $(2;4)$:

$$p(2<X<4)=\int\limits_{2}^{4} f(x)dx=\int\limits_{2}^{3}\frac{x-1}{2}dx+\int\limits_{3}^{4} 0dx=\frac{(x-1)^2}{4}\Bigg|_{2}^{3}=\frac{(3-1)^2}{4}-\frac{(2-1)^2}{4}=\frac{3}{4}$$

или по формуле (21):

$$p(2<X<4)=F(4)-F(2)=1\Big|_{x=4}-\left(\frac{x^2}{4}-\frac{x}{2}+\frac{1}{4}\right)\Bigg|_{x=2}=1-\left(\frac{4}{4}-\frac{2}{2}+\frac{1}{4}\right)=\frac{3}{4}.$$

<u>Ответ:</u> $F(x)=\begin{cases} 0, & x\le 1 \\ \dfrac{x^2}{4}-\dfrac{x}{2}+\dfrac{1}{4}, & 1<x\le 3, \\ 1, & 3<x \end{cases}$ $p(2<X<4)=\dfrac{3}{4}$.

<u>Пример.</u> Случайная величина X распределена равномерно на интервале $(3;12)$. Требуется:

1) найти функцию плотности $f(x)$, функцию распределения $F(x)$ и построить их графики;
2) вычислить числовые характеристики – математическое ожидание, дисперсию и среднее квадратическое отклонение;
3) найти вероятность попадания случайной величины в интервал $(1;8)$.

<u>Решение.</u>

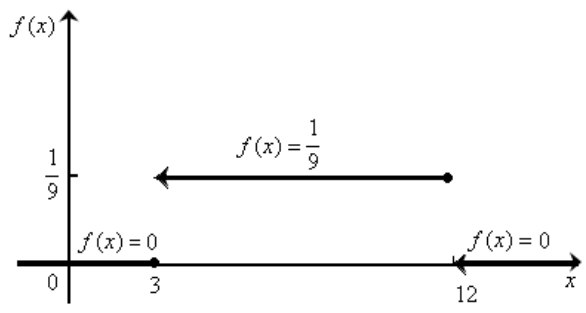
1) Функция плотности распределения случайной величины X на промежутке $(3;12]$ равна $f(x)=1/(12-3)=1/9$, вне данного промежутка $f(x)=0$: $f(x)=\begin{cases} 1/9, & x\in(3;12] \\ 0, & x\notin(3;12] \end{cases}$.

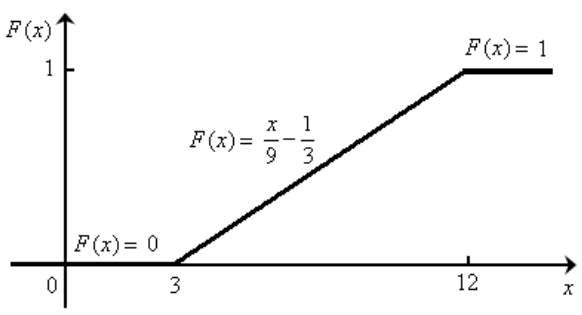
Найдем функцию распределения $F(x)$: если $x\le 3$, то $F(x)=\int\limits_{-\infty}^{x} f(z)dz=\int\limits_{-\infty}^{x} 0dz=0$;

если $3<x\le 12$, то $F(x)=\int\limits_{-\infty}^{x} f(z)dz=\int\limits_{-\infty}^{3} 0dz+\int\limits_{3}^{x}\dfrac{1}{9}dz=0+\left(\dfrac{z}{9}\right)\Bigg|_{3}^{x}=\dfrac{x}{9}-\dfrac{1}{3}$;

если $12<x$, то $F(x)=\int\limits_{-\infty}^{x} f(z)dz=\int\limits_{-\infty}^{3} 0dz+\int\limits_{3}^{12}\dfrac{1}{9}dz+\int\limits_{12}^{+\infty} 0dz=1$.

