**Министерство образования и науки Российской Федерации**
Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования
«**Ростовский государственный строительный университет»**

А.Р.Губеладзе, Е.Н.Яговкина, И.О.Губеладзе

# Введение в теорию ошибок измерений

Утверждено на заседании редакционно-издательского совета
университета в качестве учебного пособия

**Ростов-на-Дону**

**2 0 1 5**

**УДК** 528.1
**ББК** 26.104

Рецензент: профессор кафедры
«Горное дело»
Южно-Российского Государственного Политехнического Университета
имени М.И. Платова Ю.В. Посыльный

Губеладзе А.Р., Яговкина Е.Н., Губеладзе И.О. Введение в теорию ошибок измерений: Учебное пособие.- Ростов н/ Д: Рост. гос. строит. ун-т, 2015.-162 с.: ил.

Учебное пособие составлено по основным разделам программы курса “Теория математической обработки геодезических измерений” для студентов геодезических специальностей. Изложены основы теории ошибок измерений с элементами теории вероятностей. Рассмотрены способы обнаружения и устранения случайных и систематических ошибок, принципы обработки результатов равноточных и неравноточных измерений. Приведены методы оценки параметров распределения ошибок наблюдений.

В учебном пособии рассмотрены вопросы применения математической статистики к исследованию результатов измерений и их ошибок. Основное внимание уделяется прикладной стороне вероятностно-статистического анализа для оценки точности измерений.

Приведены общие закономерности распределения случайных величин и определение их числовых характеристик. Рассмотрены вопросы, связанные с методикой исследования ошибок измерений.

В заключительной части изложены вопросы дисперсионного и корреляционного анализов результатов геодезических измерений. Приведена качественная теория систематических ошибок измерений.

Данное учебное пособие предназначено для студентов геодезических специальностей вузов.

## Введение

При любой точности измерений полученные результаты содержат в себе ошибки, вызванные изменяющимися условиями, в которых производились эти измерения. Необходимую точность можно обеспечить методикой измерений с учетом комплекса всех условий. Но при этом возникают задачи получения наиболее надежных окончательных результатов с достаточной степенью точности.

В общем виде решаемые задачи формулируются следующим образом:

1) изучение законов распределения ошибок измерений и нахождение надежной меры точности полученных значений;

2) установление допусков, ограничивающих применение результатов измерений в заданных пределах;

3) нахождение наиболее надежного значения из всех результатов многократных измерений одной величины;

4) оценка точности как самих измерений, так и функций от измеренных значений;

5) алгоритмическое обеспечение математической обработки результатов геодезических измерений.

Геодезист должен разработать такую методику выполнения работ, при которой ошибки не превышали установленных допусков. Истинные значения измеряемых величин неизвестны, однако можно приблизиться к ним с той или иной точностью. В практической деятельности перед геодезистом возникают следующие задачи:

1) установление необходимой и достаточной точности измерений;

2) определение методов и средств для получения установленной точности;

3) использование подходящих критериев (допусков), позволяющих говорить о надежности результатов измерений;

4) выбор методов и средств обработки измеренных значений для получения наилучших результатов;

5) определение качества и точности выполненных измерений и полученных после обработки величин.

При этом всегда имеется в виду, что получение чрезмерной точности приведет к дополнительным затратам труда и использованием сложного и дорогостоящего оборудования. С другой стороны, недостаточная точность измерения может привести к ошибочным решениям, на исправление которых потребуются значительные затраты.

Изучением качества геодезических измерений, законов возникновения и действия неизбежных малых ошибок, разработкой правил оценки и расчетов необходимой точности измерений, а также методов и способов вычислений, позволяющих получить наилучшие результаты более экономически рациональными методами, и занимается теория математической обработки результатов геодезических измерений.

Вероятностно-статистический анализ дает возможность установить динамические или функциональные связи ошибок как между собой, так и с параметрами, характеризующими условия измерений. Случайная ошибка возникает под воздействием различных факторов, каждый из которых в свою очередь является источником возникновения более малых случайных ошибок. Таким образом, случайная ошибка является как бы результирующей пренебрегаемо малых ошибок по сравнению с суммарной. Эта теория аддитивности случайных ошибок получает вероятностное толкование.

Теория ошибок измерений изучает закономерности действия ошибок, способы нахождения наиболее достоверных результатов измерений и оценки точности выполненных работ. Поэтому для оценки точности необходимо использовать такие методы, которые использовали бы наиболее полно всю полученную в результате производства работ информацию.

Учитывая, что результаты измерений сопровождаются случайными ошибками, характер действия которых можно установить чаще всего из большого числа измерений, для исследования таких ошибок прибегают к вероятностностатистическим методам. Установленные статистическим путем закономерности распределения случайных ошибок позволяют уменьшить их влияние на результаты измерений.

Математическая обработка и анализ результатов измерений являются необходимыми как студентам, так и аспирантам, и инженерам-исследователям. Недостаточное знание современных методов математической обработки приводит к использованию упрощенных приемов анализа результатов измерений. Любые измерения сопровождаются ошибками, имеющими различную природу и оказывающими неодинаковое влияние на результаты измерений.

В данное пособие включены элементы теории вероятностей, так как математические методы обработки результатов измерений основываются на вероятностном представлении. Здесь важным является то, что исследование результатов измерений производится с определенной степенью надежности, т.е. получаемые выводы даются с определенной доверительной вероятностью.

Пособие написано в соответствии с учебной программой для студентов геодезических специальностей высших учебных заведений.

# 1. Основные понятия теории вероятностей

## 1.1. Определение вероятности события

Под событием понимается любой факт, который может произойти в результате испытаний. Различают события совместные и несовместные. События являются совместными, если в одном и том же испытании появление одного из них сопровождается появлением других. В противном случае события будут несовместными.

**Пример 1**. Завод выпускает геодезические инструменты трех видов:

$A$ - теодолит 2Т30П относится к I типу;

$B$ - теодолит 3Т5КП относится ко II типу;

$C$ - теодолит 3Т2КП относится к III типу.

Очевидно, что эти события являются несовместными.

При выполнении поверок могут быть обнаружены следующие несоответствия: ось установочного уровня не перпендикулярна оси вращения теодолита; визирная ось не перпендикулярна оси вращения зрительной трубы, которая в свою очередь не перпендикулярна оси вращения инструмента. Эти несоответствия могут быть обнаружены одновременно у всех трех инструментов. Следовательно, мы имеем дело с совместными событиями.

Событие называется достоверным, если оно произойдет в любом случае в данном испытании.

Событие называется невозможным, если оно не может произойти в условиях данного испытания.

**Пример 2.** В урне находятся только черные шары. Событие $A$ - появление черного шара. Событие $B$ - появление шара любого другого цвета (белого).

Осуществление события $A$ при выемке шара является достоверным, а события $B$ - невозможным.

Событие называется возможным или случайным, если в процессе испытаний оно может либо осуществиться, либо не осуществиться.

**Пример 3.** Производится стрельба из орудия при одном и том же положении. “Попадание в цель” или “промах” являются случайными событиями.

События могут быть равновозможными, если каждое из них в условиях конкретного испытания не имеет преимуществ перед другими, например, выпадение герба или цифры при бросании монеты.

Важным понятием является полная группа событий. Если события образуют полную группу, то в результате эксперимента одно из них обязательно должно произойти.

**Пример 4.** При бросании монеты единственно возможными событиями являются: “появление герба” и “появление цифры”. Появление других событий не возможно.

События, образующие полную группу, называются случаями или шансами.

Противоположным событием $\bar{A}$ является такое, которое должно обязательно наступить, если не произошло событие $A$. В предыдущем примере “появление герба” и “появление цифры” являются противоположными событиями.

Для сравнения между собой событий по степени возможности их появления используется количественная мера, называемая вероятностью события.

Вероятностью события является численная мера степени объективной возможности появления этого события.

Случай называется благоприятным или благоприятствующим некоторому событию, если его появление влечет за собой появление данного события. Если опыт сводится к “схеме случаев” или “схеме урн”, то вероятность может быть вычислена как отношение числа благоприятных событий к общему числу всех возможных случаев, т.е.

$$P(A) = \frac{m}{n}, \tag{1.1}$$

где $P(A)$ - вероятность появления события $A$;

$m$ - число случаев, благоприятных событию $A$;

$n$ - число всех возможных испытаний.

**Пример 5.** В урне находятся 20 синих и 40 красных шаров. Определить вероятность появления двух шаров синего цвета при одновременной выборке шаров из ящика.

**Решение.** Посчитаем число всех возможных случаев $n$ и число случаев $m$, благоприятствующих появлению двух синих шаров:

$$n = C_{60}^{2} = \frac{60!}{2!\cdot\left(60-2\right)!},$$

$$m = C_{20}^{2} = \frac{20!}{2!\cdot\left(20-2\right)!};$$

следовательно,

$$P\left(A\right) = \frac{m}{n} = \frac{C_{20}^{2}}{C_{60}^{2}} = \frac{\dfrac{20!}{2!\cdot\left(20-2\right)!}}{\dfrac{60!}{2!\cdot\left(60-2\right)!}} = \frac{19\cdot 20}{59\cdot 60} \approx 0{,}11.$$

Ответ: $P(A) \approx 0{,}11$ .

Из формулы (1.1) следует, что вероятность события является положительным числом и изменяется в пределах

$$0 \le P(A) \le 1 . \qquad (1.2)$$

Рассмотрим следующие свойства вероятности.

**Свойство 1.** Если все случаи являются благоприятствующими данному событию *A*, то это событие обязательно произойдет. Данное событие является достоверным и вероятность его появления $P(A) = 1$, так как $m = n$, то

$$P(A) = \frac{m}{m} = 1.$$

**Свойство 2.** Если нет ни одного случая, благоприятствующего данному событию *A*, то это событие в результате данного испытания произойти не может. Рассматриваемое событие является невозможным и его вероятность $P(A) = 0$, поскольку $m = 0$:

$$P(A) = \frac{0}{n} = 0 .$$

**Свойство 3.** Вероятность событий, образующих полную группу, равна единице, так как появление хотя бы одного из них в результате испытания будет достоверным событием.

**Свойство 4.** Вероятность противоположного события $\overline{A}$ вычисляется так же, как и вероятность события *A*

$$P(\overline{A}) = \frac{n-m}{n} = 1 - \frac{m}{n} \ . \qquad (1.3)$$

Отсюда

$$P(\overline{A}) = 1 - P(A) \ . \qquad (1.4)$$

**Пример 6.** В ящике находятся тщательно перемешанных 10 черных и 12 белых шаров. Найти вероятность того, что наудачу извлеченный шар окажется белым.

**Решение.** Появление белого шара обозначим как событие $A$. Общее число случаев составит $n = 22$, а благоприятных событию $A$ - $m = 12$. Следовательно,

$$P(A) = \frac{12}{22} = 0{,}54 \ .$$

Ответ: $P(A) = 0{,}54$ .

**Пример 7.** В урне 10 шаров: 6 белых и 4 черных. Вынули два шара. Какова вероятность, что оба шара белые?

**Решение.** Определим число всех случаев

$$n = C_{10}^2 = \frac{10 \cdot 9}{1 \cdot 2} = 45 .$$

Число благоприятных событий

$$m = C_6^2 = \frac{6 \cdot 5}{1 \cdot 2} = 15 .$$

Следовательно,

$$P(A) = \frac{m}{n} = \frac{15}{45} = \frac{1}{3} = 0{,}333 .$$

Ответ: $P(A) = 0{,}333$ .

**Пример 8.** В ящике находится 10 бракованных и 15 годных деталей. Из ящика вынимают три детали. Определить:

1) какова вероятность того, что все три детали будут годные;
2) какова вероятность, что годными окажутся две детали?

**Решение.** Событие $A$ - все три детали годные. Тогда общее число случаев

$$n = C_{25}^3 = \frac{25 \cdot 24 \cdot 23}{1 \cdot 2 \cdot 3} = 2300 .$$

Число случаев, благоприятных событию,

$$m = C_{15}^3 = \frac{15 \cdot 14 \cdot 13}{1 \cdot 2 \cdot 3} = 455 .$$

Искомая вероятность

$$P(A) = \frac{455}{2300} = 0{,}198 .$$

Ответ: $P(A) = 0, 198$.

## 1.2. Сложение вероятностей

Суммой двух событий *A* и *B* называется третье событие *C*, состоящее в появлении хотя бы одного из событий *A* и *B*. Если события несовместимы, то появление этих событий вместе отпадает и сумма событий *A* и *B* заключается в появлении либо события *A*, либо события *B*.

Суммой нескольких событий называется событие, состоящее в появлении хотя бы одного из этих событий.

**Теорема 1**. Вероятность появления одного из двух несовместных событий, безразлично какого, равна сумме вероятностей этих событий:

$$P(A + B) = P(A) + P(B)\ . \qquad (1.5)$$

Теорема справедлива для любого числа несовместных событий

$$P\left(\sum_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i). \qquad (1.6)$$

**Следствие 1.** Если несовместные события $A_1$, $A_2$, ... , $A_n$ образуют полную группу, то сумма вероятностей этих событий равна единице, т.е.

$$P(A_1) + P(A_2) + ... + P(A_n) = 1. \qquad (1.7)$$

**Следствие 2.** Сумма вероятностей противоположных событий равна единице:

$$P(A_1) + P(A) = 1. \qquad (1.8)$$

**Теорема 2.** Вероятность суммы двух совместных событий равна сумме вероятностей этих двух событий без вероятности их совместного появления:

$$P(A+B) = P(A) + P(B) - P(AB). \qquad (1.9)$$

## 1.3. Умножение вероятностей

Произведением двух случайных событий *A* и *B* называется сложное событие *C*, заключающееся в одновременном или последовательном осуществлении обоих событий.

Произведением нескольких событий называется событие, состоящее в совместном появлении всех этих событий.

Необходимо различать зависимые и независимые события. Два события называются независимыми, если появление одного из них не изменяет вероятность появления другого.

Несколько событий называются независимыми в совокупности, если любое из них не зависит от любой комбинации остальных. События называются зависимыми, если появление одного из них изменяет вероятность появления другого.

Под условной вероятностью события $B$ по отношению к событию $A$ (обозначение $P(B/A)$) понимается вероятность осуществления события $B$, определенная в предположении, что событие $A$ произошло.

**Теорема 3.** Вероятность произведения двух событий равна произведению вероятности одного из них на условную вероятность другого, вычисленную при условии, что первое событие произошло

$$P(AB) = P(A) \cdot P(B/A) \tag{1.11}$$

или

$$P(AB) = P(B) \cdot P(A/B). \tag{1.12}$$

**Пример 9.** В ящике 6 белых и 8 черных шаров. Из ящика вынули 2 шара (не возвращая вынутый шар в ящик). Найти вероятность того, что оба шара белые.

**Решение.** Событие $A$ - появление белого шара при первом вынимании. Событие $B$ - появление белого шара при втором вынимании. Для случая зависимых событий имеем

$$P(AB) = P(A) \cdot P(B/A). \tag{1.10}$$

Так как вероятности появления первого белого шара и второго шара в предположении, что осуществилось событие $A$,

$$P(A) = \frac{6}{6+8} = \frac{3}{7} = 0{,}43 \quad \text{и} \quad P(B/A) = \frac{6-1}{6+8-1} = \frac{5}{13} = 0{,}38 .$$

Следовательно,

$$P(AB) = 0{,}43 \cdot 0{,}38 = 0{,}16.$$

Ответ: $P(AB) = 0{,}16$.

**Теорема 4.** Вероятность произведения нескольких событий равна произведению вероятности одного из этих событий на условные вероятности других

$$P(A_1A_2 \dots A_n) = P(A_1)\, P(A_2/A_1) \dots P(A_n/A_1A_2 \dots A_{n-1}). \qquad (1.13)$$

**Следствие 1.** Если появление события $A$ не зависит от события $B$, то и появление события $B$ не зависит от события $A$.

Если появление события $A$ не зависит от события $B$, то можно записать

$$P(A) = P(A/B). \qquad (1.14)$$

Аналогично и для события $B$:

$$P(B/A) = P(B). \qquad (1.15)$$

**Следствие 2.** Вероятность произведения двух независимых событий равна произведению вероятностей этих событий

$$P(AB) = P(A) \cdot P(B). \qquad (1.16)$$

Для характеристики зависимости событий С.Н.Бернштейн предложил пользоваться коэффициентом взаимосвязанности или коэффициентом коллигации

$$\lambda = \frac{P(AB)}{P(A) \cdot P(B)} = \frac{P(A/B)}{P(A)} = \frac{P(B/A)}{P(B)}. \qquad (1.17)$$

Следовательно, коэффициент коллигации является отношением условной вероятности к безусловной. Практически удобнее выражать зависимость между событиями разностью $P(AB) - P(A) \cdot P(B)$. Если события независимы, то эта разность равна нулю. Показателем зависимости служит коэффициент корреляции

$$r(AB) = \frac{P(AB) - P(A) \cdot P(B)}{\sqrt{P(A) \cdot P(B) \cdot P(\overline{A}) \cdot P(\overline{B})}}. \qquad (1.18)$$

**Пример 10.** Три стрелка стреляют по цели. Вероятность попадания в цель первым стрелком равна 0,68, вторым - 0,82 и третьим - 0,90. Определить вероятность того, что все три стрелка одновременно попадут в цель.

**Решение.** Обозначим $A$- попадание в цель первым стрелком; $B$ - то же, вторым стрелком и $C$ - третьим, т.е.

$$P(A) = 0{,}68;\ P(B) = 0{,}82;\ P(C) = 0{,}90.$$

Тогда

$$P(ABC) = 0{,}68 \cdot 0{,}82 \cdot 0{,}90 = 0{,}50.$$

Ответ: $P(ABC) = 0{,}50$.

**Пример 11.** В условиях предыдущей задачи вычислить вероятность того, что хотя бы один стрелок попадет в цель.

**Решение.** Определим вероятность промаха каждым стрелком

$$P(\bar{A}) = 1 - 0{,}68 = 0{,}32; \quad P(\bar{B}) = 1 - 0{,}82 = 0{,}18; \quad P(\bar{C}) = 1 - 0{,}10\ .$$

Тогда вероятность одновременного промаха всех трех стрелков

$$P(\bar{A}\ \bar{B}\ \bar{C}) = 0{,}32 \cdot 0{,}18 \cdot 0{,}10 = 0{,}006.$$

Следовательно, искомая вероятность

$$P(ABC) = 1 - P(\bar{A}\ \bar{B}\ \bar{C}) = 1 - 0{,}006 = 0{,}994.$$

Ответ: $P(ABC) = 0{,}994$.

### 1.4. Формула полной вероятности. Формула Бейеса

**Теорема 5.** Если известно, что событие $A$ может произойти вместе с одним из событий $B_1, B_2, \ldots, B_n$, образующих полную группу несовместных событий, то вероятность события $A$ равна сумме произведений вероятностей каждого из событий $B_1, B_2, \ldots, B_n$ на соответствующую условную вероятность события $A$:

$$P(A) = \sum P(B_i) \cdot P(A / B_i)\ . \tag{1.19}$$

Эта формула носит название формулы полной вероятности.

Условная вероятность события $B_i$ в предположении, что событие $A$ имеет место, определяется по формуле Бейеса

$$P(B_i / A) = \frac{P(AB_i)}{P(A)} = \frac{P(B_i) \cdot P(A / B_i)}{\sum\limits_{i=1}^{n} P(B_i) \cdot P(A / B_i)}\ . \tag{1.20}$$

Вероятности $P(B_i/A)$, вычисленные по формуле Бейеса, называют вероятностями гипотез.

**Пример 12.** Имеется четыре урны. В первой урне 1 белый и 1 черный шары. Во второй - 2 белых и 3 черных шара. В третьей - 3 белых и 5 черных шаров. В четвертой - 4 белых и 7 черных шаров. Событие $B_i$ - выбор $i$-й урны ($i$ = 1, 2, 3, 4). Дано, что вероятность выбора $i$-й урны равна $\frac{i}{10}$, т.е.

$$P(B_1) = \frac{1}{10}; \quad P(B_2) = \frac{1}{5}; \quad P(B_3) = \frac{3}{10}; \quad P(B_4) = \frac{2}{5}.$$

Выбирая наугад одну из урн, вынимаем один шар. Вычислить вероятность того, что этот шар будет белый.

**Решение.** Согласно условию

$$P(A/B_1)=\frac{1}{2};\quad P(A/B_2)=\frac{2}{5};\quad P(A/B_3)=\frac{3}{8};\quad P(A/B_4)=\frac{4}{11}.$$

Вероятность появления белого шара определяем по формуле полной вероятности

$$P(A) = P(B_1)\ P(A/B_1)+P(B_2)\ P(A/B_2)+P(B_3)\ P(A/B_3)+P(B_4)\ P(A/B_4).$$

Следовательно,

$$P(A)=\frac{1}{10}\cdot\frac{1}{2}+\frac{1}{5}\cdot\frac{2}{5}+\frac{3}{10}\cdot\frac{3}{8}+\frac{2}{5}\cdot\frac{4}{11}=\frac{1707}{4400}=0{,}39.$$

Ответ: $P(A) = 0{,}39$ .

**Пример 13.** Имеются три одинаковых ящика. В первом ящике 20 белых шаров. Во втором - 10 белых и 10 черных шаров. В третьем - 20 черных шаров. Из выбранного наугад ящика вынули белый шар. Определить вероятность того, что белый шар вынут из первого ящика.

**Решение.** Гипотеза $B_1$ - выбор первого ящика;
Гипотеза $B_2$ - выбор второго ящика;
Гипотеза $B_3$ - выбор третьего ящика.

Событие $A$ - появление белого шара.

$P(B_1)=P(B_2)=P(B_3)=\frac{1}{3}$ - выбор любого из ящиков; $P(A/B_1) = 1$ - вероятность извлечения белого шара из первого ящика; $P(A/B_2)=\frac{10}{20}=\frac{1}{2}$ - вероятность извлечения белого шара из второго ящика; $P(A/B_3) = 0$ - вероятность извлечения белого шара из третьего ящика.

Искомую вероятность $P(B_1/A)$ находим по формуле Бейеса

$$P(B_1/A)=\frac{1\cdot\frac{1}{3}}{1\cdot\frac{1}{3}+\frac{1}{2}\cdot\frac{1}{3}+0\cdot\frac{1}{3}}=\frac{2}{3}=0{,}67.$$

Ответ: $P(B_1/A) = 0{,}67$.

## 2. Распределение случайной величины
### 2.1. Случайная величина

Случайной величиной называется величина, которая может принимать в результате опыта то или иное значение, о величине которого заранее ничего не известно. В отличие от случайного события случайная величина количественно характеризует результат испытания (опыта). Случайные величины подразделяются на дискретные и непрерывные.

Дискретной случайной величиной называется величина, которая на заданном интервале принимает конечное или бесконечное счетное множество значений.

Непрерывной случайной величиной называют величину, принимающую на определенном интервале бесчетное множество значений.

Дискретная случайная величина $X$ определяется последовательностью значений $x_1$, $x_2$, ... , $x_n$ , называемых вариантами, и вероятностями их появления $P(x_1)$, $P(x_2)$, ... , $P(x_n)$. Такие два ряда характеризуют распределение случайной величины (табл. 1).

Таблица 1

| $x$ | $x_1$ | $x_2$ | . . . | $x_i$ | . . . | $x_n$ |
|---|---|---|---|---|---|---|
| $P(x)$ | $P(x_1)$ | $P(x_2)$ | . . . | $P(x_i)$ | . . . | $P(x_n)$ |

Все возможные значения случайной величины $X$ составляют полную группу несовместных событий. Следовательно,

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

Практически при многократных измерениях полученные результаты $x_1$, $x_2$, ... , $x_n$ группируют по классам (интервалам), определяя частоту $n_1$, $n_2$, ... , $n_k$ появления этих результатов в соответствующем классе и вычисляя эмпирические (приближенные) значения вероятностей - частости

$$p_j = \frac{n_j}{n}, \qquad (2.2)$$

где $j = 1, 2, \ldots, k$ и

$$n = \sum_{j=1}^{k} n_j . \qquad (2.3)$$

Такое эмпирическое распределение можно представить в виде упорядоченного или вариационного (возрастающего) ряда (табл. 2).

Таблица 2

| Интервал | От $x_0$ до $x_1$ | От $x_1$ до $x_2$ | . . . | От $x_{n-1}$до $x_n$ | Σ |
|---|---|---|---|---|---|
| $n_j$ | $n_1$ | $n_2$ | . . . | $n_k$ | $n$ |
| $p_j$ | $p_1$ | $p_2$ | . . . | $p_k$ | 1,0 |

Применение систем классов необходимо при обработке больших рядов распределения случайной величины. Для определения шага интервала используют формулу Г. Стерджесса

$$\Delta x = \frac{x_{\max} - x_{\min}}{1 + 3{,}2 \lg n}, \qquad (2.4)$$

где $x_{max}$ , $x_{min}$ - соответственно максимальное и минимальное значения случайной величины;

$n$ - число наблюдений.