Таким образом, $F(x)=\begin{cases} 0, & x\le 3, \\ \dfrac{x}{9}-\dfrac{1}{3}, & 3<x\le 12, \\ 1, & 12<x. \end{cases}$





2) Найдем числовые характеристики: математическое ожидание

$$M(x)=\int\limits_{-\infty}^{\infty} xf(x)dx=\int\limits_{-\infty}^{3} x\cdot 0dx+\int\limits_{3}^{12} x\frac{1}{9}dx+\int\limits_{12}^{\infty} x\cdot 0dx=\int\limits_{3}^{12} x\frac{1}{9}dx=\frac{x^2}{18}\Bigg|_{3}^{12}=\frac{135}{18}=7{,}5;$$

дисперсия $D(x)=M(X^2)-(M(X))^2$, где $M(X^2)=\int\limits_{-\infty}^{\infty}x^2 f(x)dx=\int\limits_{-\infty}^{3}x^2\cdot 0dx+$

$+\int\limits_{3}^{12}x^2\frac{1}{9}dx+\int\limits_{12}^{\infty}x^2\cdot 0dx=\int\limits_{3}^{12}\frac{x^2}{9}dx=\left.\frac{x^3}{27}\right|_3^{12}=63$, тогда $D(x)=63-(7{,}5)^2\approx 6{,}75$;

среднее квадратическое отклонение $\sigma(x)=\sqrt{D(x)}=\sqrt{6{,}75}\approx 2{,}6$.

Вычислить числовые характеристики в данном случае можно было по формулам: $M(X)=\frac{a+b}{2}=\frac{3+12}{2}=7{,}5$, $D(X)=\frac{(b-a)^2}{12}=\frac{(12-3)^2}{12}=\frac{81}{12}=6{,}75$, $\sigma(X)=\frac{b-a}{2\sqrt{3}}=\frac{12-3}{2\sqrt{3}}\approx 2{,}6$.

3) Вероятность того, что случайная величина X примет значение, принадлежащее интервалу $(1;8)$: $p(1<X<8)=\int\limits_{1}^{8}f(x)dx=\int\limits_{1}^{3}0dx+\int\limits_{3}^{8}\frac{1}{9}dx=0+\left.\frac{x}{9}\right|_3^8=\frac{5}{9}$

или по формуле (21): $p(1<X<8)=F(8)-F(1)=\left.\left(\frac{x}{9}-\frac{1}{3}\right)\right|_{x=8}=\frac{5}{9}$.

<u>Ответ:</u> $f(x)=\begin{cases}1/9, & x\in(3;12]\\ 0, & x\notin(3;12]\end{cases}$, $F(x)=\begin{cases}0, & x\le 3\\ \frac{x}{9}-\frac{1}{3}, & 3<x\le 12\\ 1, & 12<x\end{cases}$, $M(x)=7{,}5$, $D(x)=6{,}75$, $\sigma(x)\approx 2{,}6$, $p(1<X<8)=5/9$.

<u>Пример.</u> Диаметр деталей, вытачиваемых на станке, является случайной величиной, распределенной по нормальному закону с параметрами $a=20$ *мм* и $\sigma=2$ *мм*. Деталь считается стандартной, если ее диаметр d лежит в диапазоне $19<d<22$. Найти:

1) вероятность того, что деталь будет стандартной;
2) вероятность того, что отклонение диаметра от ожидаемого (от математического ожидания $a=20$ *мм*) не превышает 1 *мм*;
3) написать функцию плотности распределения и нарисовать ее график.

<u>Решение.</u> 1) Вероятность того, что деталь будет стандартной, где $a=20$, $\sigma=2$, $\alpha=19$ и $\beta=22$ определяется (3.9):

$$p(19<X<22)=\Phi\left(\frac{22-20}{2}\right)-\Phi\left(\frac{19-20}{2}\right)=\Phi(1)-\Phi(-0{,}5)=$$
$$=\Phi(1)+\Phi(0{,}5)=0{,}3413+0{,}1915=0{,}5328.$$

Мы воспользовались тем, что функция $\Phi(x)$ нечетна, а значения $\Phi(1)$ и $\Phi(0{,}5)$ нашли по таблице значений функции Лапласа [1, прил. 2].

<u>Пример</u> **.** Математическое ожидание нормально распределенной случайной величины X равно 6, а среднее квадратическое отклонение равно 2. Требуется найти: а) вероятность того, что X примет значение, принадлежащее интервалу (3; 10); б) вероятность того, что абсолютная величина отклонения X от математического ожидания окажется меньше 5.

Решение. а) Воспользуемся формулой (3.9). В нашем случае, $a = 6$, $\sigma = 2$, $\alpha = 3$, $\beta = 10$.

$$P(\alpha < X < \beta) = \Phi\left(\frac{10-6}{2}\right) - \Phi\left(\frac{3-6}{2}\right) = \Phi(2) - \Phi(-1{,}5) = \Phi(2) + \Phi(1{,}5) =$$

$$= 0{,}4772 + 0{,}4332 = 0{,}9104.$$

Значения $\Phi(2)$ и $\Phi(1{,}5)$ определены по таблице [1, прил.2].

б) Воспользуемся формулой (3.10). По условию задачи имеем

$$a = 6, \quad \delta = 5, \quad \sigma = 2.$$

$$P(|X - 6| < 5) = 2\Phi\left(\frac{5}{2}\right) = 2\Phi(2{,}5) = 2 \cdot 0{,}4938 = 0{,}9876.$$

II. Элементы математической статистики

Рассмотрим некоторое социальное явление, мнение о котором можно составить по опросу населения. Пусть дано большое множество объектов (листов опроса населения) – данное множество объектов называется **генеральной совокупностью**. Нет возможности изначально перечислить все ответы в листах-опросах, т.е. составить полную группу исходов, но возможно одинаковым ответам присвоить одинаковые номера, таким образом, перевести все ответы на язык цифр. Далее необходимо работать с генеральной совокупностью как со случайной величиной, принимающей численные значения, которые называют **вариантами**. Поскольку листов опроса очень большое количество, берется сравнительно небольшая часть генеральной совокупности. Часть генеральной совокупности называется **выборкой** (количество элементов в ней есть **объем выборки** n), по которой делается вывод об всей генеральной совокупности: по какому закону она распределена, каковы числовые характеристики распределения. При этом важно, чтобы полученные результаты обладали достаточной достоверностью.

Обычно статистическое распределение задается с помощью таблицы:

X	x_1	x_2	…	x_i	…	x_k
n	n_1	n_2	…	n_i	…	n_k

. (2.1)

Частоты n_i показывают, сколько раз встретилось значение варианты x_i:

$$\sum_{i=1}^{k} n_i = n. \tag{2.2}$$

Графически распределению (2.1), (2.2) соответствует **полигон частот** – ломаная, отрезки которой последовательно соединяют две соседние точки $A_1(x_1;n_1)$, $A_2(x_2;n_2)$, $A_3(x_3;n_3)$,…, $A_k(x_k;n_k)$ на плоскости.

Если в законе распределения (2.1) перейти к относительным частотам

$$w_i = \frac{n_i}{n} \quad \left(\sum_{i=1}^{k} w_i = 1\right), \tag{2.3}$$

которые являются аналогами вероятностей, то получим статистическое распределение выборки через относительные частоты:

X	x_1	x_2	…	x_i	…	x_k
w	w_1	w_2	…	w_i	…	w_k

(2.4)

Статистическому распределению графически соответствует **полигон относительных частот** – ломаная, отрезки которой последовательно соединяют две соседние точки $B_1(x_1;w_1)$, $B_2(x_2;w_2)$, $B_3(x_3;w_3)$,…, $B_k(x_k;w_k)$ на плоскости.

Поскольку в (2.4) указаны аналоги вероятности каждой из вариант, то можно ввести аналоги числовых характеристик $M(x)$, $D(x)$ и $\sigma(x)$:

выборочное среднее

$$\bar{x}_в = \frac{\sum_{i=1}^{k} x_i n_i}{n}, \tag{2.5}$$

выборочная дисперсия

$$D_в = \overline{(x_в^2)} - (\bar{x}_в)^2, \text{ где } \overline{(x_в^2)} = \frac{\sum_{i=1}^{k} x_i^{\,2} n_i}{n}, \tag{2.6}$$

выборочное среднее квадратическое отклонение

$$\sigma_в = \sqrt{D_в}\,. \tag{2.7}$$

Для <u>непрерывной</u> случайной величины статистическое распределение задается в виде последовательности интервалов и соответствующих частот:

$(x_1;x_2]$	$(x_2;x_3]$	$(x_3;x_4]$	…	$(x_i;x_{i+1}]$	…	$(x_k;x_{k+1}]$
n_1	n_2	n_3	…	n_i	…	n_k

(2.8)

Частоты n_i показывают, сколько раз встретилось значение варианты x_i в полуинтервале $(x_i;x_{i+1}]$, формула (2.2) справедлива и в этом случае.