Практически длина и количество интервалов зависят от точности измерений и объема выборки, что может привести к неоднозначности выводов. Поэтому для характеристики распределения случайной величины $X$ пользуются функциями распределения. Различают интегральную и дифференциальную функции распределения. Первая называется функцией распределения или дистрибуантой, а вторая - плотностью вероятности или плотностью распределения.

Интегральная функция распределения $F(x_j)$ определяется вероятностью того, что случайная величина $X$ принимает значение, которое меньше действительного числа $x_j$, т.е.

$$F(x_j) = P(\, X < x_j \,)\,. \qquad (2.5)$$

Для дискретных случайных величин функция распределения находится по формуле

$$P(\, X < x_j \,) = P(x_1) + P(x_2) + \ldots + P(x_{j-1}) = \sum_{i=1}^{j-1} P(x_{j-1})\,. \qquad (2.6)$$

Функция (2.6) обладает всеми свойствами, присущими вероятности, т.е.

$$0 \le F(x_j) \le 1. \qquad (2.7)$$

Функция распределения - неубывающая и если $x_2 > x_1$ , то

$$F(x_2) > F(x_1)\ . \qquad (2.8)$$

Кроме рядов распределения большую роль при исследовании случайных величин играют накопленные частоты, которые определяются суммированием частот, начиная с частоты первого разряда и кончая частотой данного класса (интервала)

$$N_j = n_1 + n_2 + \ldots + n_j\ . \qquad (2.9)$$

Отношение накопленной частоты к сумме всех частот называется накопленной частостью или эмпирической функцией распределения $F_n\ (x)$

$$F_n(x) = \frac{N_j}{n}\ . \qquad (2.10)$$

На первоначальной стадии исследования эмпирических рядов распределения случайной величины $X$ можно использовать графический метод. Чаще всего строят следующие графики: полигон (рис.1), гистограмму (рис. 2), ступенчатый график накопленной частости (рис.3).

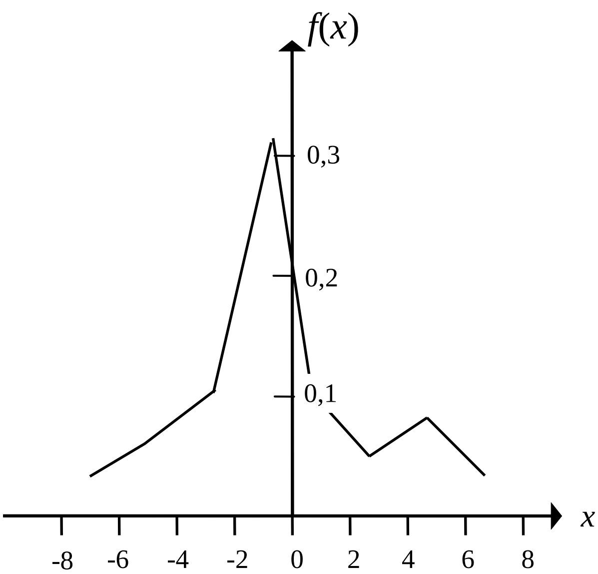


Рис. 1

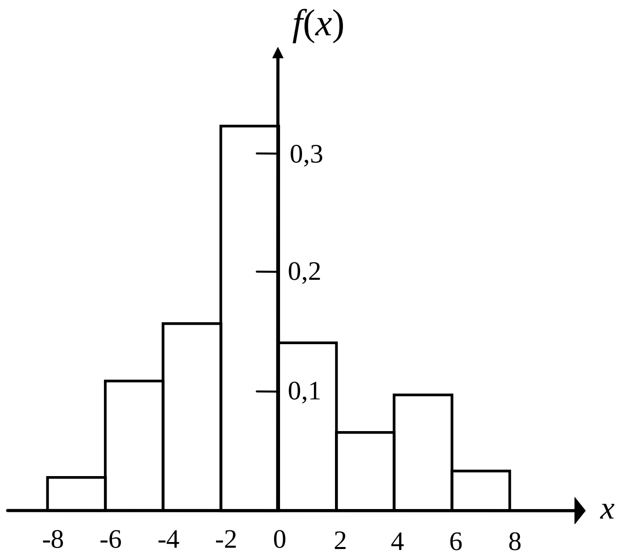


Рис. 2

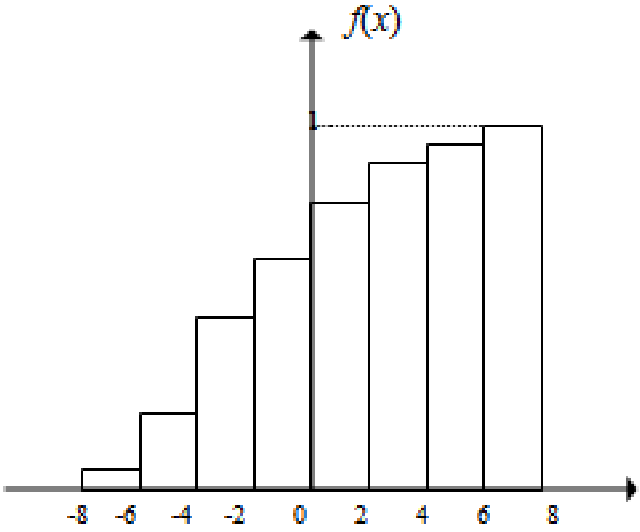

Рис. 3

Графики строятся системе прямоугольных координат. По оси абсцисс откладывают значения случайной величины, а ординат - частости $p_j$. Полигон имеет вид точечного графика, а гистограмма и ступенчатый график - системы прямоугольников, площади которых пропорциональны соответствующим значениям частостей или эмпирической функции распределения. Полигон и гистограмма дают приближенное представление о функции распределения. Поэтому, чтобы получить более или менее правдоподобное представление о плотности распределения по гистограмме и о функции распределения по ступенчатому графику, производят сглаживание или выравнивание статистических рядов. Для этого используют системы интерполяционных кривых распределения.

Рассмотрим свойства функций распределения.

**Свойство 1.** Дифференциальная функция распределения неотрицательна

$$f(x) \geq 0. \qquad (2.10)$$

Геометрически это означает, что кривая распределения расположена либо над осью абсцисс, либо на этой оси.

**Свойство 2.** Вероятность попадания непрерывной случайной величины $X$ в интервал $[\alpha, \beta)$ равна определенному интегралу от дифференциальной функции, взятому в пределах от $\alpha$ до $\beta$:

$$P(\alpha < X < \beta) = \int_{\alpha}^{\beta} f(x)dx. \qquad (2.11)$$

**Свойство 3.** Интегральная функция распределения может быть выражена через дифференциальную по формуле

$$F(x) = \int_{-\infty}^{x} f(x)\,dx \,. \qquad (2.12)$$

**Свойство 4.** Интеграл в бесконечных пределах от дифференциальной функции равен единице:

$$\int_{-\infty}^{\infty} f(x)\,dx = 1 \,. \qquad (2.13)$$

## 2.2. Центры группирования случайной величины

Законы распределения характеризуют случайную величину с вероятностной точки зрения. Но при этом возникает необходимость определять действительное значение этой величины и произвести оценку точности результатов измерения определяемого параметра. Это достигается отысканием центра группирования результатов измерения и показателя, характеризующего рассеивание таких результатов около центра группирования. Такие показатели называются числовыми характеристиками.

К числовым характеристикам центра группирования относятся математическое ожидание, начальные моменты первого порядка, медиана и мода. Математическое ожидание характеризует положение случайной величины на числовой оси, около которого размещаются все возможные значения этой величины.

Если имеем $x_1$, $x_2$, ... , $x_n$ - результаты измерения какой-либо величины с соответствующими вероятностями $p(x_1)$, $p(x_2)$, ... , $p(x_n)$ и образующих полную группу событий, т.е. $\sum_{i=1}^{n} p(x_i) = 1$, то

$$M(x) = \sum_{i=1}^{n} x_i \cdot p(x_i) \qquad (2.14)$$

Согласно теории ошибок получилось средневзвешенное значение из результатов измерений. При этом вес результата $x_i$ принят пропорционально вероятности $p(x_i)$.

Следовательно, математическим ожиданием дискретной случайной величины называется сумма произведений всех ее значений на соответствующие вероятности. В случае непрерывной случайной величины вместо $p(x_i)$ необходимо воспользоваться элементом вероятности $f(x)dx$, а сумму заменить интегралом. Таким образом, для непрерывной случайной величины математическое ожидание

$$M(x) = \int_{-\infty}^{\infty} x \cdot f(x)dx \ . \qquad (2.15)$$

Математическое ожидание непрерывной случайной величины $X$, значения которой принадлежат интервалу [ $a$, $b$ ], определяется

$$M(x) = \int_{a}^{b} x \cdot f(x)dx \ . \qquad (2.16)$$

Отметим свойства математического ожидания.

**Свойство 1.** Математическое ожидание суммы двух случайных величин равно сумме их математических ожиданий:

$$M(x+y) = M(x) + M(y) \ . \qquad (2.17)$$

**Свойство 2.** Математическое ожидание произведения двух независимых случайных величин равно произведению их математических ожиданий:

$$M(xy) = M(x) \cdot M(y) \ . \qquad (2.18)$$

**Свойство 3.** Математическое ожидание постоянной величины равно самой постоянной:

$$M(c) = c. \qquad (2.19)$$

**Свойство 4.** Постоянный множитель случайной величины может быть вынесен за знак математического ожидания:

$$M(cx) = c \cdot M(x) \ . \qquad (2.20)$$

**Свойство 5.** Математическое ожидание отклонения случайной величины от ее математического ожидания равно нулю:

$$M[X - M(x)] = M(x) - M(x) = 0 \ . \qquad (2.21)$$

Для определенных условий измерения математическое ожидание, в основном, будет величиной постоянной и около него колеблются все возможные значения случайной величины, в том числе и арифметические средние, вычисленные по результатам многократных наблюдений. В случае равноточных измерений центром группирования будет простая арифметическая средина

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \,. \qquad (2.22)$$

Для ряда равноточных измерений какой-либо случайной величины простая арифметическая средина является приближенным значением математического ожидания, т.е. $\overline{x} \approx M(x)$.

Математическое ожидание или генеральная средняя - это теоретическая или вероятнейшая характеристика центра группирования.

Арифметическая средняя $\overline{x}$ является эмпирической, статистической или выборочной числовой характеристикой центра группирования. В общем случае можно говорить об эмпирическом, статистическом или выборочном математическом ожидании. Для этого вероятности $p(x_i)$ заменим на приближенные им значения - частости $p_i$

$$\overline{x} = M^{*}(x) = \sum_{i=1}^{n} x_i \cdot p_i \,. \qquad (2.23)$$

При неограниченном возрастании числа измерений частости приближаются к вероятностям, и следовательно,

$$\overline{x} = M^{*}(x) \to M(x) \,. \qquad (2.24)$$

Иногда в качестве центров группирования результатов измерения используют медиану и моду. Установлено, что вероятность получить значения $x$ меньше медианы $Me(x)$ равна вероятности появления значения $x$ больших медианы, т.е. в этом случае справедливо равенство

$$P\,[\,x < Me(x)\,] = P\,[\,x > Me(x)\,] \,. \qquad (2.25)$$

Для вариационного ряда результатов равноточных измерений $x_1 < x_2 < \ldots < x_n$ приближенным значением теоретической медианы является срединное значение ряда, которое является эмпирическим, статистическим или выборочным

$$m_e \approx Me(x) \,. \qquad (2.26)$$

При симметричном распределении случайной величины

$$Me(x) = M(x). \qquad (2.27)$$

Медиана с геометрической точки зрения является абсциссой точки, в которой площадь, ограниченная кривой, делится пополам. Поскольку вся площадь под кривой равна единице, то функция распределения в данной точке равна 0,5

$$F(Me) = P(x < Me) = 0,5. \qquad (2.28)$$

Для проверки симметричности распределения случайной величины и дополнительного контроля вычислений определяют моду.

Теоретическая или вероятностная мода $Mo(x)$ - это наиболее вероятностное значение случайной величины. Геометрически мода представляет собой абсциссу точки, соответствующей максимуму кривой распределения. Различают одномодальные, двухмодальные и полимодальные распределения. Следует отметить, что если распределение одномодальное и симметричное, то математическое ожидание, медиана и мода имеют одно и тоже значение. В практике обработки результатов измерений находят эмпирическую, статистическую или выборочную моду

$$m_o \approx Mo(x) . \qquad (2.29)$$

## 2.3. Дисперсия и среднее квадратическое отклонение

Мерой рассеивания случайной величины около математического ожидания служат дисперсия и стандарт (среднее квадратическое отклонение).

Теоретическая или вероятностная дисперсия случайной величины $X$ определяется по формуле

$$D(x) = \sigma^2 = M[x - M(x)]^2 . \qquad (2.30)$$

Для дискретных значений случайной величины $x_1, x_2, \ldots, x_n$ с вероятностями $p(x_1), p(x_2), \ldots, p(x_n)$ дисперсия равна сумме произведений квадратов отклонений случайной величины от ее математического ожидания на соответствующие вероятности

$$D(x) = \sum_{i=1}^{n} p(x_i) \cdot \left[ x - M(x) \right]^2 . \qquad (2.31)$$

Для непрерывной случайной величины, если закон распределения задан в виде плотности вероятности, дисперсия составит

$$D(x) = \int_{-\infty}^{\infty} [x - M(x)]^2 \cdot f(x)dx. \qquad (2.32)$$

Существенным недостатком дисперсии является размерность квадрата. Среднее квадратическое отклонение, представляющее собой положительный корень квадратный из дисперсии

$$\sigma(x) = \sqrt{D(x)}, \qquad (2.33)$$

лишено указанного недостатка. Рассмотрим некоторые свойства дисперсии.

**Свойство 1.** Дисперсия суммы двух независимых случайных величин равна сумме дисперсий этих величин:

$$D(x + y) = D(x) + D(y). \qquad (2.34)$$

**Свойство 2.** Дисперсия случайной величины равна разности между математическим ожиданием квадрата случайной величины $X$ и квадратом ее математического ожидания:

$$D(x) = M(x^2) - [M(x)]^2. \qquad (2.35)$$

**Свойство 3.** Дисперсия постоянной величины равна нулю:

$$D(c) = 0. \qquad (2.36)$$

**Свойство 4.** Постоянный множитель случайной величины можно выносить за знак дисперсии, предварительно возведя его в квадрат:

$$D(cx) = c^2 \cdot D(x) \qquad (2.37)$$

**Свойство 5.** Дисперсия разности двух независимых случайных величин равна сумме их дисперсий:

$$D(x - y) = D(x) + D(y). \qquad (2.38)$$

**Свойство 6.** Дисперсия произведения двух независимых случайных величин $X$ и $Y$ определяется по формуле:

$$D(xy) = D(x) \cdot D(y) + [M(x)]^2 \cdot D(y) + [M(y)]^2 \cdot D(x). \qquad (2.39)$$

Для эмпирической или выборочной дисперсии, если $\Delta_i = x_i - a$, где $a$ - истинное значение величины, т.е. $M(x) = a$, то

$$S(x) = \frac{1}{n}\sum (x_i - a)^2 . \qquad (2.40)$$

Эмпирический или выборочный стандарт отдельного результата

$$s_x = m = \sqrt{S(x)} . \qquad (2.41)$$

В практике геодезических измерений эмпирический стандарт чаще всего называют средней квадратической ошибкой. Если известно истинное значение измеряемой величины, то

$$m = \sqrt{\frac{[\Delta\Delta]}{n}} . \qquad (2.42)$$

При неизвестном действительном значении измеряемой величины средняя квадратическая ошибка определяется через отклонения результатов измерений от арифметической средины $v_i$

$$m = \sqrt{\frac{[vv]}{n-1}} . \qquad (2.43)$$

Использование для оценки точности средней квадратической ошибки дает ряд преимуществ:

1) на ее величину наиболее сильно влияют большие по абсолютной величине ошибки;

2) средняя квадратическая ошибка наиболее устойчива и при небольшом числе измерений можно получить наилучшее представление о точности измерений;

3) средняя квадратическая ошибка используется и при установлении предельной ошибки измерений.

Пусть даны результаты независимых равноточных измерений $x_1, x_2, \ldots , x_n$ нормально распределенной случайной величины $X$. Определим дисперсию и стандарт для простой арифметической средины. Представим значение средней арифметической в виде линейной функции

$$\overline{x} = \frac{1}{n}x_1 + \frac{1}{n}x_2 + \ldots + \frac{1}{n}x_n. \qquad (2.44)$$

Согласно (2.34) и (2.37) получим

$$D(\overline{x}) = \frac{1}{n^2}D(x_1) + \frac{1}{n^2}D(x_2) + \ldots + \frac{1}{n^2}D(x_n) = \frac{1}{n^2}\sum_{i=1}^{n} D(x_i). \qquad (2.45)$$

Если случайные величины $x_i$ обладают одинаковым распределением, то они имеют одинаковые математические ожидания и дисперсии, т.е.

$$D(x_1) = D(x_2) = \ldots = D(x_n) = D(x).$$

Отсюда

$$D(\overline{x}) = \frac{D(x)}{n}, \qquad (2.46)$$

а следовательно,

$$\sigma(\overline{x}) = \frac{\sigma(x)}{\sqrt{n}}, \qquad (2.47)$$

где $\sigma(x) = \sigma(x_1) = \sigma(x_2) = \ldots = \sigma(x_n)$.

## 2.4. Моменты

Среди числовых характеристик важную роль играют начальные и центральные моменты.

Начальным моментом $\nu_s$ $s$-го порядка случайной величины $X$ называется математическое ожидание $s$-й степени этой случайной величины

$$\nu_s = M(x^s). \qquad (2.48)$$

Для дискретной случайной величины начальный момент $s$-го порядка

$$\nu_s = \sum_{i=1}^{n} x_i^s \cdot p(x_i) \qquad (2.49)$$

Очевидно, что начальный момент первого порядка соответствует математическому ожиданию случайной величины, т.е.

$$\nu_1 = M(x). \qquad (2.50)$$

Для непрерывных случайных величин

$$\nu_s = \int_{-\infty}^{\infty} x^s \cdot f(x)\,dx. \qquad (2.51)$$

Центральным моментом $\mu_s$ $s$-го порядка случайной величины $X$ называется математическое ожидание $s$-й степени отклонения величины $X$ от ее математического ожидания:

$$\mu_s = M[x - M(x)]^s. \tag{2.52}$$

Для дискретной случайной величины

$$\mu_s = \sum \left(x_i - a\right)^s \cdot p(x_i), \tag{2.53}$$

где $M(x) = a$ .

Если $s = 2$, это значит, что центральный момент второго порядка представляет собой дисперсию случайной величины

$$D(x) = \mu_2 .$$

Центральный момент s-го порядка для непрерывной случайной величины $X$ имеет вид

$$\mu_s = \int_{-\infty}^{\infty} \left[x - M(x)\right]^s f(x)\,dx. \tag{2.54}$$

Центральные моменты второго и более высоких порядков являются характеристиками рассеивания случайной величины.

При обработке результатов измерений обычно вычисляются эмпирические значения моментов. Эмпирический начальный момент $s$-го порядка

$$\nu_s^* = \frac{1}{n}\sum_{i=1}^{n} x_i^s . \tag{2.55}$$

При $s = 1$ начальный момент первого порядка равен простой арифметической средине

$$\nu_1^* = \overline{x} .$$

Эмпирический центральный момент $s$-го порядка

$$\mu_s^* = \frac{1}{n}\sum_{i-1}^{n} \left(x_i - a\right)^s . \tag{2.56}$$

Если математическое ожидание $M(x) = a$ неизвестно, а определено среднее арифметическое, то эмпирический центральный момент вычисляется согласно

$$\mu_s^* = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^s. \qquad (2.57)$$

При $s$ = 2 получаем формулы для определения эмпирической дисперсии

$$\begin{aligned} \mu_2^* &= S(x) = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - a\right)^2; \\ \mu_2^* &= S(x) = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2. \end{aligned} \qquad (2.58)$$

Между центральными и начальными моментами существуют следующие функциональные зависимости:

$$\mu_o = 1\ ; \qquad (2.59)$$

$$\mu_1 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - a\right) = 0\ ; \qquad (2.60)$$

$$\mu_2 = \nu_2 - \nu_1^2; \qquad (2.61)$$

$$\mu_3 = \nu_3 - 3\nu_2 \cdot \nu_1^2 + 2\nu_1^3; \qquad (2.62)$$

$$\mu_4 = \nu_4 - 4\nu_3 \cdot \nu_1 + 6\nu_2 \cdot \nu_1^2 - 3\nu_1^4. \qquad (2.63)$$

В случае больших рядов распределения возникает необходимость разбиения результатов измерения на классы (интервалы). В этом случае для каждого класса определяют частоту $n_i$ или частость $p_i$. Поскольку при этом находят шаг класса, то вычисляют относительную середину класса

$$y_j = \frac{x_j - c}{\Delta x}, \qquad (2.64)$$

где $x_j$ - середина установленного класса (интервала);

$c$ - “ложный нуль” (обычно принимается середина интервала с наибольшей частотой).

В дальнейшем используются относительные начальные моменты

$$\tilde{\nu}_s = \frac{1}{n}\sum_{j=1}^{k} n_j \cdot y_j^s. \qquad (2.65)$$

Тогда вероятнейшее значение измеряемой величины определится согласно

$$\overline{x} = \Delta x \cdot \tilde{\nu}_1 + c. \qquad (2.66)$$

Эмпирические центральные моменты второго, третьего и четвертого порядков вычисляются из следующих зависимостей

$$\mu_2^* = \Delta x^2 \cdot \left(\tilde{\nu}_2 - \tilde{\nu}_1^{\,2}\right); \qquad (2.67)$$

$$\mu_3^* = \Delta x^3 \cdot \left(\tilde{\nu}_3 - 3\tilde{\nu}_2 \cdot \tilde{\nu}_1 + 2\tilde{\nu}_1^3\right); \qquad (2.68)$$

$$\mu_4^* = \Delta x^4 \cdot \left(\tilde{\nu}_4 - 4\tilde{\nu}_3 \cdot \tilde{\nu}_1 + 6\tilde{\nu}_2 \cdot \tilde{\nu}_1^2 - 3\tilde{\nu}_1^4\right). \qquad (2.69)$$

Моменты применяют не только для вычисления параметров распределения и оценки точности, но и для характеристики рядов распределения. Однако с помощью моментов невозможно установить неизвестную функцию распределения, так как разные распределения могут иметь одинаковые моменты. Только при определенных условиях, используя различные критерии, можно установить закон распределения той или иной случайной величины. При этом большое внимание уделяется сопоставлению эмпирического распределения с теоретическим.

**Пример 14.** Эмпирическое распределение случайной непрерывной величины в геодезии получаем, производя измерения какой-либо физической величины. Результаты измерения угла (секунды) приведены в табл. 3.

Таблица 3

| N | Секунды измерен. угла | N | Секунды измерен. угла | N | Секунды измерен. угла | N | Секунды измерен. угла | N | Секунды измерен. угла |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 19,9 | 21 | 19,0 | 41 | 20,5 | 61 | 19,4 | 81 | 20,2 |
| 2 | 22,7 | 22 | 20,9 | 42 | 22,9 | 62 | 22,6 | 82 | 23,7 |
| 3 | 20,3 | 23 | 21,0 | 43 | 21,8 | 63 | 21,8 | 83 | 23,0 |
| 4 | 21,0 | 24 | 20,1 | 44 | 23,0 | 64 | 23,8 | 84 | 21,3 |
| 5 | 18,5 | 25 | 21,0 | 45 | 22,4 | 65 | 25,5 | 85 | 21,5 |
| 6 | 20,0 | 26 | 20,8 | 46 | 21,5 | 66 | 21,1 | 86 | 20,3 |
| 7 | 21,5 | 27 | 21,4 | 47 | 21,1 | 67 | 21,3 | 87 | 22,8 |
| 8 | 23,4 | 28 | 23,3 | 48 | 20,4 | 68 | 21,1 | 88 | 20,4 |
| 9 | 20,8 | 29 | 22,9 | 49 | 25,6 | 69 | 21,8 | 89 | 22,0 |
| 10 | 20,5 | 30 | 21,3 | 50 | 19,0 | 70 | 22,8 | 90 | 21,2 |
| 11 | 22,3 | 31 | 21,9 | 51 | 22,5 | 71 | 20,9 | 91 | 20,6 |
| 12 | 22,9 | 32 | 18,2 | 52 | 22,4 | 72 | 20,0 | 92 | 21,0 |
| 13 | 23,2 | 33 | 22,3 | 53 | 22,9 | 73 | 23,6 | 93 | 21,4 |
| 14 | 21,2 | 34 | 19,6 | 54 | 23,0 | 74 | 23,1 | 94 | 21,2 |
| 15 | 21,3 | 35 | 21,0 | 55 | 19,1 | 75 | 21,3 | 95 | 21,7 |
| 16 | 24,7 | 36 | 22,4 | 56 | 20,7 | 76 | 21,7 | 96 | 18,9 |
| 17 | 21,1 | 37 | 24,4 | 57 | 23,5 | 77 | 20,1 | 97 | 19,4 |
| 18 | 23,6 | 38 | 24,5 | 58 | 22,6 | 78 | 21,8 | 98 | 23,3 |
| 19 | 20,0 | 39 | 20,0 | 59 | 21,8 | 79 | 21,8 | 99 | 19,5 |
| 20 | 19,3 | 40 | 21,3 | 60 | 21,6 | 80 | 18,7 | 100 | 21,6 |

**Решение.** Если число значений наблюдаемого признака относительно большое, как например, 50 и более, то с целью упрощения вычислений, связанных с определением эмпирических числовых характеристик $\overline{x}$ , $s$ и прибегают к группировке. Шкала интересующего нас признака подразделяется на некоторое число интервалов; затем рассматривают не отдельные величины, а группы значений, попавших в последовательно расположенные интервалы.

Рассмотрим методику групповой обработки статистических данных, приведенных в табл. 3.

В этой таблице наименьшее значение угла равно 18,2″ и наибольшее 25,6″. Для проведения группирования установим шаг интервала, пользуясь формулой (2.4)

$$\Delta x = \frac{25{,}6 - 18{,}2}{1 + 3{,}2\lg 100} = 1.$$

Установив шаг и границы интервалов, вычисляем частоты, частости и величину накопленной частости. Результаты вычислений представлены в табл. 4.

Таблица 4

| Номера интервалов | Границы интервалов $\alpha_i$ | $\beta_i$ | Частоты $n_i$ | Частости $p_i$ | Накоплена частость Fn(x) |
|---|---|---|---|---|---|
| 1 | 2 | | 3 | 4 | 5 |
| 1 | от 18,0 | до 19,0 | 7 | 0,07 | 0,07 |
| 2 | 19,0 | 20,0 | 10 | 0,10 | 0,17 |
| 3 | 20,0 | 21,0 | 15 | 0,15 | 0,32 |
| 4 | 21,0 | 22,0 | 34 | 0,34 | 0,66 |
| 5 | 22,0 | 23,0 | 12 | 0,12 | 0,78 |
| 6 | 23,0 | 24,0 | 16 | 0,16 | 0,94 |
| 7 | 24,0 | 25,0 | 3 | 0,03 | 0,97 |
| 8 | 25,0 | 26,0 | 3 | 0,03 | 1,00 |
| | | | Σ 100 | 1,00 | |

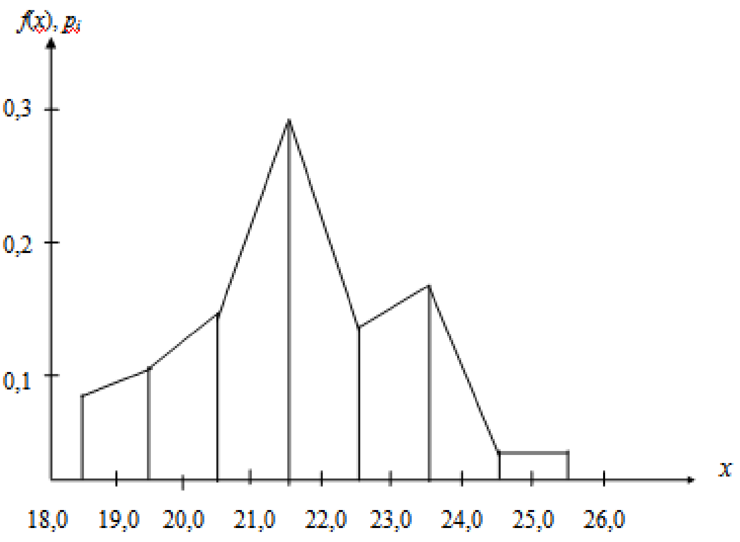


Рис. 4

Определим относительную середину интервала

$$y_i = \frac{x_i - c}{\Delta x},$$

где $c$ - “ложный нуль”, а $\Delta x$ - ширина интервала.

Относительные начальные моменты вычисляют по формуле

$$\overline{v}_r = \frac{1}{n}\sum_{i=1}^{n} n_i y_i^r .$$

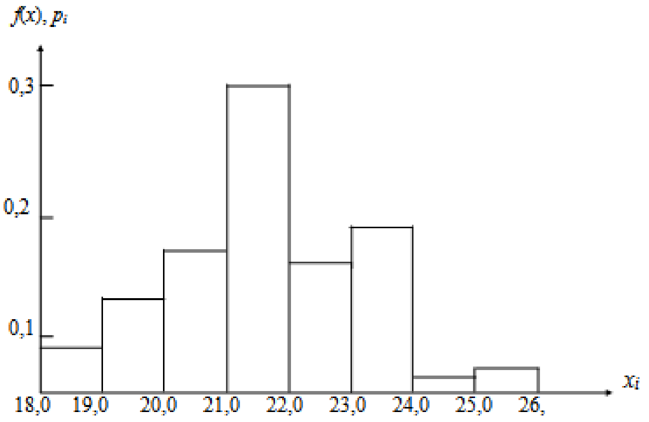


Рис. 5

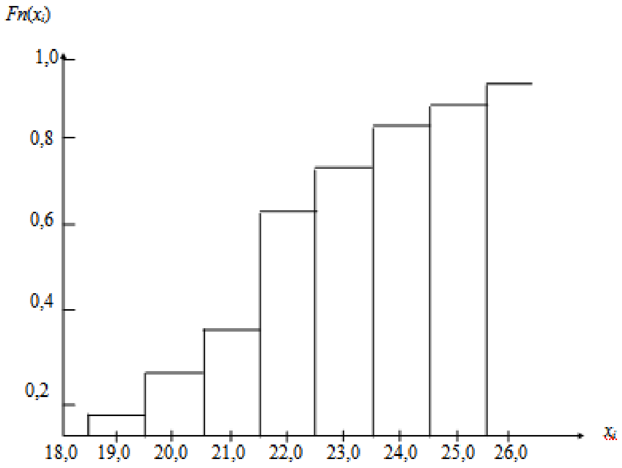


Рис. 6

Вычисление относительных начальных моментов выполнено в табл. 5.

Таблица 5

| Номер интер вала | Середина интервала $x_i$ | Относит. середина интервала $y_i$ | Частота $n_i$ | $n_i y_i$ | $n_i y_i^2$ | $n_i y_i^3$ | $n_i y_i^4$ |
|---|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 1 | 18,5 | -3 | 7 | -21 | 63 | -189 | 567 |
| 2 | 19,5 | -2 | 10 | -20 | 40 | -80 | 160 |
| 3 | 20,5 | -1 | 15 | -15 | 15 | -15 | 15 |
| 4 | 21,5 | 0 | 34 | 0 | 0 | 0 | 0 |
| 5 | 22,5 | 1 | 12 | 12 | 12 | 12 | 12 |
| 6 | 23,5 | 2 | 16 | 32 | 64 | 128 | 256 |
| 7 | 24,5 | 3 | 3 | 9 | 27 | 81 | 243 |
| 8 | 25,5 | 4 | 3 | 12 | 48 | 192 | 768 |
| Суммы | | | 100 | 9 | 269 | 129 | 2021 |
| Относительные начальные моменты | | | | 0,09 | 2,69 | 1,29 | 20,21 |

Вычислим дисперсию по формуле

$$S(x) = \mu_2 = 2{,}68,$$

откуда среднее квадратическое отклонение

$$s_x = \sqrt{S(x)} = 1{,}64''.$$

Асимметрия и эксцесс в данном примере будут:

$$A_s = \frac{\mu_3}{s_x^{\;3}} = \frac{0{,}56}{4{,}41} = 0{,}13; \qquad E_s = \frac{\mu_4}{s_x^{\;4}} - 3 = \frac{20{,}21}{7{,}23} = 0{,}20.$$

Ответ: $\overline{x}$ = 21,6”; $S(x)$ = 2,68 ; $s_x$ = 1,64”.

# 3. Кривые распределения
## 3.1. Биномиальное распределение

Конечным результатом исследования распределения случайной величины является определение кривой распределения значений этой величины. Основными типами распределения являются биномиальное распределение Бернулли, нормальное распределение Лапласа-Гаусса и распределение Пуассона.

Биномиальным распределением является распределение *к* числа появления события *A* при $n$ независимых испытаниях, причем вероятность появления данного события в каждом из них постоянна и равна $p$. При этом вероятность противоположного события $q = 1 - p$.

Вероятность $P_{n,k}$ появления такого события определенное число раз при $n$ независимых испытаниях определится по формуле Бернулли

$$P_{n,k} = C_n^k \cdot p^k \cdot q^{n-k}, \qquad (3.1)$$

где $C_n^k = \frac{n!}{k!(n-k)!}$ – число сочетаний из $n$ по $k$.

При больших выборках использование формулы Бернулли затруднено. Тогда для вычисления факториалов прибегают к формуле Стирлинга

$$n! \approx \sqrt{2\pi \cdot n} \cdot n^n \cdot e^{-n}\left(1 + \frac{1}{12n} + \frac{1}{288n^2} + \ldots\right). \qquad (3.2)$$

Выражение в скобках при $n \to \infty$ большого влияния на результат не оказывает, следовательно, формула Стирлинга примет вид

$$n! \approx \sqrt{2\pi \cdot n} \cdot \left(\frac{n}{e}\right)^n. \qquad (3.3)$$

Вероятнейшее число $k_o$ появления положительных ошибок в случае $n$ измерений может быть найдено либо из таблицы распределений, либо путем отыскания отношения смежных членов

$$\frac{P_{n,k}}{P_{n,k-1}} = \frac{p}{q}\left(\frac{n+1}{k} - 1\right). \qquad (3.4)$$

При $k = 1$ данное соотношение равно $\frac{np}{q}$ и при $k = n$ соответственно $\frac{p}{qn}$.

Найдем числовые характеристики случайной величины, подчиняющейся биномиальному распределению.

Начальные моменты $\nu_s$ $s$-го порядка вычисляются по формуле

$$\nu_s = \sum_{k=0}^{n} k^s C_n^k p^k q^{n-k} . \qquad (3.5)$$

Для вычисления сумм в формуле (3.5) продифференцируем несколько раз по $p$ следующее выражение:

$$(p+q)^n = \sum_{k=0}^{n} C_n^k p^k q^{n-k} .$$

В результате получим

$$n(p+q)^{n-1} = \sum_{k=0}^{n} k C_n^k p^{k-1} q^{n-k} , \qquad (3.6)$$

$$n(n-1)(p+q)^{n-1} = \sum_{k=0}^{n} k(k-1) C_n^k p^{k-2} q^{n-k} , \qquad (3.7)$$

$$n(n-1)(n-2)(p+q)^{n-3} = \sum_{k=0}^{n} k(k-1)(k-2) C_n^k p^{k-3} q^{n-k} , \qquad (3.8)$$

$$n(n-1)(n-2)(n-3)(p+q)^{n-4} = \sum_{k=0}^{n} k(k-1)(k-2)(k-3) C_n^k p^{k-4} q^{n-k} . \qquad (3.9)$$

Умножая левую и правую части равенства (3.6) на $p$, имеем

$$np(p+q)^{n-1} = \sum_{k=0}^{n} k C_n^k p^k q^{n-k} \qquad (3.10)$$

Поскольку правая часть равенства (3.10) соответствует начальному моменту первого порядка, следовательно, с учетом $p + q = 1$ будем иметь

$$\nu_1 = np(p+q)^{n-1} = np \qquad (3.11)$$

или

$$M(x) = np . \qquad (3.12)$$

Умножим обе части равенства (3.7) на $p^2$

$$n(n-1)p^2(p+q)^{n-2} = \sum_{k=0}^{n} k(k-1) C_n^k p^k q^{n-k} .$$

Учитывая, что $p + q = 1$, имеем

$$n^2 p^2 - np^2 = \sum_{k=0}^{n} k^2 C_n^k p^k q^{n-k} - \sum_{k=0}^{n} k C_n^k p^k q^{n-k} .$$

Используя равенства (3.5) и (3.11), получим начальный момент второго порядка

$$\nu_2 = n^2 p^2 - np^2 + np\,. \qquad (3.13)$$

Умножим обе части равенства (3.8) на $p^3$

$$n(n-1)(n-2)p^3(p+q)^{n-3} = \sum_{k=0}^{n} k(k-1)(k-2)C_n^k p^k q^{n-k}$$

или

$$n^3p^3 - 3n^2p^3 + 2np^3 = \sum_{k=0}^{n} k^3 C_n^k p^k q^{n-k} - 3\sum_{k=0}^{n} k^2 C_n^k p^k q^{n-k} + 2\sum_{k=0}^{n} m C_n^k p^k q^{n-k}\,.$$

Учитывая равенства (3.5) и (3.11), получим выражение начального момента третьего порядка

$$\nu_3 = n^3p^2 - 3n^2p^3 + 2np^2 + 3n^2p^2 - 3np^2 + np$$

или

$$\nu_3 = npq\,(q-p) + 3n^2p^2q + n^3p^3\,. \qquad (3.14)$$

Умножая обе части равенства (3.9) на $p^4$, получим

$$n(n-1)(n-2)(n-3)p^4(p+q)^{n-4} = \sum k(k-1)(k-2)(k-3)C_n^k p^k q^{n-k}$$

или

$$n^4p^4 - 6n^3p^4 + 11n^2p^4 - 6np^4 = \sum_{k=0}^{n} k^4 C_n^k p^k q^{n-k} - 6\sum_{k=0}^{n} k^3 C_n^k p^k q^{n-k} +$$

$$+11\sum_{k=0}^{n} k^2 C_n^k p^k q^{n-k} - 6\sum_{k=0}^{n} k C_n^k p^k q^{n-k}.$$

Учитывая равенства (3.6) - (3.9) и (3.11), (3.13) - (3.14), определяем

$$\nu_4 = n^4p^4 - 6n^3p^4 + 11n^2p^4 - 6np^4 + 6n^3p^3 - 18n^2p^3 + 12np^3 + 7n^2p^2 - 7np^2 + np$$

или

$$\nu_4 = npq\,(1-6pq) + n^2p^2q\,(7q-4p) + 6n^3p^3q + n^4p^4\,. \qquad (3.15)$$

Соответственно центральные моменты первых пяти порядков будут иметь следующий вид :

$$\begin{aligned} &\mu_0 = 1\,; \\ &\mu_1 = 0\,; \\ &\mu_2 = npq\,; \\ &\mu_3 = npq\,(q-p)\,; \\ &\mu_4 = npq\,[3(n-2)pq+1]\,. \end{aligned} \qquad (3.16)$$

Таким образом, дисперсия и среднее квадратическое отклонение для биномиального распределения

$$\mu_2 = \sigma^2 = D(x) = npq; \qquad (3.17)$$

$$\sigma = \sqrt{npq}. \qquad (3.18)$$

Показатели асимметрии и эксцесса для биномиального распределения имеют вид

$$A = \frac{q - p}{\sqrt{npq}}; \qquad (3.19)$$

$$E = \frac{1 - 6pq}{\sqrt{npq}}. \qquad (3.20)$$

При симметричном распределении, т.е. $p = q = 0{,}5$

$$\begin{aligned} M(x) &= 0{,}5\, n; \\ D(x) &= 0{,}25\, n; \\ \sigma(x) &= 0{,}5\sqrt{n}. \end{aligned} \qquad (3.21)$$

Биномиальному распределению подчиняются только дискретные случайные величины. В геодезической практике этому распределению соответствуют положительные или отрицательные случайные ошибки измерений.

## 3.2. Равномерное распределение

В геодезической практике встречаются ошибки измерений, обладающие равномерным или равновероятным распределением. Примером могут служить ошибки округления, а также ошибки визирования в случае, когда толстая нить покрывает тонкую визирную цель.

Непрерывная случайная величина $X$ имеет равномерное распределение на интервале $[a, b]$, если на этом интервале плотность распределения случайной величины постоянна, а вне его равна нулю, т.е.

$$\varphi(x) = \begin{cases} 0 & \text{при} \quad x < a \\ c & \text{при} \quad a < x < b \\ 0 & \text{при} \quad x > b \end{cases}, \qquad (3.22)$$

где $c$ = $const$. Равномерное распределение иногда называют законом равномерной плотности. Графики равномерного распределения - на рис. 7 и 8.

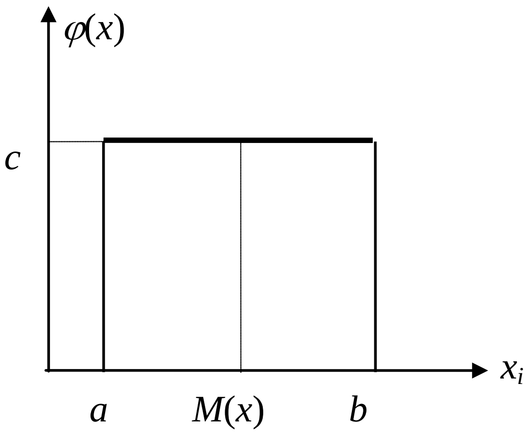


Рис. 7

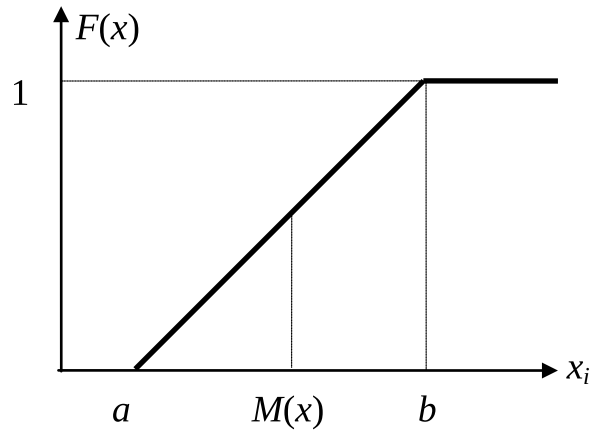


Рис. 8

Найдем значение постоянной $c$. Так как площадь, ограниченная кривой распределения, равна единице, и все значения случайной величины принадлежат интервалу $[a, b]$, то должно выполняться равенство

$$\int_a^b \varphi(x)dx = 1 \qquad (3,23)$$

или

$$\int_a^b c dx = 1,$$

отсюда

$$c = \frac{1}{b-a}\ . \qquad (3.24)$$

Итак, закон равномерного распределения аналитически можно записать в виде

$$\varphi(x) = \begin{cases} 0 & \text{при} \quad x < a \\ \dfrac{1}{b\text{-}a} & \text{при} \quad a < x < b \\ 0 & \text{при} \quad x > b\ . \end{cases} \qquad (3.25)$$

Функция распределения $F(x)$ для равномерного распределения на интервале $[a, b]$

$$F(x)=\int_a^b \varphi(x)dx=\int_a^b \frac{1}{b-a}dx=\frac{\mathrm{x}}{\mathrm{b-a}}\Bigg|_a^x=\frac{x-a}{b-a}. \qquad (3.26)$$

При $x < a$ имеем $F(x) = 0$, а при $x > b$ имеем $F(x) = 1$. Математическое ожидание находится по формуле

$$M(x)=\int_a^b x\varphi(x)\,dx=\frac{1}{b-a}\int_{\mathrm{a}}^b x\,dx=\frac{b^2-a^2}{2(b-a)}=0{,}5(a+b) \qquad (3.27)$$

и соответствует медиане. Закон равномерной плотности моды не имеет. Дисперсия определяется согласно

$$D(x)=\int_a^b \{x-M(x)\}^2\varphi(x)\,\mathrm{dx}=\frac{1}{b-a}\int_a^b\left(x-\frac{a+b}{2}\right)^2dx=\frac{(b-a)^2}{12}. \qquad (3.28)$$

Отсюда

$$\sigma(x)=\frac{b-a}{2\sqrt{3}}\ . \qquad (3.29)$$

В случае симметричного распределения случайной величины, т.е. при $a = -b$,

$$M(x)=0;\ \ D(x)=\frac{b^2}{3};\ \ \sigma(x)=\frac{b}{\sqrt{3}}. \qquad (3.30)$$

### 3.3. Нормальное распределение

Нормальное распределение - наиболее часто встречающийся вид распределения. С ним приходится иметь дело при анализе ошибок измерений, при исследовании осадки сооружений и т.д.

Следует иметь в виду, что нормальному закону распределения подчиняются только непрерывные случайные величины. Поэтому распределение нормальной совокупности может быть задано в виде плотности распределения

$$\varphi(x)=\frac{1}{\sigma(x)\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2(x)}} \qquad (3.31)$$

или в виде функции распределения

$$F(x)=\int_{-\infty}^x \varphi(x)\,\mathrm{dx}=\frac{1}{\sigma(x)\sqrt{2\pi}}\int_{-\infty}^x e^{-\frac{(x-a)^2}{2\sigma^2(x)}}dx. \qquad (3.32)$$

Нормальное распределение двухпараметрическое с параметрами: математическое ожидание $M(x) = a$ и среднее квадратическое отклонение или стандарт $\sigma(x)$.

График плотности нормального распределения (рис. 9) называют нормальной кривой распределения.

Из формулы (3.31) следует, что нормальное распределение симметрично относительно ординаты, соответствующей значению $x = a$. Ордината в данной точке соответствует максимуму и равна $\frac{1}{\sigma(x)\sqrt{2\pi}}$. При удалении от точки $a$ по оси абсцисс плотность вероятности убывает и при $x \to \pm\infty$ кривая асимптотически приближается к оси абсцисс. При изменении центра группирования кривая распределения смещается вдоль оси абсцисс, не изменяя своей формы, которую характеризует параметр $\sigma(x)$. Площадь, ограниченная кривой распределения, не изменяется и остается равной единице (рис. 10). Определим математическое ожидание и дисперсию случайной величины, подчиняющейся закону нормального распределения

$$M(x) = \int_{-\infty}^{+\infty} x\,\varphi(x)\,dx = \frac{1}{\sigma(x)\sqrt{2\pi}} \int_{-\infty}^{+\infty} x\, e^{-\frac{(x-a)^2}{2\sigma^2(x)}} dx\,. \qquad (3.33)$$

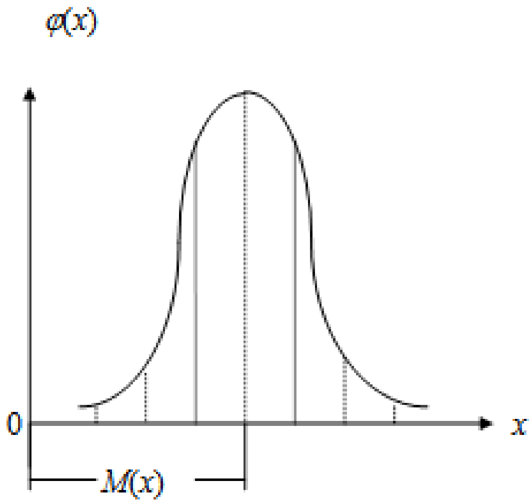


Рис. 9

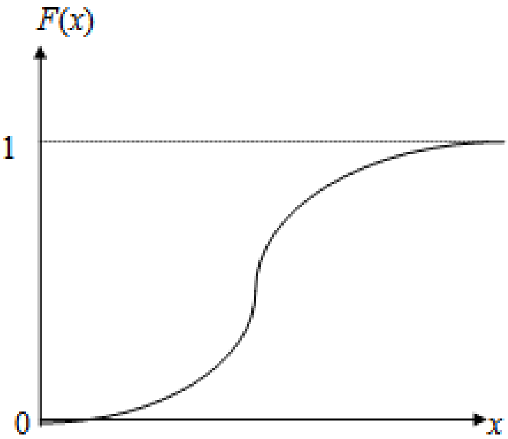


Рис. 10

Введем переменную $t = \frac{x-a}{\sigma(x)}$, которую назовем нормированной случайной величиной, откуда имеем $x - a = \sigma(x)\cdot t$ и $dx = \sigma(x)\cdot dt$.

Следовательно, вместо зависимости (3.33) получим

$$M(x) = \frac{\sigma(x)}{\sigma(x)\sqrt{2\pi}}\int_{-\infty}^{+\infty}\left\{\sigma(x)\ t + a\right\}\ e^{-\frac{t^2}{2}}dt = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\sigma(x)\ t\, e^{-\frac{t^2}{2}}dt\ + \frac{a}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{t^2}{2}}dt = a.$$

В правой части первое слагаемое равно нулю, а второе - $a$. Итак, математическое ожидание $M(x) = a$.