Количество интервалов k и длины интервалов h определяются формулами

$$k \approx 1 + 3{,}2 \cdot \lg n, \tag{2.9}$$

$$h = \frac{x_{\max} - x_{\min}}{k}, \tag{2.10}$$

где n – объем выборки, $x_{\max}$ и $x_{\min}$, соответственно, максимальное и минимальное значения случайной величины для данной выборки.

Интервальное распределение выборки (2.8) легко сводится к дискретному распределению (2.1), если в качестве вариант выбрать середины полуинтервалов $(x_i;x_{i+1}]$:

$y_1 = \frac{x_1 + x_2}{2}$	$y_2 = \frac{x_2 + x_3}{2}$	…	$y_i = \frac{x_i + x_{i+1}}{2}$	…	$y_k = \frac{x_k + x_{k+1}}{2}$
n_1	n_2	…	n_i	…	n_k

(2.11)

.

По формулам (2.5)-(2.7) можно вычислить числовые характеристики непрерывной случайной величины.

Графически интервальное распределение задается **гистограммой** – ступенчатой фигурой, основанием каждой ступеньки которой служат отрезки длины h: $[x_1; x_2]$, $[x_2; x_3]$,…, $[x_k; x_{k+1}]$, а соответствующие высоты равны $h_i = w_i / h$. Гистограмма относительных частот является аналогом функции плотности, так как площадь под ней всегда равна единице.

Аналогично функции распределения случайной величины вводится эмпирическая функция распределения

$$F^*(x) = \frac{n_x}{n}, \tag{2.12}$$

где n_x – число вариант меньших x, n – объем выборки.

Полученные по формулам (2.5)-(2.7) числовые оценки $x_в$, $D_в$ и $\sigma_в$ сами являются случайными величинами. При разных выборках одного и того же объема их значения будут различными. Необходимо установить, с какой точностью δ они оценивают параметры $M(x)$, $D(x)$ и $\sigma(x)$ генеральной совокупности. Путь θ некоторый параметр генеральной совокупности, а θ_n^* точечная оценка параметра θ по некоторой выборке объема n. В теории вероятностей оценить предоставляется возможным только <u>вероятность</u> того события, что отклонение полученного по выборке параметра θ_n^* от истинного θ не превосходит величину δ (получить оценку с точностью δ):

$$p(|\theta_n^* - \theta| < \delta) = \gamma. \tag{2.13}$$

Вероятность γ называется **доверительной вероятностью**, или надежностью, а интервал $(\theta_n^* - \delta; \theta_n^* + \delta)$ **доверительным интервалом** с заданной надежностью γ.

Укажем доверительный интервал с заданной надежностью γ для оценки математического ожидания a генеральной совокупности, распределенной по нормальному закону:

если известно среднее квадратическое отклонение σ, то верна оценка

$$\bar{x}_в - \frac{t \cdot \sigma}{\sqrt{n}} < a < \bar{x}_в + \frac{t \cdot \sigma}{\sqrt{n}}, \tag{2.14}$$

где n – объем выборки, а t находится из равенства $\Phi(t) = \gamma / 2$ по таблице значений функции Лапласа $\Phi(t)$ [1. прил. 2];

если σ неизвестно, то так как $\sigma_в$ смещенная влево оценка, то σ заменяется на исправленное среднее квадратическое отклонение s:

$$s = \sqrt{\frac{n}{n-1}}\sigma_в, \tag{2.15}$$

и доверительный интервал для математического ожидания оценивается неравенствами:

$$\bar{x}_в - \frac{t_\gamma \cdot s}{\sqrt{n}} < a < \bar{x}_в + \frac{t_\gamma \cdot s}{\sqrt{n}}, \tag{2.16}$$

где $t_\gamma = t(\gamma, n)$ находится из таблицы [1. прил. 3].