Для дисперсии

$$D(x) = \int_{-\infty}^{+\infty}(x-a)^2\varphi(x)\ dx = \frac{1}{\sigma(x)\sqrt{2\pi}}\int_{-\infty}^{+\infty}(x-a)^2\, e^{-\frac{(x-a)^2}{2\sigma^2(x)}}dx = \frac{\sigma^3(x)}{\sigma(x)\sqrt{2\pi}}\int_{-\infty}^{+\infty}t^2 e^{-\frac{t^2}{2}}dt =$$

$$= \frac{\sigma^2(x)}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}t\cdot t\, e^{-\frac{t^2}{2}}dt.$$

Интегрируя по частям и обозначив $u = t$ и $t\, e^{-\frac{t^2}{2}}dt = dv$, получим $du = dt$, $v = \int_{-\infty}^{+\infty}t\, e^{-\frac{t^2}{2}}dt$. Окончательно имеем $D(x) = \sigma^2(x) = \sigma^2$.

Главная особенность закона нормального распределения состоит в том, что он является предельным.

Нормальное распределение является достаточно правдоподобной моделью возникновения случайных ошибок измерений, в которой рассматривается возможность появления ошибок от $-\infty$ до $+\infty$. При этом используется нормированная функция плотности распределения с параметрами $M(t) = 0$ и $\sigma(t) = 1$

$$\varphi(t) = \frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}. \qquad (3.34)$$

Нормированную плотность распределения используют для расчета теоретической кривой распределения, соответствующей данному эмпирическому ряду. Значения нормированной плотности, определенные согласно (3.34), для различных параметров $t$ приведены в табл. 1 приложения. Следует иметь ввиду, что нормированная плотность распределения является четной функцией, т.е. $\varphi(-t) = \varphi(t)$ .

Нормированная нормальная функция распределения

$$F(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{t^2}{2}} dt \quad . \tag{3.35}$$

## 3.4. Вычисление вероятностей попадания случайной величины в интервал

Знание функции плотности распределения случайных ошибок измерений дает возможность для решения целого ряда задач, возникающих при различных геодезических работах: определение наиболее вероятного значения при многократных измерениях; установление предельных значений (допусков) для конкретного вида работ; вычисление вероятности появления случайной ошибки в определенном интервале; выявление предельных значений, за которыми ошибки можно квалифицировать как грубые. Для этого воспользуемся функцией Лапласа, имеющей следующий вид:

$$\Phi(t) = \frac{2}{\sqrt{2\pi}} \int_{0}^{t} e^{-\frac{t^2}{2}} dt. \tag{3.36}$$

Вероятность того, что любая случайная величина $X$ примет значение, принадлежащее интервалу ($\alpha$, $\beta$), можно выразить через функцию распределения

$$P(\alpha < x < \beta) = F(\beta) - F(\alpha). \tag{3.37}$$

Перейдем к нормированной случайной величине $t$. Неравенства $\alpha < x < \beta$ и $\frac{\alpha - a}{\sigma} < \frac{x - a}{\sigma} < \frac{\beta - a}{\sigma}$ равносильны. Поэтому вероятности этих неравенств равны между собой:

$$P(\alpha < x < \beta) = P\left(\frac{\alpha - a}{\sigma} < \frac{x - a}{\sigma} < \frac{\beta - a}{\sigma}\right). \tag{3.38}$$

Используя равенство ( 3.37 ), запишем

$$P(t_\alpha < t < t_\beta) = F(t_\beta) - F(t_\alpha), \tag{3.39}$$

где $t_\alpha = \dfrac{\alpha - a}{\sigma}$; $t_\beta = \dfrac{\beta - a}{\sigma}$.

Считая, что нормированная случайная величина $t$ имеет нормальное распределение, и, используя определение интегральной функции, получим

$$F(t) = \int_{-\infty}^{t} \varphi(t)dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{t^2}{2}} dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} e^{-\frac{t^2}{2}} dt + \frac{1}{\sqrt{2\pi}} \int_{0}^{t} e^{-\frac{t^2}{2}} dt. \tag{3.40}$$

Первое слагаемое числено равно половине площади под нормированной кривой распределения; второе - является выражением функции Лапласа. Итак, для нормально распределенной случайной величины имеем

$$F(t) = \frac{1}{2} + \frac{1}{2}\Phi(t). \tag{3.41}$$

Используя формулу (3.41) в выражении (3.39), получим

$$P\left(t_\alpha < t < t_\beta\right) = \frac{1}{2}\Phi(t_\beta) - \frac{1}{2}\Phi(t_\alpha). \tag{3.42}$$

При этом следует иметь ввиду, что функция Лапласа является нечетной, т.е. $\Phi(-t) = -\Phi(t)$.

### 3.5. Определение вероятности отклонения случайной величины от ее математического ожидания

Определим вероятность того, что нормально распределенная величина $X$ отклоняется от своего математического ожидания на величину, меньшую чем $\varepsilon$, т.е. найдем вероятность осуществления неравенства $a - \varepsilon < x < a + \varepsilon$. Перейдя к нормированной случайной величине $t$, имеем

$$t = \frac{\varepsilon}{\sigma}. \tag{3.43}$$

Согласно равенству (3.42)

$$P\left(t_\alpha < t < t_\beta\right) = \frac{1}{2}\Phi(t_\beta) - \frac{1}{2}\Phi(t_\alpha) = \frac{1}{2}\Phi\left(\frac{\varepsilon}{\sigma}\right) - \frac{1}{2}\Phi\left(-\frac{\varepsilon}{\sigma}\right) = \Phi\left(\frac{\varepsilon}{\sigma}\right). \tag{3.44}$$

Окончательно получим

$$P\left(\left|x - a\right| < \varepsilon\right) = \Phi\left(\frac{\varepsilon}{\sigma}\right). \tag{3.45}$$

При $M(x) = a = 0$ равенство (3.45) примет вид

$$P(|x| < \varepsilon) = \Phi\left(\frac{\varepsilon}{\sigma}\right). \qquad (3.46)$$

Выразив отклонение случайной величины $X$ в долях среднего квадратического отклонения, т.е. положив $\varepsilon = t\sigma$, равенство (3.45) можно записать

$$P(|x - a| < t\sigma) = \Phi(t). \qquad (3.47)$$

Таким образом, значение $\Phi(t)$ при заданном $t$ определяет вероятность того, что отклонение нормально распределенной величины по абсолютному значению меньше $t\sigma$.

При $t = 1$ имеем $P(|x - a| < \sigma) = \Phi(1) = 0{,}6827;$

при $t = 2$ имеем $P(|x - a| < 2\sigma) = \Phi(2) = 0{,}9545;$

при $t = 3$ имеем $P(|x - a| < 3\sigma) = \Phi(3) = 0{,}9973.$

Из последнего равенства следует, что практически рассеивание случайной величины укладывается на участке $M(x) \pm 3\sigma$. Вероятность того, что случайная величина $X$ попадает за этот интервал очень мала, а именно равна 0,0027, т.е. это событие можно считать практически невозможным. На приведенном рассуждении основано “правило трех сигм”, которое формулируется следующим образом: если случайная величина имеет нормальное распределение, то отклонение этой величины от ее математического ожидания по абсолютной величине не превосходит утроенного среднего квадратического отклонения.

**Пример 15**. Вычислить вероятность попадания значения $X$ горизонтального угла в интервал от 42°35′19,5″ до 42°35′22,5″, если вероятнейшее значение этого угла составило 42°35′21,5″, а среднее квадратическое отклонение равно 1,25″.

**Решение**. Для определения вероятности попадания нормально распределенной случайной величины $X$ в заданный интервал воспользуемся функцией Лапласа. При этом согласно соотношению (3.42)

$$P(\alpha < X < \beta) = \frac{1}{2}\Phi(t_\beta) - \frac{1}{2}\Phi(t_\alpha),$$

в котором $t_\alpha = \dfrac{\alpha - a}{\sigma}$ и $t_\beta = \dfrac{\beta - a}{\sigma}$,

где $a$ = 42°35′21,5” и $\sigma$ = 1,25”.

Тогда

$$P\left(19{,}5'' < X < 22{,}5''\right) = \frac{1}{2}\Phi\left(\frac{22{,}5-21{,}5}{1{,}25}\right) - \frac{1}{2}\Phi\left(\frac{19{,}5-21{,}5}{1{,}25}\right) =$$

$$= \frac{1}{2}\Phi(0{,}8) - \frac{1}{2}\Phi(-1{,}6) = 0{,}288 + 0{,}445 = 0{,}733.$$

Согласно табл. 1 приложения значения функции $\Phi(t)$ определяются по параметрам $t_\beta = 0{,}8; t_\alpha = 1{,}6$ и соответственно равны 0,576 и 0,890.

Ответ: $P\ (19{,}5'' < X < 22{,}5'') = 0{,}733$.

Большинство ошибок измерений носит случайный характер и подчиняется нормальному закону распределения. Отклонение от этого закона чаще всего свидетельствует о действии систематических ошибок. Для установления такого отклонения используются особые параметры, получившие название асимметрии и эксцесса.

Асимметрию применяют для оценки симметричности распределения. Наиболее просто для оценки симметричности использовать центральный момент третьего порядка $\mu$.

Коэффициентом асимметрии $A_s$ называют отношение центрального момента третьего порядка к кубу среднего квадратического отклонения

$$A_s = \frac{\mu_3}{\sigma^3}. \tag{3.48}$$

Если распределение симметрично и подчиняется нормальному закону, то $A_s = 0$. Дифференциальная кривая нормального распределения имеет характерную куполообразную форму. Если дифференциальная кривая распределения оказывается скошенной в ту или иную сторону от математического ожидания, то такое распределение называют асимметричным. В первом случае, когда мода предшествует медиане, то асимметрия положительна, т.е. $A_s > 0$, а во втором случае, когда медиана предшествует моде, то асимметрия отрицательна, т.е. $A_s < 0$ (рис. 11). Коэффициент асимметрии не имеет ни верхней, ни нижней границы, что снижает его ценность как меры асимметрии. Практически коэффициент асимметрии редко бывает по своей величине большим, а для умеренно асимметричных рядов он обычно меньше единицы.

При вычислении эмпирического коэффициента асимметрии используют эмпирические значения центрального момента третьего порядка и среднего квадратического отклонения (среднюю квадратическую ошибку). Сходимость эмпирической асимметрии к теоретической возрастает при увеличении числа измерений $n$. При малом числе измерений стандарт асимметрии

$$\sigma_A = \sqrt{\frac{6(n-1)}{(n+1)(n+3)}}. \tag{3.49}$$

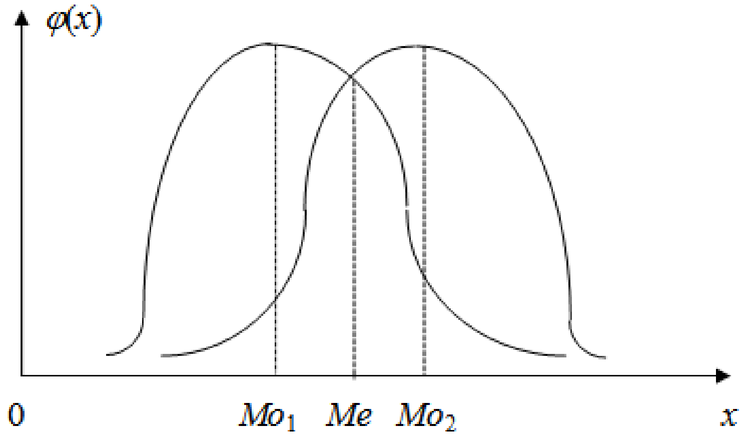


**Рис. 11**

Если для обработки используется выборка большого объема, то

$$\sigma_A \approx \sqrt{\frac{6}{n+3}} \approx \sqrt{\frac{6}{n}}. \tag{3.50}$$

Асимметрию считают существенной тогда, когда имеет место следующее условие:

$$|A_s| > t \cdot \sigma_A, \tag{3.51}$$

где $t$ - определяется для заданного уровня значимости $q$ согласно табл.1 приложения. Если выполняется условие (3.51), то асимметрией можно пренебречь.

Эксцессом или коэффициентом крутости $E_s$ называется уменьшенное на три единицы отношение центрального момента четвертого порядка к четвертой степени среднего квадратического отклонения

$$E_s = \frac{\mu_4}{\sigma^4} - 3 \ . \qquad (3.52)$$

За стандартное значение эксцесса принимают $E_s = 0$.

Дифференциальная кривая имеет форму, показанную на рис. 3. Кривые, у которых $E_s < 0$, по сравнению с нормальной имеют более плоскую вершину и называются плосковершинными. Кривые, у которых $E_s > 0$, более крутые, имеют более острую вершину и называются островершинными (рис. 12).

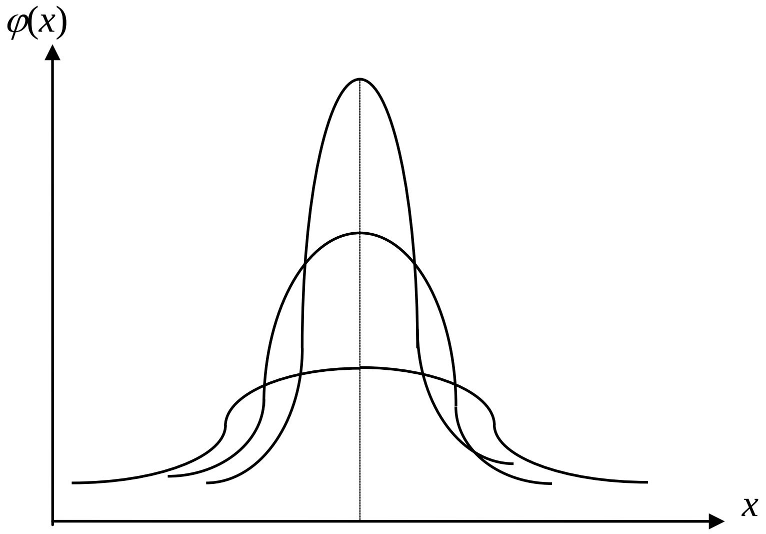


Рис. 12

Стандарт эмпирического эксцесса получаем

$$\sigma_E = \sqrt{\frac{24n(n-2)(n-3)}{(n-1)^2(n+3)(n+5)}} \approx \sqrt{\frac{24(n-3)}{(n+3)(n+5)}} \approx \sqrt{\frac{24}{n}} \, . \qquad (3.53)$$

Эмпирический эксцесс считается существенным, если он значительно превышает стандарт $\sigma_E$ , т.е. выполняется неравенство

$$|\,E_s\,| > 3\sigma_E \, . \qquad (3.54)$$

### 3.6. Распределение Пуассона

Применение нормального распределения при изучении рядов распределения случайной величины не охватывает всего многообразия распределений. Так, например, биномиальное распределение приближается к нормальному тем лучше, чем больше $n$ и чем меньше $p$ или $q$. Однако это не имеет места, если наряду с увеличением одна из величин $p$ или $q$ стремится к нулю. В этом случае распределение в пределе дает распределение Пуассона.

Преобразуем выражение для вероятности (3.1), приняв, что при возрастании $n$ сохраняется равенство $M(x) = a = np$

$$P_{k,n} = \frac{n(n-1)(n-2)\dots(n-k+1)(n-k)\dots 2\cdot 1}{k!(n-k)} p^k q^{n-k} =$$

$$= \frac{1}{k!} n(n-1)(n-2)\ .\ .\ .\ (n-k+1)\left(\frac{a}{n}\right)^k \left(1-\frac{a}{n}\right)^{n-k} =$$

$$= \frac{1}{k!}\cdot\frac{n}{n}\cdot\frac{n-1}{n}\cdot\frac{n-2}{n}\dots\frac{n-k+1}{n} a^k \left(1-\frac{a}{n}\right)^n \left(1-\frac{a}{n}\right)^{-k}.$$

Найдем пределы для каждого члена полученного выражения

$$\lim_{n\to\infty}\frac{n}{n} = 1;\quad \lim_{n\to\infty}\frac{\mathrm{n - 1}}{n} = \lim_{n\to\infty}\left(1-\frac{1}{n}\right) = 1;\quad \lim_{n\to\infty}\frac{\mathrm{n - 2}}{n} = \lim_{n\to\infty}\left(1-\frac{2}{n}\right) = 1;$$

$$\lim_{n\to\infty}\frac{\mathrm{n - k + 1}}{n} = \lim_{n\to\infty}\left(1-\frac{k-1}{n}\right) = 1;\quad \lim_{n\to\infty}\left(1-\frac{a}{n}\right)^{\mathrm{-k}} = 1.$$

Обозначив $-\frac{a}{n} = \frac{1}{x}$, установим предел для выражения $\left(1-\frac{a}{n}\right)^n$

$$\lim_{n\to\infty}\left(1-\frac{a}{n}\right)^n = \lim_{x\to\infty}\left(1+\frac{1}{x}\right)^{-ax} = \lim_{x\to\infty}\left[\left(1+\frac{1}{x}\right)^x\right]^{-a} = e^{-a},$$

так как $\lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x = e = 2{,}71828\dots$ . Следовательно, при $n \to \infty$ и $p$ или $q \to \infty$ получим

$$P_k = \frac{a^k e^{-a}}{k!}. \qquad (3.55)$$

Формула (3.55) выражает распределение Пуассона. Закон Пуассона описывает число появления событий $k$, происходящих за одинаковые промежутки времени при условии, что события происходят независимо друг от друга с постоянной интенсивностью. Вероятность $p$ появления события в каждом отдельном случае постоянна и мала. Поэтому закон Пуассона получил название закона распределения редких явлений.

Распределение Пуассона широко применяется при контроле производства и качества продукции. Кроме того, данный закон используют для оценки месторождений полезных ископаемых с редко встречающимися элементами и минералами.

**Пример 16.** Пусть в 1 м$^3$ песков россыпи в среднем содержится 3 частицы золота. Определить распределение золота в 1 м$^3$. Объем участка, для которого находится распределение, равен 10 м$^3$.

**Решение.** Количество частиц на всем участке составит $n = 10 \cdot 3 = 30$. Вероятность попадания одной частицы золота на участок объемом в 1м$^3$ составит 1м$^3$/10м$^3$ = 0,1. Имея ввиду случайное размещение золота по участку в 1м$^3$ может встретиться $k = 0, 1, 2, \ldots, 30$ частиц. Вероятность событий определяется по формуле (3.55). В табл. 6 приводятся вычисленные значения вероятностей распределения частиц золота в 1м$^3$ при $n = 30$ и числе частиц $k = 0, 1, \ldots, 8$.

Таблица 6

| $k$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|---|
| $P_k$ | 0,04 | 0,14 | 0,23 | 0,24 | 0,18 | 0,10 | 0,05 | 0,02 | 0,01 |

Вычислим математическое ожидание согласно формуле (3.11)

$$M(x) = a = np = 30 \cdot 0{,}1 = 3.$$

Соответственно эмпирические значения дисперсии и стандарта, вычисленные по формулам (3.16) и (3.17), будут равны:

$$D(x) = npq = 30 \cdot 0{,}1 \cdot 0{,}9 = 2{,}7\ ;$$

$$\sigma = \sqrt{D(x)} = \sqrt{2{,}7} = 1{,}64\,.$$

Как видно из табл. 6 наилучшая возможность встретить в 1м$^3$ три частицы золота, а уже семь и более частиц вероятность падает до нуля.

Ответ: $M(x) = 3$ шт. ; $D(x) = 2{,}7$ ; $\sigma = 1{,}64$ шт.

### 3.7. Гамма- и бета-распределение

Одним из часто встречающихся типов распределений являются гамма- и бета-распределения. Гамма-функция для какого-либо положительного числа $p$ представляет собой интеграл

$$\partial(p) = \int_0^\infty x^{p-1} e^{-x} dx\,. \tag{3.56}$$

Интеграл существует при любом $p > 0$.

Между $\Gamma(p + 1)$ и $\Gamma(p)$ существует очень простое соотношение, которое можно получить следующим образом. Так как согласно (3.56) имеем

$$\Gamma(p+1) = \int_0^\infty x^p e^{-x} dx. \qquad (3.57)$$

При интегрировании получаем

$$\int_0^\infty x^p e^{-x} dx = \left[\; -x^p e^{-x} \;\right]_0^\infty + p\int_0^\infty x^{p-1} e^{-x} dx.$$

При $p > 0$ выражение в скобках обращается в нуль при любом пределе интегрирования. Следовательно,

$$\int_0^\infty x^p e^{-x} dx = p\int_0^\infty x^{p-1} e^{-x} dx ,$$

т.е. имеет место равенство

$$\Gamma(p + 1) = p\Gamma(p) , \qquad (3.58)$$

которое выражает основное свойство рассматриваемой функции и получило название функционального уравнения $\Gamma$-функции.

Распределение непрерывной случайной величины $x$ с плотностью вероятности

$$\varphi(x, p) = \frac{x^{p-1} e^{-x}}{\Gamma(p)} \qquad (3.59)$$

называется гамма-распределением $\gamma(p)$

$$F(x, p) = \frac{1}{\Gamma(p)} \int_0^\infty x^{p-1} e^{-x} dx . \qquad (3.60)$$

График функции $y = x^{p-1} e^{-x}$ приведен на рис. 13.

Моменты $\gamma(p)$ - распределения относительно нуля равны

$$\nu_s = M(x^s) = \frac{1}{\Gamma(p)}\int_0^\infty x^{p+s-1} e^{-x} dx = \frac{\Gamma(p+s)}{\Gamma(p-1)!} = p(p+1)\; .\; .\; .\; (p+s-1). \qquad (3.61)$$

Откуда имеем

$$\nu_1 = p; \quad \nu_2 = p(p+1).$$

Дисперсия в этом случае составит

$$\sigma^2 = p(p+1) - p^2 = p .$$

Аналогично определяют центральные моменты третьего и четвертого порядков

$$\mu_3 = 2p\,; \qquad \mu_4 = 3p\,(p+2)\,.$$

Асимметрия и эксцесс в случае $\gamma(p)$ – распределения составят:

$$A_s = \frac{2}{\sqrt{p}}, \qquad E_s = \frac{6}{\sqrt{p}}. \qquad (3.62)$$

Как видно из равенств (3.62) асимметрия и эксцесс зависят только от параметра $p$.

Характеристическая функция $\gamma(p)$ – распределения

$$\varphi(t) = \frac{1}{\partial(p)} \int_0^\infty x^{p-1} e^{-(1-it)x} dx\,.$$

Применяя подстановку $z = (1 - it)^{-p}$, получим

$$\varphi(t) = \frac{1}{\Gamma(p)} \int_0^\infty \left(\frac{z}{1-it}\right)^{p-1} e^{-z} \frac{dz}{1-it} = \frac{1}{\Gamma(p)} \frac{1}{(1-it)^p} \int_0^\infty z^{p-1} e^{-z} dz,$$

т.е.

$$\varphi(t) = (1-it)^{-p}. \qquad (3.63)$$

Следовательно, если случайная величина $X$ подчиняется нормальному закону распределения, имея при этом математическое ожидание $M(x) = 0$ и дисперсию $\sigma^2$, величина

$$u = \frac{1}{2} \cdot \frac{x^2}{\sigma^2}$$

имеет гамма-распределение с параметром 1/2.

Характеристическая функция найденного $\gamma(1/2)$ – распределения соответственно окажется

$$\varphi(t) = (1-it)^{-1/2}\,. \qquad (3.64)$$

С помощью $\Gamma$ - функции можно выразить различные виды определенных интегралов. Еще большую роль играет интеграл, получивший название $B$ - функции

$$B(p,q)=\int_0^1 x^{p-1}(1-x)^{q-1}\,dx \qquad (p>0,\ q>0). \qquad (3.65)$$

Если $1-x=\dfrac{1}{1+t}$ , равенство (3.65) примет вид

$$B(p,q)=\int_0^\infty \frac{t^{p-1}}{(1+t)^{p+q}}\,dt. \qquad (3.66)$$

И при $q = 1$получим

$$B(p,1)=\int_0^1 x^{p-1}dx=\frac{1}{p}. \qquad (3.67)$$

В частности, $B(1,1) = 1$.

При $p > 1$

$$B(p,q)=\frac{q-1}{p}B(p=1,q-1). \qquad (3.68)$$

Это равенство описывает свойство $B$-функции. Значения $B$-функции могут быть определены через $\Gamma$-функцию

$$B(p,q)=\frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}. \qquad (3.69)$$

Бета-распределения распадаются на два рода. Бета-распределение первого рода $\beta_1(p,q)$ представляет распределение случайной величины $X$ с плотностью вероятности

$$\frac{x^{p-1}(1-x)^{q-1}}{B(p,q)}$$

на интервале от 0 до 1. Следовательно,

$$\beta_1(p,q)=\frac{1}{B(p,q)}\int_0^1 t^{p-1}(1-t)^{q-1}dt\ . \qquad (3.70)$$

График функции $y=x^{p-1}(1-x)^{q-1}$ приводится на рис. 14. Площадь под кривой распределения выражает бета-функцию.

Центральный момент $s$-го порядка $\beta_1(p,q)$ - распределения

$$\mu_s=\frac{1}{B(p,q)}\int_0^1 x^{p+s-1}(1-x)^{q-1}dx=\frac{B(p+s,q)}{B(p,q)}=\frac{p(p+1)...(p+s-1)}{(p+q)(p+q+1)...(p+q+s-1)}. \qquad (3.71)$$

Откуда

$$\mu_1 = \frac{p}{p+q}, \qquad \mu_2 = \frac{p(p+1)}{(p+q)(p+q+1)}.$$

Дисперсия $\beta_1(p,q)$ - распределения

$$\sigma^2 = \frac{pq}{(p+q)(p+q+1)}. \qquad (3.72)$$

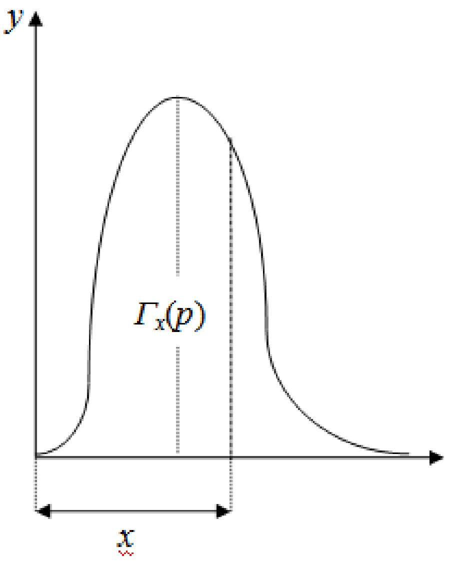


Рис. 13

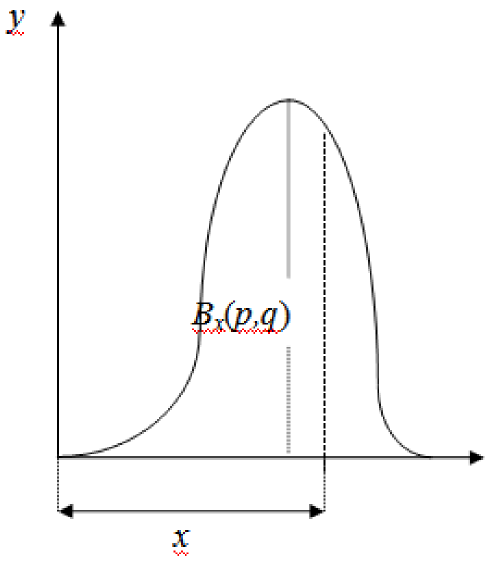


Рис. 14

Бета-распределение второго рода $\beta_2(p,q)$ с положительными параметрами $p$ и $q$ представляет собой распределение непрерывной случайной величины $X$ с плотностью вероятности

$$\frac{x^{p-1}}{B(p,q)(1+x)^{p+q}}.$$

Центральный момент $s$-го порядка $\beta_2(p,q)$ - распределения при $s < q$

$$\mu_s = \frac{1}{B(p,q)}\int_0^\infty \frac{x^{p+s-1}}{(1+x)^{p+q}}dx = \frac{p(p+1)\ldots(p+s-1)}{(q-1)(q-2)\ldots(q-s)}. \qquad (3.73)$$

В частности

$$\mu_1 = \frac{p}{q-1}, \qquad \mu_2 = \frac{p(p+q-1)}{(q-1)(q-2)}.$$

Дисперсия $\beta_2(p,q)$ - распределения

$$\sigma^2 = \frac{p(p+q-1)}{(q-1)^2(q-2)}. \qquad (3.74)$$

## 3.8. Логнормальное распределение

Логарифмически нормальным распределением называется распределение положительной величины $X$, логарифм которой $Y = \ln X$ подчиняется закону нормального распределения. В этом случае его плотность распределения случайной величины $Y$ находится согласно

$$\varphi(y) = \frac{1}{\sigma_y\sqrt{2\pi}} e^{-[y-M(y)]^2/2\sigma_y^2}. \quad (3.75)$$

Математическое ожидание такого распределения

$$M(y) = \ln M(x) - \frac{\sigma_y^2}{2} = \ln M(x) - \frac{1}{2}\ln\left\{\frac{\sigma_x^2}{[M(x)]^2} + 1\right\}. \quad (3.76)$$

Дисперсия

$$\sigma_y^2 = \ln\left\{\frac{\sigma_x^2}{[M(x)]^2} + 1\right\}. \quad (3.77)$$

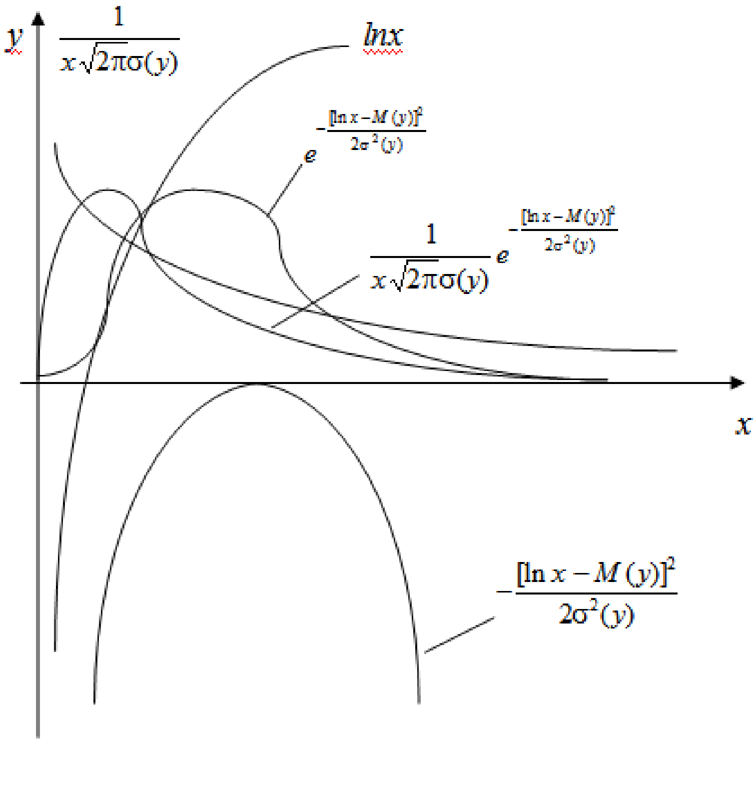


Рис. 15

В равенствах (3.76) и (3.77) математическое ожидание и дисперсия случайной величины $X$, подчиняющейся логнормальному распределению, равны соответственно:

$$M(x)=\int_0^{\infty} x\cdot\varphi(x)\,dx = e^{\frac{\sigma_y^2+M(y)}{2}}; \qquad (3.78)$$

$$\sigma_y^2=\int_0^{\infty}[x-M(x)]^2\varphi(x)\,dx = e^{\sigma_y^2+2M(y)}\left[e^{\sigma_y^2}-1\right]. \qquad (3.79)$$

Таким образом, плотность логнормально распределенной величины определится

$$\varphi(x)=\frac{1}{x\sqrt{2\pi\ln\left\{\frac{\sigma_x^2}{[M(x)]^2}+1\right\}}}e^{-\frac{\ln X-\ln M(x)+\frac{1}{2}\ln\left\{\frac{\sigma_x^2}{[M(x)]^2}+1\right\}}{2\ln\left\{\frac{\sigma_x^2}{[M(x)]^2}+1\right\}}}. \qquad (3.80)$$

На рис. 15 представлены вид кривой логнормального распределения и ее составляющие.

Логнормальным распределением можно аппроксимировать эмпирические распределения случайной величины, имеющей ярко выраженную правостороннюю асимметрию.

# 4. Основы теории ошибок измерений

## 4.1. Общие сведения об измерениях

Численное значение любой физической величины получается в результате ее измерения. Под измерением следует понимать определение численного значения физической величины с помощью специальных технических средств, или это есть процесс сравнения какой-либо величины с другой ей однородной величиной, принятой за единицу. Измеряемая величина и единица измерения не соизмеримы друг с другом, что приводит к возникновению ошибок.

В процессе измерения участвуют следующие элементы: объект измерения; наблюдатель; инструмент; внешняя среда. Все это образует условия измерения, которые и являются источниками возникновения ошибок.

Любое измерение, как бы оно тщательно не выполнялось, сопровождается ошибкой, численно равной разности между результатом измерения и истинным значением измеряемой величины. Это значение можно назвать истинной ошибкой измерения

$$\Delta_i = x_i - X, \tag{4.1}$$

где $x_i$ - результаты измерения;

$X$ - истинное значение измеряемой величины.

Значения большинства величин получают как в результате непосредственных измерений, так и с помощью вычислений, т.е. прямым и косвенным способами.

Объектами измерений могут быть как однородные, так и неоднородные величины. Например, в триангуляции измеряются однородные величины (углы), а в полигонометрии - неоднородные (углы и длины линий). Вместе с тем основные определяемые величины - координаты пунктов - в том и другом случае являются однородными.

Различают необходимые и избыточные измеренные величины. Необходимыми являются измеренные величины, достаточные для однозначного определения значений искомых величин. Измерения, выполненные сверх необходимых, будут избыточными. Они играют в теории ошибок важную роль, так как позволяют:

- контролировать качество выполненных работ, выявляя результаты с грубыми ошибками;
- оценить точность выполненных измерений;
- определять наиболее надежные значения измеряемых величин.

По отношению к точности результаты измерений можно подразделить на равноточные и неравноточные. Равноточными являются такие измерения, которые выполняются

а) одним и тем же инструментом или разными инструментами, но с одинаковой точностью;

б) одними и теми же методами или способами;

в) в одних и тех же условиях.

Если какое-либо из перечисленных пунктов не соблюдается, то измерения относятся к неравноточным.

Особым качеством результатов измерений является их взаимная независимость. Наиболее полная независимость достигается в том случае, если измерения произведены в различных условиях. В своей же массе большинство результатов измерений можно отнести к зависимым. Однако практика геодезических работ позволяет пренебречь в пределах точности измерений возникающими в этом случае зависимостями.

## 4.2. Классификация ошибок измерений

Причинами возникновения ошибок в результате измерений являются:

1) изменение величины или состояния объекта в процессе измерения;

2) личные ошибки наблюдателя;

3) инструментальные ошибки измерений;

4) влияние внешней среды.

Возникшие при этом ошибки можно подразделить на три вида: грубые, систематические и случайные.

К грубым ошибкам относятся промахи, просчеты при измерениях, а также ошибки, превосходящие допустимые значения. Грубые ошибки выявляются повторными измерениями и исключаются из результатов. Следовательно, задача сводится к организации контроля наблюдений.

Если среднее арифметическое из ошибок равноточных измерений стремится к некоторому пределу, отличному от нуля, при увеличении числа измерений до бесконечности, то такие ошибки называются систематическими. К систематическим ошибкам относятся составляющие общей ошибки измерений, которые постоянны или закономерно изменяются при повторных измерениях одной и той же величины. Систематические ошибки по характеру действия классифицируют: на сохраняющие знак и величину; меняющиеся по величине, но сохраняющие знак (одностороннее действующие); изменяющиеся по какому-либо функциональному закону. Анализ причин возникновения систематических ошибок позволяет частично или полностью исключить их из результатов измерений. Величина систематических ошибок зависит от методики измерений.

Если среднее арифметическое из ошибок равноточных измерений одной и той же величины стремится к нулю при увеличении числа измерений до бесконечности, то такие ошибки называются случайными. Случайная ошибка является той частью общей ошибки, которая меняется при повторных измерениях одной и той же величины. Случайные ошибки по величине чаще всего больше систематических, но из-за взаимных компенсаций их влияние на окончательный результат может быть слабее.

### 4.3. Свойства случайных ошибок

Случайные ошибки измерений подчиняются нормальному закону распределения с плотностью

$$\varphi(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}}, \tag{4.2}$$

где $t$ - текущее значение случайной ошибки измерений;

$\sigma$ - среднее квадратическое отклонение (стандарт) генеральной совокупности этих ошибок.

Из анализа функции (4.2) имеем:

1) функция достигает максимальных значений при $t = 0$;

2) функция четная, т.е. $f(t) = f(-t)$; отсюда следует, что кривая симметрична относительно значения $t = 0$;

3) математическое ожидание случайных ошибок измерений определится согласно

$$M(t) = \frac{1}{\sigma\sqrt{2\pi}} \int t e^{-\frac{t^2}{2\sigma^2}} dt = 0; \tag{4.3}$$

4) дисперсия случайных ошибок измерений запишется в виде

$$D(t)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty} t^2 e^{-\frac{t^2}{2\sigma^2}}dt=\sigma^2. \tag{4.4}$$

Отсюда среднее квадратическое отклонение

$$\sigma=\sqrt{D(t)}. \tag{4.5}$$

Равенство (4.2) носит название уравнения кривой Гаусса. Свойства ее совпадают с кривой нормального распределения. Отсюда получим следующие свойства случайных ошибок измерений.

1. При данных условиях измерений случайные ошибки не могут превзойти по абсолютной величине определенного предела, т.е.

$$|\Delta| = \Delta_{np} = 3\sigma . \tag{4.6}$$

2. Малые по абсолютной величине ошибки встречаются чем большие.

3. Положительные ошибки появляются так же часто, как и равные им по абсолютной величине отрицательные ошибки, т.е.

$$p\ (\Delta < 0) = p\ (\ \Delta > 0) = 0{,}5. \tag{4.7}$$

4. Среднее арифметическое из случайных ошибок измерений одной и той же величины стремится к нулю при неограниченном возрастании числа измерений, т.е.

$$\lim_{n\to\infty}\frac{[\Delta]}{n}=0. \tag{4.8}$$

Если $\lim\limits_{n\to\infty}\frac{[\Delta]}{n}\neq 0,$ то в этом случае систематическое влияние полностью не исключено.

### 4.4. Принцип арифметической средины

Пусть даны результаты равноточных измерений одной величины $x_1, x_2, \ldots , x_n$ . При этом следует иметь в виду, что истинное значение $X$ измеряемой величины известно.

Составим ряд истинных ошибок результатов измерений согласно равенству (4.1)

$$\begin{aligned} \Delta_1 &= x_1 - X; \\ \Delta_2 &= x_2 - X; \\ &\ldots\ldots\ldots\ldots\ldots \\ \Delta_n &= x_n - X. \end{aligned} \tag{4.9}$$

Равенства (4.9) почленно сложим, в результате получим следующее значение:

$$[\Delta] = [x] - nX \; .$$

Полученное равенство разделим на $n$, вследствие чего будем иметь

$$\frac{[\Delta]}{n} = \frac{[x]}{n} - X. \tag{4.10}$$

Введем обозначение $\bar{x} = \frac{[x]}{n}$ – среднее арифметическое значение измеряемой величины. Тогда

$$\frac{[\Delta]}{n} = \bar{x} - X \; . \tag{4.11}$$

На основании четвертого свойства случайных ошибок можно утверждать, что $\bar{x} \to X$ при $n \to \infty$, т.е. среднее арифметическое из результатов равноточных измерений стремится к истинному значению этой величины при неограниченном возрастании числа измерений.

Среднее арифметическое из данного ряда равноточных измерений принимается за наиболее надежное значение и при конечном числе измерений.

## 4.5. Критерии точности результатов измерений

Для оценки точности отдельного измерения некоторой величины необходимо определить возможные отклонения результатов измерения этой величины от ее истинного значения. Суждение о точности выполненных измерений можно получить по степени различия результатов измерений в ряду: чем больше разбросаны результаты в ряду, тем сильнее рассеивание (дисперсия) ряда, тем менее точны измерения.

Возьмем два ряда измерений, выполненных в различных условиях, истинные ошибки которых будут:

$$-3, +2, +4, -2, -1, 0, +4, -3, +2, -3;$$

$$0, -1, +7, +2, -1, -1, -8, 0, +3, -1.$$

Сопоставляя ряды, видим, что разброс ошибок во втором ряду больше, чем в первом. Следовательно, условия измерений в первом случае более благоприятны.

Найдем среднюю ошибку $\vartheta$, которая определяется по абсолютным значениям ошибок ряда

$$\vartheta = \frac{[|\Delta|]}{n}. \tag{4.12}$$

В результате вычислений получим следующие значения средних ошибок $\vartheta_1 = 2{,}4$; $\vartheta_2 = 2{,}4$. Как видим, оба значения одинаковы. Это объясняется тем, что $\vartheta$ недостаточно чувствительно к наличию крупных ошибок.

В качестве меры точности Гаусс предложил принять среднее квадратическое отклонение или среднюю квадратическую ошибку согласно равенству (4.5)

$$m = \sqrt{\frac{[\Delta\Delta]}{n}}. \tag{4.13}$$

Вся теория математической обработки результатов измерений построена на использовании средней квадратической ошибки, которая обладает рядом преимуществ:

1) при вычислении нет необходимости учитывать знак величины $\Delta_i$ отдельных ошибок;

2) большие по абсолютному значению величины $\Delta_i$ после возведения в степень увеличивают ее численные значения, создавая своего рода “запас прочности”;

3) по формуле (4.13) значение средней квадратической ошибки $m$ получается со знаком «±», что соответствует природе случайных ошибок.

Отличием средней квадратической ошибки $m$ от стандарта $\sigma$ является то, что средняя квадратическая ошибка - величина эмпирическая, вычисляемая из ограниченного числа измерений, а стандарт - величина постоянная, характеризующая бесконечную совокупность данного вида измерений. Это – величина теоретическая. При $n \to \infty$ , $m \to \sigma$.

Для нашего примера $m_1 = \sqrt{\frac{78}{10}} = 2{,}7$; $m_2 = \sqrt{\frac{130}{10}} = 3{,}6$.

Из сравнения $m_1$ и $m_2$ следует, что условия измерений для первого ряда лучше, чем для второго ряда.

Так как $\sigma$ всегда остается неизвестным, то приходится пользоваться его “оценкой” $m$, т.е. $m \approx \sigma$, которое в свою очередь определяется с ошибкой $m_m$

$$m_m = \frac{m}{\sqrt{2n}}. \tag{4.14}$$

Существует связь между средней и средней квадратической ошибками при $n \to \infty$

$$\vartheta = 0{,}798m \approx \frac{4}{5}m. \tag{4.15}$$

Иногда в качестве меры точности используют вероятную ошибку $r$, которая делит ряд ошибок, расположенных в порядке возрастания их абсолютных значений, пополам. При $n \to \infty$ вероятная ошибка $r$ связана со средней квадратической ошибкой зависимостью

$$r = 0{,}674m \approx \frac{2}{3}m. \tag{4.16}$$

По величине средней квадратической ошибки, определяющей условия измерения, можно установить предельную ошибку

$$\Delta_{\text{пр}} = k \cdot m. \tag{4.17}$$

Ошибки, большие предельной, считаются грубыми. Ранее было установлено:

$$P(|\Delta| < \sigma) = 0{,}6827;\ P(|\Delta| < 2\sigma) = 0{,}9545;$$

$$P(|\Delta| < 2{,}5\sigma) = 0{,}9876;\text{P}(|\Delta| < 3\sigma) = 0{,}9973,$$

т.е. случайная ошибка измерения может превосходить среднюю квадратическую ошибку $|m|$ в 32 случаях из 100, удвоенную $|2m|$ – в 5 случаях из 100, утроенную $|3m|$ – в 3 случаях из 1000. Следовательно, согласно “правилу трех сигм” для различных геодезических работ величину $k$ принимают равную 3, т.е.

$$\Delta_{np} = 3m. \tag{4.18}$$

При вычислении допустимых невязок $k = 2$ и тогда

$$\Delta_{np} = 2m. \tag{4.19}$$

Такие ошибки, как средняя $\vartheta$, средняя квадратическая $m$, вероятная $r$, истинная $\Delta$ и предельная $\Delta_{np}$, являются абсолютными ошибками. Они всегда выражены в единицах измеряемой величины, т.е. имеют одинаковую с измеряемой величиной размерность.

Часто возникают случаи, когда разные по величине объекты измеряют с одинаковыми абсолютными ошибками. Например, средняя квадратическая ошибка измерения линий длиной: $l_1$ = 100 м и $l_2$ = 1000 м, составила $m$ = 5 см. Возникает вопрос: какая же линия измерялась точнее ? Чтобы избежать неопределенности, точность измерений ряда величин оценивают в виде отношения абсолютной ошибки к значению измеряемой величины. Полученное отношение называется относительной ошибкой, которую обычно выражают дробью с числителем, равным единице.

Наименование абсолютной ошибки определяет и название соответствующей ей относительной ошибки измерения.

Пусть $x$ – результат измерения некоторой величины. Тогда

$\frac{m}{x} = \frac{1}{N_1}$ – средняя квадратическая относительная ошибка;

$\frac{\vartheta}{x} = \frac{1}{N_2}$ – средняя относительная ошибка;

$\frac{r}{x} = \frac{1}{N_3}$ – вероятная относительная ошибка;

$$\frac{\Delta}{x} = \frac{1}{N_4}$$ – истинная относительная ошибка;

$$\frac{\Delta_{пр}}{x} = \frac{1}{N_5}$$ – предельная относительная ошибка.

Знаменатель $N$ относительной ошибки необходимо округлять до двух значащих цифр с нулями.

**Пример 17**. $m_x = 0{,}3$ м; $x = 152{,}0$ м; $\frac{m_x}{x} = \frac{1}{510}$.

$m_x = 0{,}25$ м; $x = 643{,}00$ м; $\frac{m_x}{x} = \frac{1}{2600}$.

$m_x = 0{,}033$ м; $x = 795{,}000$ м; $\frac{m_x}{x} = \frac{1}{24000}$.

Ответ: $\frac{m_x}{x} = \frac{1}{510}$; $\frac{m_x}{x} = \frac{1}{2600}$; $\frac{m_x}{x} = \frac{1}{24000}$.

Как видно из примера, чем больше знаменатель дроби, тем точнее выполнены измерения.

### 4.6. Ошибки функций измеренных величин

В общем случае найдем среднюю квадратическую ошибку $M_F$ некоторой функции вида

$$F = f(x_1, x_2, \ldots, x_n), \tag{4.20}$$

где $x$ – коррелированные аргументы, связанные между собой зависимостями и полученные из наблюдений.

При этом будем иметь в виду, что средние квадратические ошибки результатов измерений $m_1$, $m_2$, ... , $m_n$ известны. Кроме этого, предположим, что $X_1$, $X_2$, ... , $X_n$ – истинные значения аргументов. Каждая величина измерялась $k$ раз, т. е.

для $X_1$ : $x_{11}, x_{12}, \ldots, x_{1k}$;

для $X_2$ : $x_{21}, x_{22}, \ldots, x_{2k}$;

. . . . . . . . . . . . . . . . .

для $X_n$ : $x_{n1}, x_{n2}, \ldots, x_{nk}$.