Доверительный интервал для оценки среднего квадратического отклонения σ нормального распределения с заданной надежностью γ находится по формуле

$$\boldsymbol{s(1-q) < \sigma < s(1+q)}, \qquad (2.17)$$

где $\boldsymbol{s}$ есть исправленное среднее квадратическое отклонение (54), а $\boldsymbol{q = q(\gamma, n)}$ находится из [1. прил. 4].

<u>Пример .</u> Задано интервальное распределение выборки:

$(y_i; y_{i+1}]$	12-15	15-18	18-21	21-24	24-27	27-30
n_i	2	6	12	19	7	4

Требуется:
1) построить гистограмму относительных частот;
2) перейти к вариантам, записать эмпирическую функцию распределения и построить ее график;
3) найти точечные оценки $\bar{x}_в$ и $\sigma_в$;
4) считая генеральную совокупность нормально распределенной, найти доверительные интервалы для a и σ с надежностью $\gamma = 0{,}95$.

<u>Решение</u>. 1) Объем выборки $n = \sum_{i=1}^{6} n_i = 2+6+12+19+7+4 = 50$. По формуле (2.3) найдем относительные частоты и выпишем таблицу с относительными частотами:

$(y_i; y_{i+1}]$	12-15	15-18	18-21	21-24	24-27	27-30
w_i	0,04	0,12	0,24	0,38	0,14	0,08

Для построения гистограммы относительных частот на оси абсцисс отложим частичные интервалы длины $h = 3$, высота ступенчатой функции на каждом из интервалов равна $h_i = w_i / h$. Для лучшей наглядности выберем разный масштаб на осях координат.

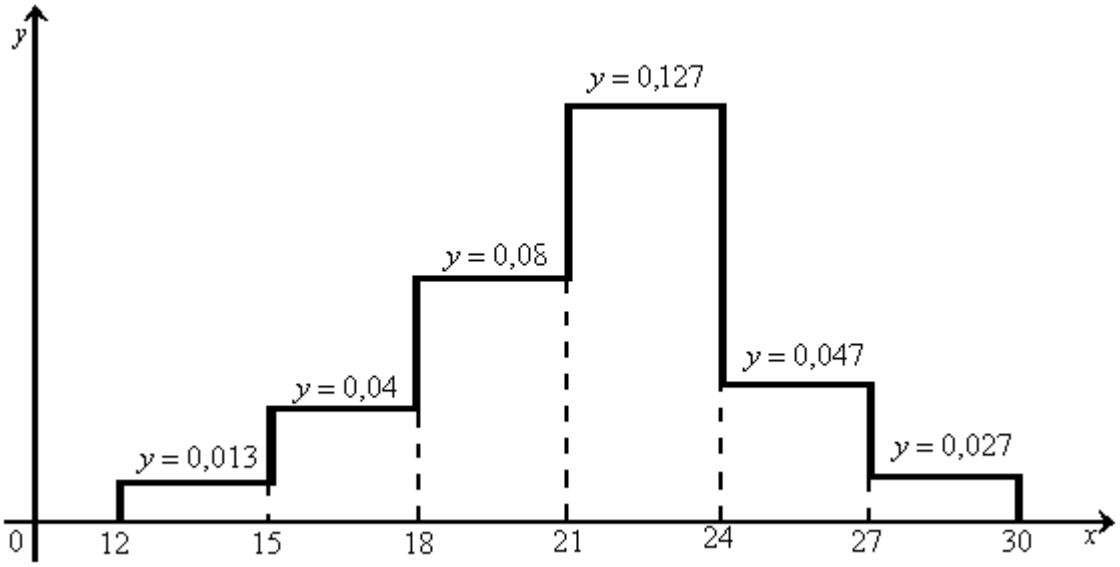


Рис. 1

2) Перейдем к вариантам (2.10), тогда данному в условии интервальному распределению выборки, будет соответствовать вариационный ряд

x_i	13,5	16,5	19,5	22,5	25,5	28,5
n_i	2	6	12	19	7	4
w_i	0,04	0,12	0,24	0,38	0,14	0,08

Аналогично тому, как находилась функция распределения $F(x)$ в примере 11, легко написать эмпирическую функцию $F^*(x)$, прибавляя к предыдущему ее значению следующее значение относительной частоты w_i (аналог вероятности p_i):