Для каждой серии наблюдений получим значения оцениваемой функции (4.20):

$$
\begin{aligned}
&F_1 = f(x_{11}, x_{12}, \ldots, x_{1n});\\
&F_2 = f(x_{21}, x_{22}, \ldots, x_{2n});\\
&\ldots\ldots\ldots\ldots\ldots\ldots\\
&F_k = f(x_{k1}, x_{k2}, \ldots, x_{kn}).
\end{aligned}
\tag{4.21}
$$

Используя равенство (4.1), получим вместо уравнений (4.21) следующие зависимости:

$$
\begin{aligned}
&F_1 = f(X_1 + \Delta_{11}, X_2 + \Delta_{21}, \ldots, X_n + \Delta_{n1});\\
&F_2 = f(X_1 + \Delta_{12}, X_2 + \Delta_{22}, \ldots, X_n + \Delta_{n2});\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&F_k = f(X_1 + \Delta_{1k}, X_2 + \Delta_{2k}, \ldots, X_n + \Delta_{nk}).
\end{aligned}
\tag{4.22}
$$

Разложив равенства (4.22) в ряд Тейлора и ограничиваясь в виду малости ошибок измерений первыми степенями разложения, получим

$$
\begin{aligned}
&F_1 = f(X_1, X_2, \ldots, X_n) + \frac{\partial f}{\partial X_1}\Delta_{11} + \frac{\partial f}{\partial X_2}\Delta_{21} + \ldots + \frac{\partial f}{\partial X_n}\Delta_{n1};\\
&F_2 = f(X_1, X_2, \ldots, X_n) + \frac{\partial f}{\partial X_1}\Delta_{12} + \frac{\partial f}{\partial X_2}\Delta_{22} + \ldots + \frac{\partial f}{\partial X_n}\Delta_{n2};\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&F_k = f(X_1, X_2, \ldots, X_n) + \frac{\partial f}{\partial X_1}\Delta_{1k} + \frac{\partial f}{\partial X_2}\Delta_{2k} + \ldots + \frac{\partial f}{\partial X_n}\Delta_{nk}.
\end{aligned}
\tag{4.23}
$$

Учитывая, что истинная ошибка функции

$$\Delta_F = f(x_1, x_2, \ldots, x_n) - f(X_1, X_2, \ldots, X_n),$$

получим систему случайных ошибок функции

$$
\begin{aligned}
&\Delta_{F1} = \frac{\partial f}{\partial X_1}\Delta_{11} + \frac{\partial f}{\partial X_2}\Delta_{21} + \ldots + \frac{\partial f}{\partial X_n}\Delta_{n1};\\
&\Delta_{F2} = \frac{\partial f}{\partial X_1}\Delta_{12} + \frac{\partial f}{\partial X_2}\Delta_{22} + \ldots + \frac{\partial f}{\partial X_n}\Delta_{n2};\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&\Delta_{Fk} = \frac{\partial f}{\partial X_1}\Delta_{1k} + \frac{\partial f}{\partial X_2}\Delta_{2k} + \ldots + \frac{\partial f}{\partial X_n}\Delta_{nk}.
\end{aligned}
\tag{4.24}
$$

Возведя каждое равенство системы (4.24) в квадрат, просуммировав полученные значения и разделив их на $k$, придем к следующему уравнению:

$$\frac{[\Delta_F\Delta_F]}{k}=\left(\frac{\partial f}{\partial X_1}\right)^2\frac{[\Delta_1\Delta_1]}{k}+\left(\frac{\partial f}{\partial X_2}\right)^2\frac{[\Delta_2\Delta_2]}{k}+...+\left(\frac{\partial f}{\partial X_n}\right)^2\frac{[\Delta_n\Delta_n]}{k}+2\left(\frac{\partial f}{\partial X_1}\right)\left(\frac{\partial f}{\partial X_2}\right)\frac{[\Delta_1\Delta_2]}{k}+... \quad (4.25)$$

В математической статистике доказывается следующая теорема: “Среднее значение произведения случайных величин равно произведению средних значений сомножителей”

$$\frac{[\Delta_1\Delta_2]}{k}=\frac{[\Delta_1]}{k}\cdot\frac{[\Delta_2]}{k}. \quad (4.26)$$

Согласно четвертому свойству случайных ошибок каждое из сомножителей равенства (4.26) стремится к нулю при неограниченном возрастании числа измерений. Поэтому можно утверждать, что и

$$\lim_{n\to\infty}\frac{[\Delta_1\Delta_2]}{k}=0. \quad (4.27)$$

Перейдем к средним квадратическим ошибкам функции независимых величин, учитывая равенство (4.13), а также условие (4.27)

$$M_F^2=\left(\frac{\partial f}{\partial X_1}\right)^2 m_1^2+\left(\frac{\partial f}{\partial X_2}\right)^2 m_2^2+...+\left(\frac{\partial f}{\partial X_n}\right)^2 m_n^2, \quad (4.28)$$

где $m_l$ – средние квадратические ошибки измеренных величин.

Если аргументы зависимы (коррелированы), то для средних значений произведений ошибок этих величин при $k\to\infty$ справедливо равенство

$$\lim_{k\to\infty}\frac{[\Delta_i\Delta_j]}{k}=r_{ij}\cdot m_i\cdot m_j, \quad (4.29)$$

где $r_{ij}$ – коэффициент корреляции между зависимыми аргументами.

С учетом равенства (4.29) выражение (4.25) примет следующий вид

$$M_F^2 = \left(\frac{\partial f}{\partial X_1}\right)^2 m_1^2 + \left(\frac{\partial f}{\partial X_2}\right)^2 m_2^2 + \ldots + \left(\frac{\partial f}{\partial X_n}\right)^2 m_n^2 +$$
$$+ 2\left(\frac{\partial f}{\partial X_1}\right)\left(\frac{\partial f}{\partial X_2}\right) r_{12} m_1 m_2 + \ldots + 2\left(\frac{\partial f}{\partial X_{n-1}}\right)\left(\frac{\partial f}{\partial X_n}\right) r_{n-1,n} m_{n-1} m_n. \quad (4.30)$$

## 4.7. Типовые примеры

Вычислим средние квадратические ошибки функций различного вида.

1. Функция имеет следующий вид

$$F = k \cdot x, \quad (4.31)$$

где $x$ – величина, измеренная с ошибкой $m$;

$k$ – постоянный коэффициент.

Частная производная от функции (4.31) по измеренной величине

$$\frac{\partial F}{\partial x} = k.$$

Используя равенство (4.28), получим

$$M_F^2 = k^2 \cdot m^2,$$

откуда

$$M_F = k \cdot m. \quad (4.32)$$

Средняя квадратическая ошибка произведения непосредственно измеренной величины на постоянный коэффициент равна произведению этой постоянной на среднюю квадратическую ошибку измеренного аргумента.

2. Исходная функция представлена следующей зависимостью:

$$F = \pm k_1 x_1 \pm k_2 x_2 \pm \ldots \pm k_n x_n, \quad (4.33)$$

где $k_i$ – некоторые постоянные коэффициенты;

$x_i$ – независимые величины, измеренные со средними квадратическими ошибками $m_i$.

Находим частные производные функции по независимым аргументам и, подставляя в формулу (4.28), получаем

$$M_F^2 = k_1^2 m_1^2 + k_2^2 m_2^2 + \ldots + k_n^2 m_n^2. \tag{4.34}$$

Если $k_1 = k_2 = \ldots = k_n = 1$, то ошибка линейной функции составит

$$M_F = \sqrt{m_1^2 + m_2^2 + \ldots + m_n^2}, \tag{4.35}$$

т.е. средняя квадратическая ошибка суммы или разности независимо измеренных величин равна корню квадратному из суммы квадратов ошибок аргументов. Если при этом $m_1 = m_2 = \ldots = m_n = 1$, то

$$M_F = M_\Sigma = m\sqrt{n}. \tag{4.36}$$

Средняя квадратическая ошибка алгебраической суммы $n$ равноточно измеренных величин в корень квадратный из $n$ раз больше ошибки одного измерения.

Следовательно, накопление ошибок происходит пропорционально корню квадратному из числа измеряемых величин.

## 4.8. Средняя квадратическая ошибка простой арифметической средины

Дан ряд результатов равноточных измерений одной величины

$$x_1,\ x_2,\ \ldots,\ x.$$

Представим формулу для вычисления простой арифметической средины в виде

$$\bar{x} = \frac{1}{n}x_1 + \frac{1}{n}x_2 + \ldots + \frac{1}{n}x_n. \tag{4.37}$$

Согласно (4.28) будем иметь

$$M^2 = \frac{1}{n^2}m_1^2 + \frac{1}{n^2}m_2^2 + \ldots + \frac{1}{n^2}m_n^2, \tag{4.38}$$

где $m_1 = m_2 = \ldots = m_\text{n} = m$ . Следовательно, равенство (4.57) примет следующий вид

$$M^2 = \frac{m^2}{n}$$

и окончательно запишем

$$M = \frac{m}{\sqrt{n}}. \qquad (4.39)$$

Итак, средняя квадратическая ошибка простой арифметической средины в корень квадратный раз из числа измерений меньше средней квадратической ошибки одного измерения.

### 4.9. Вывод формулы Бесселя

Выполнив $n$ повторных измерений одной и той же величины, вычислим отклонения результатов этих измерений от арифметической средины

$$v_i = x_i - \bar{x}.$$

Необходимо получить формулу для оценки точности результата измерения через отклонения от арифметической средины, но предварительно рассмотрим свойства отклонений $v_i$.

**Первое свойство**. Алгебраическая сумма отклонений результатов равноточных измерений одной и той же величины от простой арифметической средины равна нулю при любом числе измерений.

Возьмем ряд отклонений от арифметической средины

$$\begin{aligned} v_1 &= x_1 - \bar{x}; \\ v_2 &= x_2 - \bar{x}; \\ &\ldots\ldots\ldots\ldots \\ v_n &= x_n - \bar{x}. \end{aligned} \qquad (4.40)$$

Суммируя левую и правую части равенства (4.40), получим

$$[v] = [x] - n\bar{x}.$$

Полученное равенство разделим на $n$

$$\frac{[v]}{n} = \frac{[x]}{n} - \bar{x}. \qquad (4.41)$$

Поскольку правая часть равенства (4.41) равна нулю, то и $[v] = 0$.

**Второе свойство**. Сумма квадратов отклонений результатов равноточных измерений от простой арифметической средины меньше суммы квадратов отклонений этих же результатов от любой другой величины, не равной простой арифметической средины, т.е. если $x' \neq \bar{x}$, то

$$[\,vv\,] < [\,\varepsilon\varepsilon\,], \tag{4.42}$$

где

$$v_i = x_i - \bar{x}\,, \qquad \varepsilon_i = x_i - x'\,.$$

Докажем второе свойство отклонений $v_i$ алгебраическим путем. Вычтем $v_i$ из $\varepsilon_i$ по частям из правого равенства левое

$$\varepsilon_i - v_i = x_i - x' - x_i + \bar{x} = \bar{x} - x' = c\,. \tag{4.43}$$

Для $n$ числа наблюдений получим ряд отклонений $\varepsilon_i$

$$\begin{aligned} \varepsilon_1 &= v_1 + c\,, \\ \varepsilon_2 &= v_2 + c\,, \\ &\dots\dots\dots \\ \varepsilon_n &= v_n + c\,. \end{aligned} \tag{4.44}$$

Возведем в квадрат равенства (4.44) и, суммируя левую и правую части, получим

$$[\,\varepsilon\varepsilon\,] = [\,vv\,] + nc^2 + 2c\,[\,v\,]. \tag{4.45}$$

Но член $2c\,[\,v\,] = 0$ по первому свойству отклонений $v_i$, тогда

$$[\,\varepsilon\varepsilon\,] = [\,vv\,] + nc^2. \tag{4.46}$$

Из формулы (4.46) следует, что

$$[\,\varepsilon\varepsilon\,] > [\,vv\,]$$

на положительное число $nc^2$, вне зависимости от того, $x'$ больше $\bar{x}$, или меньше. Таким образом, при $x' \neq x$

$$[\,vv\,] < [\,\varepsilon\varepsilon\,]\,, \tag{4.47}$$

что и требовалось доказать.

Для решения поставленной в начале данного параграфа задачи определим связь между истинными ошибками $\Delta_i$ и отклонениями $v_i$. Напишем

$$\Delta_i = x_i - X\,; \qquad v_i = x_i - \bar{x}\,.$$

Составим разность

$$\Delta_i - v_i = \bar{x} - X. \tag{4.48}$$

Возведем в квадрат равенства (4.48) и почленно просуммируем

$$[\Delta\Delta] = [vv] + 2[v](\bar{x} - X) + n \cdot (\bar{x} - X)^2\,. \tag{4.49}$$

Согласно первому свойству отклонений имеем

$$[\Delta\Delta] = [vv] + n \cdot (\bar{x} - X)^2 . \quad (4.50)$$

В равенстве (4.50) левая часть согласно формуле Гаусса равна

$$[\Delta\Delta] = n \cdot m^2 . \quad (4.51)$$

Разность $\bar{x} - X$ данного равенства соответствует средней квадратической ошибке значения среднего арифметического из результатов измерений, т.е. равна *M*. Тогда с учетом вышеизложенного получим вместо равенства (4.69) следующее

$$n \cdot m^2 = [vv] + m^2$$

или

$$m^2 (n-1) = [vv] .$$

Откуда средняя квадратическая ошибка результата измерения составит

$$m = \sqrt{\frac{[vv]}{n-1}} . \quad (4.52)$$

Средняя квадратическая ошибка вычисления ошибки согласно формуле (4.52) равна

$$m_m = \frac{m}{\sqrt{2(n-1)}} . \quad (4.53)$$

**Пример 18.** Определить среднюю квадратическую ошибку вычисления длины окружности, если ее радиус *R* измерен со средней квадратической ошибкой $m_R$ = 0,03 м.

**Решение.** Длина окружности определяется согласно формуле

$$L = 2\pi \cdot R, \quad (4.54)$$

тогда на основании (4.32) имеем

$$m_L = 2\pi \cdot m_R . \quad (4.55)$$

Подставляя в формулу (4.55) численные значения, получим

$m_L$ = 2·3,14·0,03 = 0,19 м.

Ответ: $m_L$ = 0,19 м.

**Пример 19**. Два угла треугольника измерены со средними квадратическими ошибками $m_1$ = 6,0" и $m_2$ = 8,0". Найти среднюю квадратическую ошибку третьего угла, вычисленного по двум измеренным углам.

**Решение**. В данном случае величина третьего угла определится из

$$\beta_3 = 180^\circ - (\beta_1 + \beta_2), \quad (4.56)$$

тогда согласно (4.35) имеем

$$m_3^2 = m_1^2 + m_2^2 . \quad (4.57)$$

Подставляя численные значения, получим

$$m_3 = \sqrt{6^2 + 8^2} = 10''.$$

Ответ: $m_3$ = 10”.

**Пример 20**. Измерены две линии: одна длиной 210,0 м со средней квадратической ошибкой $m_1$ = 9,0 см; другая - длиной 180,0 м с ошибкой $m_2$ = 8,0 см. С какой точностью будут вычислены расстояния, равные; 1) сумме двух линий; 2) разности этих линий?

**Решение**. Средняя квадратическая ошибка алгебраической суммы двух линий согласно (4.35)

$$m = \sqrt{6^2 + 9^2} = 12 \text{ см}.$$

Так как $S_1 + S_2$ = 210 + 180 = 390 м,

$S_1 - S_2$ = 210 - 180 = 30 м,

то относительная ошибка суммы и разности соответственно будут равны

$$\frac{1}{N_1} = \frac{0{,}12}{390} = \frac{1}{3200}; \qquad \frac{1}{N_2} = \frac{0{,}12}{30} = \frac{1}{250}.$$

Ответ: $\frac{1}{N_1} = \frac{1}{3200}; \qquad \frac{1}{N_2} = \frac{1}{250}.$

**Пример 21.** Определить среднюю квадратическую ошибку измерения угла теодолитом одним полным приемом, если средняя квадратическая ошибка измерения одного направления равна $m_0$ = 30”.

**Решение**. Средняя квадратическая ошибка угла из одного полуприема, полученного как разность двух направлений, будет равна

$$m_1 = m_0 \cdot \sqrt{2}. \tag{4.58}$$

Угол $\beta$, измеренный одним полным приемом, вычисляется как среднее из значений углов, полученных в полуприемах,

$$\beta = \frac{\beta_1 + \beta_2}{2} = \frac{1}{2}\beta_1 + \frac{1}{2}\beta_2. \tag{4.59}$$

Согласно формуле (4.34) получим

$$m_\beta^2 = \frac{1}{4}m_1^2 + \frac{1}{4}m_1^2 = \frac{1}{2}m_1^2,$$

тогда с учетом (4.58) , будем иметь

$$m_\beta = \frac{m_1}{\sqrt{2}} = \frac{m_0\sqrt{2}}{\sqrt{2}} = m_0 = 30''.$$

Ответ: $m_\beta$ = 30”.

**Пример 22.** Проложен ход геометрического нивелирования по ровной местности протяжением $L$ = 1600 м при длине визирного луча $l$ = 50 м. Найти среднюю квадратическую ошибку суммы превышений по всему ходу, если средняя квадратическая ошибка превышения на станции $m_h$ = 1,3 мм.

**Решение**. Согласно (4.36) имеем

$$m_{\Sigma h} = m_h \cdot \sqrt{n}, \qquad (4.60)$$

где $n$ – число станций.

Учитывая, что нивелирование по ровной местности осуществлялось методом из середины, число станций составит

$$n = \frac{L}{2l} = \frac{1600}{2 \cdot 50} = 16.$$

Тогда

$$m_{\Sigma h} = 1{,}3 \cdot \sqrt{16} = 5{,}2 \text{ мм}.$$

Ответ: $m_{\Sigma h}$ = 5,2 мм.

**Пример 23**. Для определения площади $P$ прямоугольника измерены две стороны $a$ = 144,28 м со средней квадратической ошибкой $m_a$ = 0,12 м и $b$ = 353,22 м со средней квадратической ошибкой $m_b$ = 0,09 м. Вычислить площадь и ее среднюю квадратическую ошибку.

**Решение.** 1-й вариант: Составим функцию

$$P = a\,b\,. \qquad (4.61)$$

Перейдем к средним квадратическим ошибкам согласно (4.28)

$$m_P^2 = a^2 m_b^2 + b^2 m_a^2. \qquad (4.62)$$

Подставим в (4.62) численные значения

$$m_P = \sqrt{144{,}28^2 \cdot 0{,}09^2 + 353{,}22^2 \cdot 0{,}12^2} = \sqrt{168{,}62 + 1796{,}61} = \sqrt{1965{,}23} = 44{,}33 \text{м}^2.$$

Вычислим площадь прямоугольника

$$P = 144{,}28 \cdot 353{,}22 = 50962{,}58 \text{ м}^2.$$

Относительная ошибка определения площади прямоугольника составит

$$\frac{1}{N} = \frac{m_P}{P} = \frac{44{,}33}{50962{,}58} = \frac{1}{1100}.$$

2-й вариант: логарифмируем функцию (4.61)

$$ln\,P = ln\,a + ln\,b$$

и, переходя к ошибкам, имеем

$$\frac{m_P^2}{P^2} = \frac{m_a^2}{a^2} + \frac{m_b^2}{b^2}. \qquad (4.63)$$

Откуда

$$\frac{m_P}{P} = \sqrt{\left(\frac{0{,}12}{144{,}28}\right)^2 + \left(\frac{0{,}09}{353{,}22}\right)^2} = \sqrt{75{,}66 \cdot 10^{-8}} = 8{,}7 \cdot 10^{-4}.$$

$$\frac{1}{N_P} = \frac{8{,}7}{10000} = \frac{1}{1100}.$$

Откуда

$$m_P = 44{,}33 \text{ м}^2 .$$

Ответ: $P = 50962{,}58$ м$^2$ ; $m_P = 44{,}33$ м$^2$ ; $\dfrac{1}{N_P} = \dfrac{1}{1100}$.

## 4.10. Оценка точности функции при наличии систематических ошибок

Закономерности возникновения случайных и систематических ошибок различны. Систематические ошибки подчиняются функциональным закономерностям, а случайные - статистическим. Рассмотрим влияние систематических ошибок на точность определения одного результата. Общая ошибка в этом случае составит

$$\Delta_i = \varepsilon_i + \Theta_i, \tag{4.64}$$

где $\varepsilon_i$ - случайная составляющая общей ошибки измерения;

$\Theta_i$ - систематическая составляющая.

В геодезической практике принято, если один из источников общей ошибки характеризуется средней квадратической ошибкой, не превышающей 1/3 средней квадратической ошибки, характеризующей другой источник, то первым можно пренебречь

$$m_\Theta = \frac{m_\varepsilon}{3}, \tag{4.65}$$

тогда

$$m = \sqrt{m_\varepsilon^2 + \frac{m_\varepsilon^2}{9}} = m_\varepsilon\sqrt{1{,}11} = 1{,}05 m_\varepsilon,$$

т.е. значение средней квадратической ошибки $m$, если пренебречь систематической ошибкой, уменьшится всего на 5%.

Определим совместное влияние случайных и систематических ошибок на отдельные результаты измерений. Будем считать, что общая ошибка содержит постоянную систематическую часть Θ. Тогда ряд ошибок измерений предстанет в виде

$$\begin{aligned} &\Delta_1 = \varepsilon_1 + \Theta_1; \\ &\Delta_2 = \varepsilon_2 + \Theta_2; \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ &\Delta_n = \varepsilon_n + \Theta_n, \end{aligned} \tag{4.66}$$

где $\Delta_i$ - общая ошибка измерений;

$\varepsilon_i$ - случайная составляющая;

$\Theta_i$ - систематическая составляющая.

Равенства (4.85) почленно возведем в квадрат

$$\begin{aligned} &\Delta_1\Delta_1 = \varepsilon_1\varepsilon_1 + 2\varepsilon_1\Theta_1 + \Theta_1^2; \\ &\Delta_2\Delta_2 = \varepsilon_2\varepsilon_2 + 2\varepsilon_2\Theta_2 + \Theta_2^2; \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &\Delta_n\Delta_n = \varepsilon_n\varepsilon_n + 2\varepsilon_n\Theta_n + \Theta_n^2. \end{aligned} \tag{4.67}$$

Учитывая, что $\Theta_1 = \Theta_2 = \ldots = \Theta_n = \Theta$, просуммируем полученные значения, а сумму разделим на $n$, т.е.

$$\frac{[\Delta\Delta]}{n} = \frac{[\varepsilon\varepsilon]}{n} + 2\Theta\frac{[\varepsilon]}{n} + \Theta^2. \tag{4.68}$$

Второй член правой части равенства (4.68) согласно четвертому свойству случайных ошибок равен нулю. Переходя к средним квадратическим ошибкам, получим

$$m^2 = m_\varepsilon^2 + \Theta^2. \tag{4.69}$$

Аналогично определяется средняя квадратическая ошибка функции

$$M_F^2 = M_{F\varepsilon}^2 + \Theta_F^2. \tag{4.70}$$

Первое слагаемое определяется согласно (4.28)

$$M_{F\varepsilon}^2 = \left(\frac{\partial f}{\partial x_1}\right)^2 m_1^2 + \left(\frac{\partial f}{\partial x_2}\right)^2 m_2^2 + \ldots + \left(\frac{\partial f}{\partial x_n}\right)^2 m_n^2. \tag{4.71}$$

Величина систематической ошибки определится из следующего выражения

$$\Theta_F = \left(\frac{\partial f}{\partial x_1}\right)\cdot\Theta_1 + \left(\frac{\partial f}{\partial x_2}\right)\cdot\Theta_2 + \ldots + \left(\frac{\partial f}{\partial x_n}\right)\cdot\Theta_n. \qquad (4.72)$$

Тогда

$$M_F^2 = \left(\frac{\partial f}{\partial x_1}\right)^2 m_1^2 + \left(\frac{\partial f}{\partial x_2}\right)^2 m_2^2 + \ldots + \left(\frac{\partial f}{\partial x_n}\right)^2 m_n^2 + \left[\left(\frac{\partial f}{\partial x_1}\right)\cdot\Theta_1 + \left(\frac{\partial f}{\partial x_2}\right)\cdot\Theta_2 + \ldots + \left(\frac{\partial f}{\partial x_n}\right)\cdot\Theta_n\right]^2. \qquad (4.73)$$

Пусть слагаемые $x_1$, $x_2$, ... , $x_n$ результаты равноточных измерений со средней квадратической ошибкой $m_\varepsilon$ и постоянной систематической ошибкой $\Theta$. Тогда при

$$y = x_1 + x_2 + \ldots + x_n$$

имеем

$$M_y^2 = M_\varepsilon^2 + M_\Theta^2, \qquad (4.74)$$

где $M_\varepsilon = m_\varepsilon\sqrt{n}$, а $M_\Theta = \Theta \cdot n$.

Тогда средняя квадратическая ошибка суммы составит

$$M_y = \sqrt{m_\varepsilon^2 \cdot n + \Theta^2 \cdot n^2}. \qquad (4.75)$$

**Пример 24.** В современном высокоточном геометрическом нивелировании систематическая часть общей ошибки на 1 км составляет 1/10 части средней квадратической случайной ошибки. Сравним влияние случайной и систематической ошибок измерений.

**Решение**. На основании (4.75) имеем

$$m_h = \sqrt{s\cdot m^2 + \frac{s^2}{100}m^2}, \qquad (4.76)$$

где s - протяженность нивелирного хода, в км;

$m$- средняя квадратическая ошибка нивелирования 1 км.

Вычислим случайную и систематическую части ошибок нивелирования в ходах различной протяженности:

1) на 10 км $\quad s\cdot m^2 = 10\cdot m^2;\quad \frac{s^2}{100}m^2 = m^2;$

2) на 100 км $\quad s\cdot m^2 = 100\cdot m^2;\quad \frac{s^2}{100}m^2 = 100\cdot m^2;$

3) на 300 км $\quad s\cdot m^2 = 300\cdot m^2;\quad \frac{s^2}{100}m^2 = 900\cdot m^2.$

В первом случае случайная составляющая превышает систематическую, во втором - они равны, а в третьем - уже систематическая превышает случайную.

## 4.11. Принцип равных влияний

При проектировании геодезических работ рассчитывают точность предстоящих измерений, пользуясь формулами теории ошибок измерений. При этих расчетах встречается задача, в которой по известному виду функции *F* требуется рассчитать точность измерений аргументов. Чтобы иметь определенное решение, применяют принцип равных влияний.

Пусть дана функция общего вида

$$y = f(x_1, x_2, \dots, x_n). \qquad (4.77)$$

На основании равенства (4.28) средняя квадратическая ошибка функции общего вида определится

$$M_y^2 = \left(\frac{\partial f}{\partial x_1}\right)^2 m_1^2 + \left(\frac{\partial f}{\partial x_2}\right)^2 m_2^2 + \dots + \left(\frac{\partial f}{\partial x_n}\right)^2 m_n^2.$$

Предположим, что

$$\frac{M_y^2}{n} = \left(\frac{\partial f}{\partial x_1}\right)^2 m_1^2 = \left(\frac{\partial f}{\partial x_2}\right)^2 m_2^2 = \dots = \left(\frac{\partial f}{\partial x_n}\right)^2 m_n^2.$$

Откуда

$$\frac{M_y}{\sqrt{n}} = \left|\left(\frac{\partial f}{\partial x_1}\right) \cdot m_1\right| = \left|\left(\frac{\partial f}{\partial x_2}\right) \cdot m_2\right| = \dots = \left|\left(\frac{\partial f}{\partial x_n}\right) \cdot m_n\right|. \qquad (4.78)$$

Эти соотношения позволяют вычислить средние квадратические ошибки аргументов при заданном $M_y$, т.е. определить необходимую точность измерений.

**Пример 25.** Ошибка определения опознака полярным способом вычисляется по формуле

$$m_o = \sqrt{\frac{s^2}{2\rho^2} m_\beta^2 + \frac{m_s^2}{2}}. \qquad (4.79)$$

Заданная точность $m_o$ = 1 м. Определить, с какой точностью нужно измерять углы и линии?

**Решение.** 1. Примем s = 1 км. Тогда на основании (4.78) имеем

$$\frac{m_o}{\sqrt{2}} = \frac{s}{\sqrt{2}\rho} m_\beta = \frac{m_s}{\sqrt{2}}; \quad \text{откуда} \quad m_o = m_s = 1 \text{ м}, \text{ а}$$

$$m_\beta = \frac{m_o \cdot \rho}{s} = \frac{1 \cdot 3438}{1000} = 3{,}4'.$$

Относительная ошибка измерения линии составит $\frac{1}{N_s} = \frac{m_s}{s} = \frac{1}{1000}$.

2. Примем s = 5 км. Тогда

$\frac{m_o}{\sqrt{2}} = \frac{m_s}{\sqrt{2}}$; откуда $m_s = m_o = 1$ м. $m_\beta = \frac{m_o \cdot \rho}{s} = \frac{1 \cdot 3438}{5000} = 0{,}7'$. Относительная ошибка $\frac{1}{N_s} = \frac{1}{5000}$.

Ответ: 1. $m_\beta = 3{,}4'$; $\frac{1}{N_s} = \frac{1}{1000}$; 2. $m_\beta = 0{,}7'$; $\frac{1}{N_s} = \frac{1}{5000}$.

### 4.12. Неравноточные измерения и их веса

В практике математической обработки результатов геодезических измерений встречаются случаи, когда приходится анализировать накопленный материал с неодинаковым числом измерений, инструментами разной точности, с различным числом измеряемых величин и различной протяженности, выполненных в разных условиях. Такие измерения можно отнести к неравноточным. Возникает задача определения по результатам неравноточных наблюдений наиболее надежного значения измеряемой величины, оценки точности этих измерений, а также уравнивания различных геодезических сетей.

Достоинство результата измерения, меру его надежности обозначают числом, называемым весом этого измерения, т.е. весом называют степень доверия к результату измерения, выраженную числом.

Чем лучше условия измерения, чем надежнее результат, тем больше его вес, т.е. тем больше наше доверие к нему. Таким образом, вес характеризует условия измерения. Но определенным условиям соответствует определенная средняя квадратическая ошибка. Чем меньше средняя квадратическая ошибка, тем надежнее результат, а следовательно, тем больше его вес. Исходя из сказанного, за веса результатов измерений принимают величины обратно пропоциональные квадратам соответствующих им средних квадратических ошибок.

Пусть некоторая величина измерялась неравноточно $n$ раз $x_1$ , $x_2$, . . . , $x_n$ , со средними квадратическими ошибками $m_1$, $m_2$, . . . , $m_n$, тогда веса результатов измерений будут равны

$$p_1 = \frac{C}{m_1^2}; \;\; p_2 = \frac{C}{m_2^2}; \; \dots \; ; \; p_n = \frac{C}{m_n^2}, \qquad (4.80)$$

где С - коэффициент пропорциональности, который может быть выбран любым, но одинаковым для данного ряда измерений.

При установлении весов необходимо соблюдать следующие условия:

1) средние квадратические ошибки, по которым определяются веса, должны быть найдены из достаточно большого числа наблюдений;

2) из измерений, по которым вычисляются средние квадратические ошибки, должны быть исключены систематические ошибки.

Согласно (4.80) при различных значениях $C$ получаем и различные веса, однако соотношение между ними остается неизменным. Отсюда следует, что веса данного ряда измерений являются величинами относительными и их можно одновременно увеличивать или уменьшать в одинаковое число раз.

## 4.13. Общая арифметическая средина и ее свойства

Пусть дан ряд результатов измерений

$$x_1 , \; x_2 \; , \; \dots \; , \; x_{\mathrm{n}} \; ;$$
$$m_1 , m_2 \; , \; \dots \; , \; m_{\mathrm{n}} \; .$$

Требуется найти наиболее надежное значение $\bar{x}$ измеренной величины. Выразим искомую величину в виде линейной функции

$$\bar{x} = k_1 x_1 + k_2 x_2 + \ldots + k_n x_n , \qquad (4.81)$$

где $k_i$ являются некоторой функцией величин $m_i$ и связаны условием

$$k_1 + k_2 + \ldots + k_n = 1. \qquad (4.82)$$

Тогда

$$M^2 = k_1{}^2 m_1{}^2 + k_2{}^2 m_2{}^2 + \ldots + k_n{}^2 m_n{}^2 . \qquad (4.83)$$

Функция приведет нас к надежному результату, если ее средняя квадратическая ошибка будет наименьшей, т. е.

$$M^2 = min\ . \tag{4.84}$$

Задачу решим по методу Лагранжа

$$\Phi\left(k_1, k_2, ..., k_n\right) = m_1^2 k_1^2 + m_2^2 k_2^2 + ... + m_n^2 k_n^2 - 2\lambda\left(k_1 + k_2 + ... + k_n - 1\right), \tag{4.85}$$

где $\lambda$ - подлежащий определению коэффициент, называемый множителем Лагранжа.

Условие (4.86) определяется точкой экстремума функции Лагранжа Ф.

$$\begin{aligned} &\frac{\partial\Phi}{\partial k_1} = 2k_1 m_1^2 - 2\lambda = 0; \\ &\frac{\partial\Phi}{\partial k_2} = 2k_2 m_2^2 - 2\lambda = 0; \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &\frac{\partial\Phi}{\partial k_n} = 2k_n m_n^2 - 2\lambda = 0. \end{aligned} \tag{4.86}$$

Отсюда

$$\begin{aligned} &k_1 = \frac{\lambda}{m_1^2}; \\ &k_2 = \frac{\lambda}{m_2^2}; \\ &\ldots\ldots\ldots \\ &k_n = \frac{\lambda}{m_n^2}. \end{aligned} \tag{4.87}$$

Подставив эти значения в уравнение (4.81), получим

$$\bar{x} = \lambda\left(\frac{1}{m_1^2} x_1 + \frac{1}{m_2^2} x_2 + ... + \frac{1}{m_n^2} x_n\right). \tag{4.88}$$

Из равенств (4.87) найдем

$$\left[k\right] = \lambda\left[\frac{1}{m^2}\right]. \tag{4.89}$$

С учетом равенства (4.82) определим множитель Лагранжа $\lambda$

$$\lambda = \frac{1}{\left[\frac{1}{m^2}\right]}. \qquad (4.90)$$

Полученное значение $\lambda$ подставим в уравнение (4.88)

$$\bar{x} = \frac{\frac{1}{m_1^2}x_1 + \frac{1}{m_2^2}x_2 + \ldots + \frac{1}{m_n^2}x_n}{\frac{1}{m_1^2} + \frac{1}{m_2^2} + \ldots + \frac{1}{m_n^2}}. \qquad (4.91)$$

Учитывая (4.80), равенство (4.91) можно записать в следующем виде:

$$\bar{x} = \frac{p_1 x_1 + p_2 x_2 + \ldots + p_n x_n}{p_1 + p_2 + \ldots + p_n} = \frac{[px]}{[p]}. \qquad (4.92)$$

Общая арифметическая средина равна сумме произведений каждого неравноточного измерения на его вес, разделенной на сумму весов.

Рассмотрим свойства отклонений от общей арифметической средины.

**Первое свойство**. Алгебраическая сумма произведений отклонений результатов неравноточных измерений от общей арифметической средины на соответствующие веса равна нулю при любом числе наблюдений, т.е.

$$[pv] = 0. \qquad (4.93)$$

Дан ряд отклонений от арифметической средины с соответствующими весами $p_1, p_2, \ldots, p_n$.

$$\begin{array}{ll} p_1 & v_1 = x_1 - \bar{x}; \\ p_2 & v_2 = x_2 - \bar{x}; \\ \ldots & \ldots \\ p_n & v_n = x_n - \bar{x}. \end{array} \qquad (4.94)$$

Перемножив равенства (4.94) на соответствующие веса и сложив левую и правую части, получим

$$[pv] = [px] - [p]\bar{x}.$$

Согласно (4.92) будем иметь

$$[px] = [p]\bar{x},$$

следовательно,

$$[pv] = 0. \qquad (4.95)$$

Данное свойство можно использовать для контроля вычислений общей арифметической средины.

**Второе свойство**. Сумма произведений квадратов отклонений результатов неравноточных измерений от общей арифметической средины на соответствующие веса является наименьшей, т.е

$$[\,pvv\,] = min \tag{4.96}$$

Пусть $x' \neq \bar{x}$, тогда

$$v_i' = x_i - x'; \quad v_i = x_i - \bar{x}. \tag{4.67}$$

Установим связь между отклонениями $v$ и $v'$

$$v_i' - v_i = \bar{x} - x'. \tag{4.98}$$

В равенстве (4.98) $v_i$ перенесем в правую часть. Затем, умножая на соответствующие веса, возведем в квадрат и почленно сложим

$$[pv'v'] = [pvv] + 2[pv](\bar{x} - x') + [p](\bar{x} - x')^2. \tag{4.99}$$

В правой части равенства (4.99) слагаемое $2[pv](\bar{x} - x') = 0$ согласно первому свойству отклонений (4.85). Следовательно, из равенства (4.99) следует, что

$$[pvv] < [pv'v']. \tag{4.100}$$

Данное свойство подтверждает, что если ошибки результатов неравноточных измерений подчиняются нормальному закону распределения, то наиболее надежным значением является общая арифметическая средина.

## 4.14. Средняя квадратическая ошибка единицы веса

Дан ряд результатов неравноточных измерений

$$x_1\,,\; x_2\,,\; \ldots\;,\; x_n\;;$$

$$m_1\,, m_2\,,\; \ldots\;,\; m_n\,.$$

Соответственно

$$p_1 = \frac{C}{m_1^2};\;\; p_2 = \frac{C}{m_2^2};\; \ldots\;;\;\; p_k = \frac{C}{m_k^2} = 1;\; \ldots\;;\; p_n = \frac{C}{m_n^2}. \tag{4.101}$$

Как видим из равенств (4.101), всегда найдется такое соотношение, когда $\frac{C}{m_i^2} = 1$, т.е. коэффициент С есть не что иное, как средняя квадратическая ошибка измерения, вес которой равен единице и которая в отличие от остальных средних квадратических ошибок обозначается $\mu$ и называется ошибкой единицы веса. Тогда

$$C = m_k^2 = \mu^2. \tag{4.102}$$

Следовательно,

$$p_1 = \frac{\mu^2}{m_1^2};\ \ p_2 = \frac{\mu^2}{m_2^2};\ \dots\ ;\ \ p_n = \frac{\mu^2}{m_n^2}. \tag{4.103}$$

В соответствии с (4.103) можно записать

$$\mu = m_i\sqrt{p_i} \tag{4.104}$$

или

$$m_i = \frac{\mu}{\sqrt{p_i}}, \tag{4.105}$$

т.е. средняя квадратическая ошибка любого результата измерения равна ошибке единицы веса, деленной на корень квадратный из веса соответствующего результата.

## 4.15. Вычисление весов функций. Вес функции независимых величин

Определим вес функции общего вида

$$F = f(x_1, x_2, \dots, x_n). \tag{4.106}$$

При этом известны веса аргументов $p_1$ , $p_2$ , ... , $p_n$. В случае независимых величин имеем

$$M_F^2 = \left(\frac{\partial f}{\partial x_1}\right)^2 m_1^2 + \left(\frac{\partial f}{\partial x_2}\right)^2 m_2^2 + \dots + \left(\frac{\partial f}{\partial x_n}\right)^2 m_n^2. \tag{4.107}$$

Разделим обе части равенства (4.107) на $\mu^2$

$$\frac{M_F^2}{\mu^2} = \left(\frac{\partial f}{\partial x_1}\right)^2 \frac{m_1^2}{\mu^2} + \left(\frac{\partial f}{\partial x_2}\right)^2 \frac{m_2^2}{\mu^2} + \dots + \left(\frac{\partial f}{\partial x_n}\right)^2 \frac{m_n^2}{\mu^2},$$

так как $p_i = \frac{\mu^2}{m_i^2}$, то окончательно имеем

$$\frac{1}{P_{fF}} = \left(\frac{\partial f}{\partial x_1}\right)^2 \frac{1}{p_1} + \left(\frac{\partial f}{\partial x_2}\right)^2 \frac{1}{p_2} + \ldots + \left(\frac{\partial f}{\partial x_n}\right)^2 \frac{1}{p_n}. \qquad (4.108)$$

1. Вес функции неравноточных слагаемых. Дана функция вида

$$F = x_1 + x_2 + \ . \ . \ . + x_n, \qquad (4.109)$$

где $x_i$ - результаты неравноточных измерений с соответствующими весами: $p_1, p_2, \ldots, p_n$ .

Требуется определить вес функции $F$ . Согласно (4.108) будем иметь

$$\frac{1}{P_F} = \frac{1}{p_1} + \frac{1}{p_2} + \ldots + \frac{1}{p_n}. \qquad (4.110)$$

Обратный вес суммы неравноточных слагаемых равен сумме обратных весов.

2. Вес суммы равноточных слагаемых. Дана функция вида (4.109), в которой $x_i$ являются результатами равноточных измерений, т.е. $p_1 = p_2 = \ldots = p_n = p$. Тогда вес такой функции определится из равенства (4.110)

$$\frac{1}{P_F} = \frac{1}{p} + \frac{1}{p} + \ldots + \frac{1}{p} = \frac{n}{p}. \qquad (4.111)$$

Откуда

$$P_F = P_\Sigma = \frac{p}{n}. \qquad (4.112)$$

Вес суммы равноточных слагаемых в $n$ раз меньше веса одного измерения.

3. Вес простой арифметической средины. Поскольку простая арифметическая средина вычисляется согласно формуле

$$\bar{x} = \frac{[x]}{n}, \qquad (4.113)$$

ее средняя квадратическая ошибка будет равна

$$M = \frac{m}{\sqrt{n}},$$

откуда

$$M^2 = \frac{m^2}{n}.$$

Переходя к весам, получим

$$\frac{1}{P} = \frac{1}{pn}$$

или

$$P = pn. \qquad (4.114)$$

Вес простой арифметической средины в $n$ раз больше веса одного измерения.

4. Вес и средняя квадратическая ошибка общей арифметической средины. Представив общую арифметическую средину в виде линейной функции

$$\bar{x} = \frac{p_1}{[p]} x_1 + \frac{p_2}{[p]} x_2 + \ldots + \frac{p_n}{[p]} x_n, \qquad (4.115)$$

согласно равенству (4.108) получим

$$\frac{1}{P} = \frac{p_1^2}{[p]^2} \frac{1}{p_1} + \frac{p_2^2}{[p]^2} \frac{1}{p_2} + \ldots + \frac{p_n^2}{[p]^2} \frac{1}{p_n}$$

или

$$\frac{1}{P} = \frac{[p]}{[p]^2} = \frac{1}{[p]},$$

откуда

$$P = [p]. \qquad (4.116)$$

Вес общей арифметической средины равен сумме весов результатов измерений.

Среднюю квадратическую ошибка $M$ общей арифметической средины получим согласно формуле (4.105)

$$M = \frac{\mu}{\sqrt{P}} \qquad (4.117)$$

или

$$M = \frac{\mu}{\sqrt{[p]}}. \qquad (4.118)$$

## 4.16. Вычисление ошибки единицы веса

1. Вычисление ошибки единицы веса при установлении весов по известным средним квадратическим ошибкам.

Согласно формулам (4.101) и (4.105) имеем

$$p_i = \frac{C}{m_i^2} = \frac{\mu^2}{m_i^2},$$

откуда

$$\mu = \sqrt{C}. \qquad (4.119)$$

Если при этом известны веса, то ошибку единицы веса можно вычислить по формуле

$$\mu = m_i \sqrt{p_i}. \qquad (4.120)$$

2. Вычисление ошибки единицы веса через истинные ошибки. Дан ряд неравноточных измерений

$$x_1,\ x_2,\ \dots,\ x_n \qquad (4.121)$$

с соответствующими истинными ошибками, весами и средними квадратическими ошибками

$$\Delta_1,\ \Delta_2,\ \dots,\ \Delta_n; \qquad (4.122)$$

$$p_1,\ p_2,\ \dots,\ p_n; \qquad (4.123)$$

$$m_1,\ m_2,\ \dots,\ m_n. \qquad (4.124)$$

Умножим каждый результат ряда (4.121) на $\sqrt{p_i}$, и получим новый ряд

$$x_1\sqrt{p_1};\ \ x_2\sqrt{p_2};\ \dots;\ \ x_n\sqrt{p_n}. \qquad (4.125)$$

При увеличении или уменьшении значения $x$ в произвольное число раз изменяются и истинные ошибки в соответствующее число раз

$$\Delta_1\sqrt{p_1};\ \ \Delta_2\sqrt{p_2};\ \dots;\ \ \Delta_n\sqrt{p_n}. \qquad (4.126)$$

Соответственно изменится и ряд средних квадратических ошибок

$$m_1\sqrt{p_1};\ \ m_2\sqrt{p_2};\ \dots;\ \ m_n\sqrt{p_n}, \qquad (4.127)$$

а это согласно (4.120)

$$m_1\sqrt{p_1} = m_2\sqrt{p_2} = \dots = m_n\sqrt{p_n} = \mu. \qquad (4.128)$$

Откуда следует, что ряд (4.125) является равноточным со средней квадратической ошибкой $\mu$ и истинными ошибками (4.126), тогда, использовав формулу Гаусса (4.13), получим

$$\mu = \sqrt{\frac{[p\Delta\Delta]}{n}}. \tag{4.129}$$

Надежность определения $\mu$ вычислится

$$m_\mu = \frac{\mu}{\sqrt{2n}}. \tag{4.130}$$

3. Вычисление средней квадратической ошибки измерения углов в триангуляции. Даны невязки в треугольниках $w_1, w_2, \ldots, w_n$, причем

$$w_i = \varphi_i(\beta_1, \beta_2, \beta_3) = \sum_{j=1}^{3} \beta_j - 180°. \tag{4.131}$$

Считая углы равноточно измеренными с весом $p = 1$, находим вес $p_i$ функции $\varphi_i$

$$\frac{1}{p_i} = 1 + 1 + 1 = 3$$

или

$$p_i = \frac{1}{3}.$$

Так как невязки являются истинными ошибками измерения углов в треугольнике, воспользуемся формулой (4.129) для вычисления ошибки единицы веса $\mu$

$$\mu = \sqrt{\frac{[pww]}{N}} = \sqrt{\frac{\frac{1}{3}[ww]}{N}},$$

где $N$ - число треугольников сети триангуляции.

Окончательно имеем

$$\mu = m_\beta = \sqrt{\frac{[ww]}{3N}}. \tag{4.132}$$

Формула (4.132) носит название формулы Ферреро.

4. Вычисление ошибки единицы веса через отклонения от арифметической средины. Пусть даны результаты неравноточных измерений $x_1, x_2, \dots, x_n$ с весами $p_1, p_2, \dots, p_n$, а также известны истинное значение $X$ и среднее арифметическое $\bar{x}_0$. Составим два ряда соответственно истинных ошибок и отклонений от арифметической средины

$$
\begin{array}{ll}
\Delta_1 = x_1 - X; & v_1 = x_1 - \bar{x}_0; \\
\Delta_2 = x_2 - X; & v_2 = x_2 - \bar{x}_0; \\
\dots\dots\dots\dots & \dots\dots\dots\dots \\
\Delta_n = x_n - X; & v_n = x_n - \bar{x}_0.
\end{array}
\qquad (4.133)
$$

Из первого ряда вычтем второй, в результате чего получим

$$
\begin{array}{l}
\Delta_1 = v_1 + (\bar{x}_0 - X); \\
\Delta_2 = v_2 + (\bar{x}_0 - X); \\
\dots\dots\dots\dots\dots\dots\dots\dots \\
\Delta_n = v_n + (\bar{x}_0 - X).
\end{array}
\qquad (4.134)
$$

Равенства (4.134) возведем в квадрат и умножим на соответствующие веса. Полученные равенства почленно просуммируем и в результате получим

$$[p\Delta\Delta] = [pvv] + 2[pv](\bar{x}_0 - X) + [p](\bar{x}_0 - X)^2. \qquad (4.135)$$

В правой части равенства (4.135) второе слагаемое согласно первому свойству отклонений от общей арифметической средины равно нулю. Следовательно,

$$[p\Delta\Delta] = [pvv] + [p](\bar{x}_0 - X)^2.$$

Согласно (4.129) имеем $[p\Delta\Delta] = \mu^2 n$, а выражение в круглых скобках является средней квадратической ошибкой среднего значения, т.е.

$$\bar{x}_0 - X = M = \frac{\mu}{\sqrt{[p]}} \qquad (4.136)$$

или

$$(\bar{x}_0 - X)^2 = M^2 = \frac{\mu^2}{[p]}. \qquad (4.137)$$

С учетом (4.136) равенство (4.137) примет вид

$$\mu^2 n = [pvv] + [p]\frac{\mu^2}{[p]}$$

или

$$\mu^2 n = [pvv] + \mu^2.$$

Откуда

$$\mu = \sqrt{\frac{[pvv]}{n-1}}. \qquad (4.138)$$

**Пример 26.** Определить вес одного направления, если вес угла, полученного как разность двух направлений, равен единице.

**Решение.** Зависимость между значениями угла и направления имеет вид

$$\beta = a - b,$$

где $a$ и $b$ - отсчеты по горизонтальному кругу .

Так как отсчеты по горизонтальному кругу равноточны между собой , т.е. $p_a = p_b = p$, то можно воспользоваться для определения веса угла $\beta$ формулой (4.112)

$$p_\beta = \frac{p}{n},$$

откуда вес одного направления

$$p = p_\beta \cdot n.$$

Подставляя численные значения, получим

$$p = 1 \cdot 2 = 2 .$$

Ответ: $p = 2$ .

**Пример 27**. Веса независимо измеренных углов $\alpha$, $\beta$ и $\gamma$ соответственно равны 5, 3 и 2 . Определить вес суммарного угла.

**Решение.** Воспользуемся формулой (4.110), найдем вес суммы неравноточно измеренных углов

$$\frac{1}{P_\Sigma} = \frac{1}{p_\alpha} + \frac{1}{p_\beta} + \frac{1}{p_\gamma} = \frac{1}{5} + \frac{1}{3} + \frac{1}{2} = \frac{31}{30}$$

или

$$P_\Sigma = \frac{30}{31} = 0{,}97 \approx 1.$$

Ответ: $P_\Sigma = 1$ .

**Пример 28**. Вес угла равен 9. Найти среднюю квадратическую ошибку этого угла, если ошибка единицы веса $\mu$ = 15”.

**Решение.** Из выражения (4.103) получим значение средней квадратической ошибки

$$m_\beta = \frac{\mu}{\sqrt{p_\beta}}. \qquad (4.140)$$

Подставив в равенство (4.140) численные значения, вычислим величину ошибки измеренного угла

$$m_\beta = \frac{15}{\sqrt{9}} = 5''.$$

Ответ: $m_\beta = 5''$.

**Пример 29.** Угол измерен теодолитом 9 приемами. Средняя квадратическая ошибка определения одного направления $m_0 = 7''$. Какой вес имеет полученный результат, если ошибка единицы веса $\mu = 3{,}3''$?