$F^*(x) = 0$, если $X \le 13{,}5$;

$F^*(x) = 0{,}04$, если $13{,}5 < X \le 16{,}5$;

$F^*(x) = 0{,}04 + 0{,}12 = 0{,}16$, если $16{,}5 < X \le 19{,}5$;

$F^*(x) = 0{,}16 + 0{,}24 = 0{,}4$, если $19{,}5 < X \le 22{,}5$;

$F^*(x) = 0{,}4 + 0{,}38 = 0{,}78$, если $22{,}5 < X \le 25{,}5$;

$F^*(x) = 0{,}78 + 0{,}14 = 0{,}92$, если $25{,}5 < X \le 28{,}5$;

$F^*(x) = 0{,}92 + 0{,}08 = 1$, если $28{,}5 < X$.

Наконец, построим график функции $F^*(x)$.

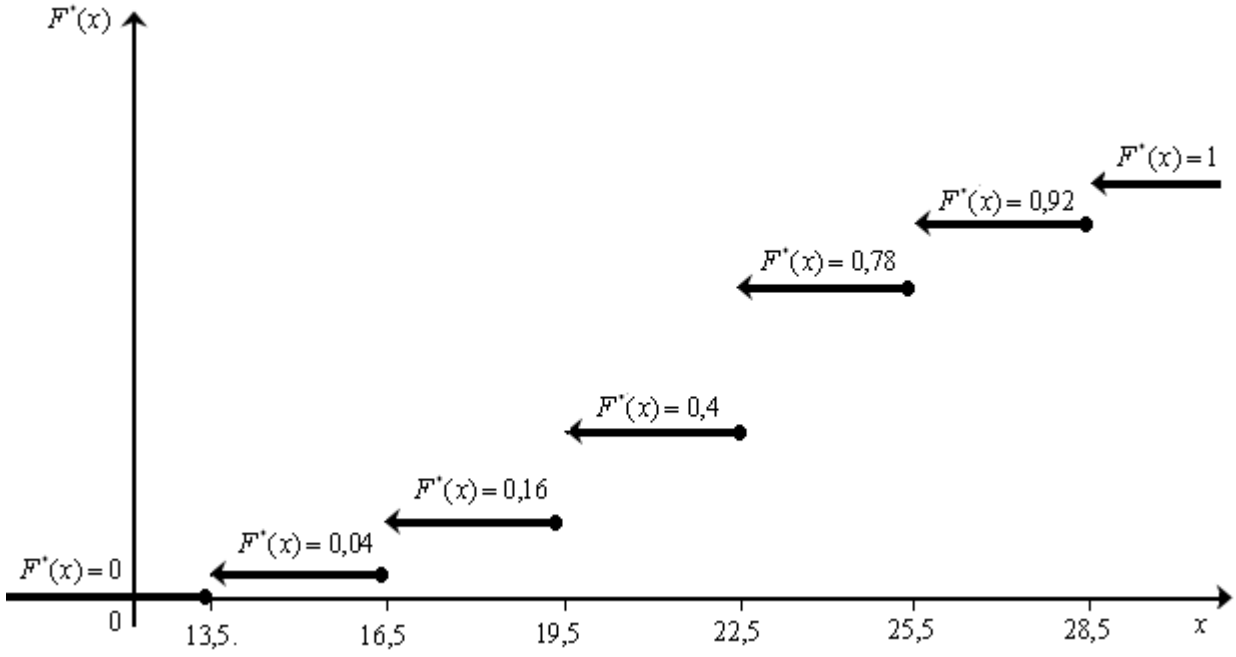


Рис. 2

3) Найдем точечные оценки $\bar{x}_в$ и $\sigma_в$ по формулам (44)-(46):

$$\bar{x}_в = \frac{\sum_{i=1}^{k} x_i n_i}{n} = \frac{(3{,}5\cdot 2 + 16{,}5\cdot 6 + 19{,}5\cdot 12 + 22{,}5\cdot 19 + 25{,}5\cdot 7 + 28{,}5\cdot 4)}{50} = 21{,}6\,;$$

$$\sigma_в(x) = \sqrt{\left(\overline{x_в^2}\right) - \left(\bar{x}_в\right)^2} =$$

$$= \sqrt{\frac{(3{,}5)^2\cdot 2 + (16{,}5)^2\cdot 6 + (19{,}5)^2\cdot 12 + (22{,}5)^2\cdot 19 + (25{,}5)^2\cdot 7 + (28{,}5)^2\cdot 4}{50} - (21{,}6)^2} = 3{,}6.$$

Обратим внимание на то, что вычисления достаточно трудоемки, несмотря на небольшой объем выборки. Укажем другой способ вычисления числовых характеристик выборки – **метод условных вариант**.

Сделаем замену переменных $u_i = \frac{x_i - x_{ср}}{h}$. В нашем случае $h = 3$, а из двух значений в середине таблицы выгодно выбрать среднее значение варианты $x_{ср} = 22{,}5$ с большей частотой, тогда $u_i = \frac{x_i - 22{,}5}{h}$ и таблица частот имеет вид

u_i	–3	–2	–1	0	1	2
n_i	2	6	12	19	7	4

Найти точечные оценки последнего вариационного ряда значительно легче:

$$\bar{u}_в = \frac{\sum_{i=1}^{k} u_i n_i}{n} = \frac{-3\cdot 2 - 2\cdot 6 - 1\cdot 12 + 0\cdot 19 + 1\cdot 7 + 2\cdot 4}{50} = -0{,}3\,;$$

$$\sigma_в(u) = \sqrt{\left(\overline{u_в^2}\right) - \left(\bar{u}_в\right)^2} = \sqrt{\frac{(-3)^2\cdot 2 + (-2)^2\cdot 6 + (-1)^2\cdot 12 + 0^2\cdot 19 + 1^2\cdot 7 + 2^2\cdot 4}{50} - (-0{,}3)^2} \approx 1{,}2\,.$$

Вернемся к первоначальной варианте, из замены $\bar{u}_в = \dfrac{\bar{x}_в - x_{ср}}{h}$ следует $\bar{x}_в = x_{ср} + \bar{u}_в \cdot h = 22{,}5 + (-0{,}3) \cdot 3 = 21{,}6$. Среднее квадратическое отклонение исходной варианты определяется по формуле $\sigma_в(x) = h \cdot \sigma_в(u) \approx 3 \cdot 1{,}2 = 3{,}6$.

4) В предыдущем пункте были найдены точечные оценки $\bar{x}_в = 21{,}6$, $\sigma_в(x) = 3{,}6$. Поскольку σ неизвестно, то по формулам (54)-(55) определим доверительный интервал для математического ожидания. Подставив значение объема выборки $n = 50$ в формулу (2.14), получим исправленное среднее квадратическое отклонение $s = \sqrt{\dfrac{50}{49}} \cdot 3{,}6 \approx 3{,}636$. Из [1. прил. 3] найдем $t_\gamma(\gamma, n) = t_\gamma(0{,}95; 50) = \ = 2{,}009$. Подставив найденные значения $\bar{x}_в = 21{,}6$, $n = 50$ и $s \approx 3{,}636$ в формулу (2.15), получим доверительный интервал для математического ожидания a:

$$21{,}6 - \frac{2{,}009 \cdot 3{,}636}{\sqrt{50}} < a < 21{,}6 + \frac{2{,}009 \cdot 3{,}636}{\sqrt{50}}.$$

После вычислений получим $20{,}57 < a < 22{,}63$.

По формуле (2.16) определим доверительный интервал для среднего квадратического отклонения σ. Из [1, прил. 4] найдем $q(\gamma; n) = q(0{,}95; 50) = 0{,}21$. Подставив найденные значения $q(\gamma; n) = 0{,}21$, $s \approx 3{,}636$ в формулу (2.16) получим доверительный интервал для среднего квадратического отклонения σ:

$$3{,}636 \cdot (1 - 0{,}21) < \sigma < 3{,}636 \cdot (1 + 0{,}21).$$

Окончательно получим $2{,}87 < \sigma < 4{,}4$.

<u>Ответ:</u> $\bar{x}_в = 21{,}6$; $\sigma_в(x) = 3{,}6$; $20{,}57 < a < 22{,}63$; $2{,}87 < \sigma < 4{,}4$.