**Решение**. Угол $\beta_1$ в полуприеме получается как разность двух направлений, тогда согласно (4.36) имеем

$$m_1 = m_0\sqrt{2},$$

а величина средней квадратической ошибки измерения угла одним полным приемом составит

$$m_\beta = \frac{m_1}{\sqrt{2}} = \frac{m_0\sqrt{2}}{\sqrt{2}} = m_0,$$

т.е.

$$m_\beta = m_0 = 7''.$$

Так как окончательный результат получен из 9 приемов, то средняя квадратическая ошибка вероятнейшего значения угла согласно (4.39)

$$M = \frac{m_\beta}{\sqrt{n}} = \frac{7}{\sqrt{9}} = 2{,}3''.$$

Вес результата вычислится по формуле

$$P = \frac{\mu^2}{M^2} = \frac{10{,}89}{5{,}29} = 2.$$

Ответ: $P = 2$.

**Пример 30**. Сколько раз необходимо измерить линию длиной 300 м, чтобы вес результата оказался равным трем, если вес результата одного измерения линии длиной 100 м равен единице?

**Решение.** Вес окончательного значения равноточно измеренной величины в $n$ раз больше веса одного измерения, т.е.

$$P = p \cdot n. \qquad (4.141)$$

Для определения числа измерений данной линии необходимо установить вес однократного измерения. Так как по условию задачи вес одного измерения линии длиной 100 м равен единице, то вес такого же определения линии длиной 300 м согласно (4.110) составит

$$\frac{1}{P_0} = \frac{1}{p} + \frac{1}{p} + \frac{1}{p} = 3,$$

где $p$ - вес результата однократного измерения линии длиной 100 м.

Значение веса в этом случае

$$P_0 = \frac{1}{3}.$$

Подставляя в формулу (4.141) численные значения, получим число измерений

$$n = 9.$$

Ответ: $n = 9$ .

**Пример 31.** Угол получен со средней квадратической ошибкой $M$ = 4,5'' Сколько приемов нужно сделать инструментом, дающим результат одного измерения со средней квадратической ошибкой $m_2$ = 11,2'', чтобы веса вероятнейших значений углов оказались одинаковыми?

**Решение**. Если $P_1 = P_2$ , то и $M_1 = M_2 = M$ = 4,5''. Тогда, используя формулу (4.29), будем иметь

$$n_2 = \frac{m_2^2}{M^2} = \frac{125{,}44}{20{,}25} = 6.$$

Ответ: $n = 6$ .

**Пример 32.** В треугольнике один угол измерен тремя приемами, второй - двумя. Найти вес третьего угла, вычисленного по двум первым, учитывая, что за единицу веса принят вес результата измерения угла одним приемом.

**Решение**. Вес первого угла $P$ = 3, а вес второго $P$ = 2. По формуле (4.110) находим

$$\frac{1}{P_3} = \frac{1}{P_1} + \frac{1}{P_2} = \frac{1}{3} + \frac{1}{2} = \frac{5}{6}$$, откуда $P_3$ = 1,2.

Ответ: $P_3$ = 1,2 .

**Пример 33.** Радиус окружности $R$ измерен с весом $P_R$ = 4. Определить веса вычисления длины окружности и площади круга.

**Решение**. Составим функциональные зависимости

$$L = 2\pi \cdot R \qquad (4.142)$$

и

$$S = \pi \cdot R^2, \qquad (4.143)$$

где $L$ - длина окружности;

$S$ - площадь круга.

Переходя к весам, получим согласно (4.108)

$$\frac{1}{P_L} = \frac{4\pi^2}{P_R} = \frac{4\pi^2}{4} = \pi^2.$$

Откуда

$$P_L = \frac{1}{\pi^2}.$$

Аналогично определим вес вычисленной площади круга

$$\frac{1}{P_S} = \frac{4\pi^2 R^2}{P_R} = \frac{4\pi^2 R^2}{4} = \pi^2 R^2$$

или

$$P_S = \frac{1}{\pi^2 R^2}.$$

Ответ: $P_L = \dfrac{1}{\pi^2}$; $P_S = \dfrac{1}{\pi^2 R^2}$.

**Пример 34**. Определить вес площади треугольника, если его основание $b = 12$ м получено с весом $P_b = 2$, а высота $h = 20$ м - с весом $P_h = 1{,}5$.

**Решение**. Напишем функцию

$$S = \frac{1}{2} b \cdot h. \tag{4.144}$$

Следовательно,

$$\frac{1}{P_s} = \frac{b^2}{4P_h} + \frac{h^2}{4P_b}. \tag{4.145}$$

Подставляя численные значения в выражение (4.145), определим обратный вес данной функции

$$\frac{1}{P_S} = \frac{12^2}{4 \cdot 1{,}5} + \frac{20^2}{4 \cdot 2{,}0} = \frac{144}{6} + \frac{400}{8} = 74$$

или

$$P_S = \frac{1}{74} = 0{,}013.$$

Ответ: $P_S = 0{,}013$.

**Пример 35.** Найти обратный вес функции

$$y = \frac{1}{4}x_1 - \frac{1}{2}x_2 + \frac{2}{3}x_3, \qquad (4.146)$$

если веса аргументов соответственно равны: $P_1 = 0{,}25$; $P_2 = 0{,}50$; $P_3 = 1{,}0$ .

**Решение**. По формуле (4.102) находим

$$\frac{1}{P_y} = \frac{1}{16}\cdot\frac{1}{P_1} + \frac{1}{4}\cdot\frac{1}{P_2} + \frac{4}{9}\cdot\frac{1}{P_3}$$

или

$$\frac{1}{P_y} = \frac{1}{16\cdot 0{,}25} + \frac{1}{4\cdot 0{,}50} + \frac{4}{9\cdot 1{,}00} = \frac{43}{36}.$$

Откуда

$$P_y = \frac{36}{43} = 0{,}84 .$$

Ответ : $P_y = 0{,}84$

## 5. Математическая обработка результатов многократных измерений одной величины

В практике геодезических измерений часто одну и ту же величину измеряют многократно. Возникает необходимость получить окончательные значения из ряда измерений, а также произвести оценку точности произведенных измерений и окончательного результата. При этом возникает задача отбраковки некачественно выполненных измерений.

В результате отбраковки можно найти:

1) наиболее надежное (среднее арифметическое) значение измеряемой величины $\bar{x}$;

2) среднюю квадратическую ошибку результата одного измерения $m$;

3) среднюю ошибку одного измерения $\vartheta$;

4) вероятную (срединную) ошибку результата измерения $r$;

5) фактические и теоретические соотношения $\vartheta/m$ и $r/m$;

6) надежность вычисления средней квадратической ошибки одного измерений $m_m$;

7) среднюю квадратическую ошибку вероятнейшего значения $M$.

## 5.1. Порядок обработки результатов равноточных измерений одной величины

1. Определяют вероятнейшее из всех результатов измерений значение, т.е. арифметическое среднее:

$$\bar{x} = \frac{[x]}{n}.$$

2. Вычисляют отклонения от арифметической средины $v_i = x_i - \bar{x}$. Для контроля подсчитывают $[ v ] = 0$.

3. Определяют величину $[ vv ]$.

4. Вычисляют средние квадратические ошибки результатов измерений и полученных значений по формулам (1.71), (1.72) и (1.58).

**Пример 36.** По результатам измерения угла, приведенным в табл. 1, найти наиболее надежное значение этого угла и оценить точность результатов измерений и вычисленного вероятнейшего значения угла.

**Решение.** Выполнять обработку результатов измерений рекомендуется согласно табл. 7.

Таблица 7

| № п.п. | Значение угла $x_i$ | Отклонения $v_i$, сек | $v_i\, v_i$ |
|---|---|---|---|
| 1 | 2 | 3 | 4 |
| 1 | 81° 35′ 26″ | +2,2 | 4,8 |
| 2 | 81 35 32 | -3,8 | 14,4 |
| 3 | 81 35 24 | +4,2 | 17,6 |
| 4 | 81 35 28 | +0,2 | 0 |
| 5 | 81 35 33 | -4,8 | 23,0 |
| 6 | 81 35 25 | +3,2 | 10,2 |
| 7 | 81 35 31 | -2,8 | 7,8 |
| 8 | 81 35 22 | +6,2 | 38,4 |
| 9 | 81 35 34 | -5,8 | 33,6 |
| 10 | 81 35 29 | -0,8 | 0,6 |
| 11 | 81 35 25 | +3,2 | 10,2 |
| 12 | 81 35 30 | -1,8 | 3,2 |
| $\bar{x}$ = | 81° 35′ 28,2″ | $[ v ] = -0.6$ | $[ vv ] = 163,8$ |

Вычисляем вероятнейшее значение угла:

$$\bar{x} = \frac{[x]}{n} = 81^\circ 35' 28{,}2''.$$

Найдем среднюю квадратическую ошибку результата измерения угла:

$$m=\sqrt{\frac{[vv]}{n-1}}=\sqrt{\frac{163,8}{12-1}}=3,9''.$$

Оценим надежность вычисления средней квадратической ошибки:

$$m_m=\frac{m}{\sqrt{2(n-1)}}=\frac{3,9}{\sqrt{2(12-1)}}=0,83''.$$

Определим среднюю квадратическую ошибку нахождения вероятнейшего значения угла:

$$m_\beta=\frac{m}{\sqrt{n}}=\frac{3,9}{\sqrt{12}}=1,1''.$$

Ответ**:** $\bar{x}=81^{\circ}35'28,2''\pm1,1''$.

## 5.2. Порядок обработки результатов неравноточных измерений одной величины

В предыдущем параграфе рассмотрено, как решается задача обработки ряда равноточных измерений одной и той же величины. Пусть теперь имеем ряд независимых измерений $x_1$, $x_2$, ... , $x_n$ с весами, соответственно равными $p_1$, $p_2$, ... , $p_n$. В этом случае для математического ожидания наилучшей оценкой будет являться общая арифметическая средина.

1. Устанавливают веса согласно выражениям:

$$p_i=\frac{C}{m_i^2} \quad \text{и} \quad p_i=\frac{\mu^2}{m_i^2}$$

или в зависимости от видов работ и применяемых методов измерений.

2. Вычисляют наиболее надежное значение согласно формуле:

$$\bar{x}=\frac{[px]}{[p]}.$$

3. Вычисляем отклонения от арифметической средины, которые обладают свойствами $[pv]=0$ и $[pvv]=\min$.

4. Вычисляют значение ошибки единицы веса, ее надежность, а также среднюю квадратическую ошибку наиболее надежного значения согласно формулам:

- ошибку единицы веса $\mu = \sqrt{\dfrac{[pvv]}{n-1}}$;

- надежность полученной ошибки $m_\mu = \dfrac{\mu}{\sqrt{2(n-1)}}$;

- среднюю квадратическую ошибку вероятнейшего значения $M = \dfrac{\mu}{\sqrt{P}}$.

**Задача 37**. Произвести математическую обработку по данным, приведенным в табл. 8, в которой даны средние значения одного и того же угла, измеренного разным числом приемов $n$. При установлении весов использовать следующее соотношение $p = n/3$. Вычисления располагаем в этой же таблице.

Таблица 8

| № пп | Значения угла $x_i$ | Число приемов $n$ | $P$ = n/3 | $v_i$, с | $p_i v_i$, с | $p_i v_i v_i$ |
|---|---|---|---|---|---|---|
| 1 | 89°47′16″ | 6 | 2 | +6 | +12 | 72 |
| 2 | 9 | 18 | 6 | -1 | -6 | 6 |
| 3 | 6 | 3 | 1 | -4 | -4 | 16 |
| 4 | 10 | 15 | 5 | 0 | 0 | 0 |
| 5 | 23 | 6 | 2 | +3 | +6 | 18 |
| 6 | 8 | 12 | 4 | -2 | -8 | 16 |
| $x_0$ | 89°47′06″ | | 20 | | -18<br>+18 | 128 |

**Решение.** Общая арифметическая средина вычисляется по формуле

$$\bar{x}_0 = \frac{[px]}{[p]} = 89^o 47'06''.$$

После вычисления отклонений от арифметической средины находят [ $pv$ ] и [ $pvv$ ].

Тогда ошибка единицы веса

$$\mu = \sqrt{\frac{[pvv]}{n-1}} = \sqrt{\frac{128}{6-1}} = 5{,}1''.$$

При этом надежность вычисления ошибки единицы веса составит

$$m_\mu = \frac{\mu}{\sqrt{2(n-1)}} = \frac{5{,}1}{\sqrt{2(6-1)}} = 1{,}6''.$$

Средняя квадратическая ошибка определения наиболее надежного результата

$$M = \frac{\mu}{\sqrt{[p]}} = \frac{5{,}1''}{\sqrt{20}} = 1{,}1''.$$

Ответ: 89° 47′ 10″ ± 1,1″.

## 5.3. Оценка точности по разностям двойных равноточных измерений

В большинстве случаев при производстве геодезических работ измерения дублируются, что приводит к накоплению информации, которую можно использовать для характеристики точности фактических результатов.

Даны результаты двойных равноточных измерений однородных величин $X_1, X_2, \ldots, X_n$:

$$\begin{aligned} &x_1', x_2', \ldots, x_n'; \\ &x_1'', x_2'', \ldots, x_n''. \end{aligned} \tag{5.1}$$

Составим разности

$$\begin{aligned} &x_1' - x_1'' = d_1; \\ &x_2' - x_2'' = d_2; \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ &x_n' - x_n'' = d_n. \end{aligned} \tag{5.2}$$

Истинное значение разности соответствует нулевому значению, так как, если бы измерения были безошибочными, то разности равнялись бы нулю. Следовательно, разности $d_i$ являются истинными ошибками. Поэтому среднюю квадратическую ошибку разности определим согласно формуле Гаусса, т.е.

$$m_d = \sqrt{\frac{[dd]}{n}}\,, \tag{5.3}$$

где $n$ - число всех разностей.

Средняя квадратическая ошибка разности двух равноточных измерений в корень квадратный из двух раз больше ошибки одного измерения, т.е.

$$m_d = m \cdot \sqrt{2}\,,$$

откуда

$$m = \sqrt{\frac{[dd]}{2n}}\,. \tag{5.4}$$

Учитывая, что $\overline{x} = \dfrac{x_i' + x_i''}{2}$ является наиболее надежным значением величины $X_i$ , то

$$M = \frac{m}{\sqrt{2}} = \frac{1}{2}\sqrt{\frac{[dd]}{n-1}}. \qquad (5.5)$$

Наличие систематических ошибок в разностях двойных измерений можно обнаружить либо по знаку, либо по величине среднего арифметического значения разностей. В качестве критерия влияния систематических ошибок используют следующее условие:

$$|[d]| \le 0{,}25[|d|]. \qquad (5.6)$$

В том случае, если условие (5.6) не соблюдается, тогда вычисляют систематическую составляющую полученных разностей

$$d_0 = \frac{[d]}{n}. \qquad (5.7)$$

Исправляя на эту величину наши разности, получим

$$\begin{aligned} &d_1 - d_0 = d_1'; \\ &d_2 - d_0 = d_2'; \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ &d_n - d_0 = d_n'. \end{aligned} \qquad (5.8)$$

Поскольку $d_0$ является средним арифметическим значением разностей, то $d_i'$ можно рассматривать как отклонения от среднего арифметического. Для подтверждения этого рассмотрим ряд наших разностей (5.8). Просуммировав данные равенства, получим

$$[d] - nd_0 = [d'].$$

В полученное равенство вместо $d_0$ подставим его значение из равенства (5.7)

$$[d] - n\frac{[d]}{n} = [d'].$$

Откуда

$$[d'] = 0. \qquad (5.9)$$

Таким свойством обладают только отклонения от среднего арифметического. Следовательно, для нахождения средней квадратической ошибки разности можно использовать формулу Бесселя

$$m_d = \sqrt{\frac{[d'd']}{n-1}} \qquad (5.10)$$

или

$$m = \sqrt{\frac{[d'd']}{2(n-1)}}. \qquad (5.11)$$

**Пример 38**. Производились двойные измерения одних и тех же линий. Оценить точность работ по разностям двойных измерений, полагая, что измерения равноточные. Все исходные данные и результаты промежуточных вычислений расположены в табл. 9.

**Решение.** Критерий допустимости $d_0$

$$\left|[d]\right| \le 0{,}25\left[|d|\right].$$

Таблица 9

| № пп | Длины линий | | Разность $d_i$, см | $d'_i$, см | $d'_i d'_i$ |
|---|---|---|---|---|---|
| | $x_i$ | $x_i$ | | | |
| 1 | 967,490 | 967,399 | +9,1 | +5,7 | 32,5 |
| 2 | 752,468 | 752,412 | +5,6 | +2,2 | 4,8 |
| 3 | 692,223 | 692,250 | -2,7 | -6,1 | 37,2 |
| 4 | 1023,536 | 1023,536 | 0 | -3,4 | 11,6 |
| 5 | 808,457 | 808,444 | +1,3 | -2,1 | 4,4 |
| 6 | 612,692 | 612,665 | +2,7 | -0,7 | 0,5 |
| 7 | 675,158 | 675,082 | +7,6 | +4,2 | 17,2 |
| | | | +26,3 | 0,2 | 108,6 |
| | | | -2,7 | | |
| | | $[d]=$ | +23,6 | | |
| | | $[\|d\|]=$ | 29,0 | | |

В нашем случае

$$\left|[d]\right| = 23{,}6 > 0{,}25\left[|d|\right] = 7{,}2;$$

следовательно, в наших разностях имеет место систематическая составляющая, величина которой определяется как

$$d_0 = \frac{[d]}{n} = \frac{23{,}6}{7} = +3{,}4 \text{ см.}$$

Все разности исправляются на полученную величину. После этого определяют $[d'd']$ и переходят к оценке точности.

Средняя квадратическая ошибка результата измерения составит

$$m = \sqrt{\frac{[d'd']}{2(n-1)}} = \sqrt{\frac{108{,}6}{12}} = 3{,}0 \text{ см.}$$

Средняя квадратическая ошибка вероятнейшего значения

$$M = \frac{1}{2}\sqrt{\frac{[d'd']}{n-1}} = \frac{1}{2}\sqrt{\frac{108{,}6}{6}} = 2{,}1 \text{ см.}$$

Ответ: $m$ = 3,0 см; $M$ = 2,1 см.

## 5.4. Оценка точности по результатам однородных двойных неравноточных измерений

Пусть даны результаты двойных неравноточных измерений $x'_i$ и $x''_i$ с соответствующими весами $p'_i$ и $p''_i$. Определим разности

$$d_i = x'_i - x''_i .$$

Веса разностей определятся соответственно

$$\frac{1}{p_{d_i}} = \frac{1}{p'_i} + \frac{1}{p''_i} \qquad \text{или} \qquad \frac{1}{p_{d_i}} = \frac{p'_i + p''_i}{p'_i \cdot p''_i} .$$

Откуда

$$p_{d_i} = \frac{p'_i \cdot p''_i}{p'_i + p''_i}. \tag{5.12}$$

Рассматривая разности как истинные ошибки, ошибку единицы веса можно найти

$$\mu = \sqrt{\frac{[p_d dd]}{n}}. \tag{5.13}$$

Если $p'_i = p''_i$, то $p_{d_i} = \frac{p_i}{2}$, тогда

$$\mu = \sqrt{\frac{[pdd]}{2n}}. \tag{5.14}$$

При наличии систематических ошибок в разностях двойных неравноточных измерений

$$d_0 = \frac{[pd]}{[p]}. \tag{5.15}$$

Тогда

$$\mu = \sqrt{\frac{[p_d d'd']}{n-1}} \tag{5.16}$$

и соответственно

$$\mu = \sqrt{\frac{[pd'd']}{2(n-1)}}\,. \tag{5.17}$$

Критерием применимости формул (5.13) и (5.14) служит условие

$$\left|\left[d\sqrt{p}\right]\right| \le 0{,}25\cdot\left[\left|d\sqrt{p}\right|\right]\,. \tag{5.18}$$

В противном случае используют формулы (5.16) и (5.17), в которых

$$d_i' = d \; - d_0\,.$$

**Пример 39.** По разностям двойных измерений оценить точность геометрического нивелирования. Исходные данные и результаты вычислений представлены в табл. 10.

Таблица 10

| № пп | Разность $d_i$, мм | Число станций $n$ | Веса $p$=10/$n$ | $d_i\sqrt{p_i}$ | $p_i d_i$ | $p_i d_i d_i$ |
|---|---|---|---|---|---|---|
| 1 | +4 | 7 | 1,43 | +4,8 | +5,7 | 22,8 |
| 2 | -14 | 27 | 0,37 | -8,5 | -5,2 | 72,8 |
| 3 | -9 | 13 | 0,77 | -7,9 | -6,9 | 62,1 |
| 4 | +15 | 25 | 0,40 | +9,5 | +6,0 | 90,0 |
| 5 | -12 | 32 | 0,31 | -6,7 | -3,7 | 44,4 |
| 6 | +11 | 15 | 0,67 | +9,0 | +7,4 | 81,4 |
| 7 | -12 | 19 | 0,53 | -8,7 | -6,7 | 76,8 |
| 8 | +13 | 18 | 0,56 | +9,7 | +7,3 | 94,9 |
| 9 | +12 | 16 | 0,62 | +9,5 | +7,4 | 88,8 |
| 10 | -7 | 23 | 0,43 | -4,6 | -3,0 | 21,0 |
| | | | 6,09 | +42,5<br>-36,4<br>+6,1<br>78,9 | | 655,0 |

**Решение.** Наличие систематических ошибок проверяем с помощью критерия

$$\left|\left[d\sqrt{p}\right]\right| = 6{,}1 < 0{,}25\left[\left|d\sqrt{p}\right|\right] = 19{,}0\,.$$

Следовательно, систематическим влиянием можно пренебречь.

Средняя квадратическая ошибка определения превышения одиночного хода из 10 станций составила

$$\mu_{10} = \sqrt{\frac{[pdd]}{2n}} = \sqrt{\frac{655}{2 \cdot 10}} = 5{,}7 \text{ мм.}$$

Средняя квадратическая ошибка определения превышения двойного нивелирного хода

$$\mu = \frac{1}{2}\sqrt{\frac{[pdd]}{n}} = \frac{1}{2}\sqrt{\frac{655}{10}} = 4{,}0 \text{ мм.}$$

Ответ: $\mu = 4{,}0$ мм .

## 6. Статистическая оценка параметров распределения
### 6.1. Понятие об оценке параметров

Вся подлежащая для изучения однородная совокупность значений случайной величины называется генеральной совокупностью. Случайно отобранная часть этой совокупности называется выборкой или выборочной совокупностью.

В том случае, если нам необходимо оценить математическое ожидание распределения случайной величины, то возникает задача определения выборочной характеристики, которая позволила бы получить наиболее полное представление об интересующем нас параметре. Поскольку сама по себе выборка случайна, то и суждение об интересующем нас параметре могут быть необъективными. Чем большего объема выборка, тем и лучше представление об исследуемом параметре.

Поэтому, чтобы оценить действительное значение какой-либо измеряемой величины, а так же произвести оценку точности полученных результатов измерений, необходимо использовать статистическое оценивание этих результатов. Полагая, что полученные в результате измерений значения являются случайными величинами и подчиняются нормальному закону распределению, определяют параметры этого распределения, которыми являются математическое ожидание и стандарт. Первый параметр характеризует количественную сторону измерений, т.е. истинное значение измеряемой величины. Второй параметр – качественную сторону, т.е. точность выполненных измерений.

Статистическая обработка результатов измерений позволяет установить приближенные значения исследуемых параметров распределения. Следовательно, возникает задача оценки точности вычисленных в результате обработки значений измеряемой величины. Термин «оценка» содержит в своем названии двойной смысл: с одной стороны позволяет судить об истинном значении измеряемой величины и ее среднем квадратической отклонении, а с другой дает возможность оценить точность выполненных измерений, а следовательно, является критерием точности полученных значений. Точность статистических оценок обеспечивается с принятой вероятностью, которая задает уровень надежности. Надежность и точность статистических оценок взаимосвязаны. При исследовании ряда распределения можно рассматривать два вида оценок – точечную и интервальную.

Приближенные значения, полученные при исследовании результатов измерений, являются оценками или статистиками. При оценке математического ожидания статистикой является арифметическое среднее, а среднего квадратического отклонения – средняя квадратическая ошибка. Причем при оценке одного и того же параметра можно воспользоваться несколькими статистиками. Так для оценки математического ожидания можно использовать арифметической срединой, медианой, модой, начальным моментом первого порядка. Как видим, возникает задача выбора той или иной статистики, при этом имеют в виду, что выбранная статистика будет наилучшей. Статистика должна обладать следующими свойствами: несмещенности, эффективности, состоятельности и достаточности.

Оценка $\hat{\Theta}$ параметра $\Theta$ называется несмещенной, если ее математическое ожидание равно оцениваемому параметру, т.е.

$$M\left(\hat{\Theta}\right)=\Theta. \tag{6.1}$$

В этом случае систематическая составляющая должна отсутствовать.

Свойство эффективности состоит в том, что несмещенная статистика $\hat{\Theta}$ будет наилучшей, если ее дисперсия будет наименьшей.

Свойством состоятельности обладает статистика, которая при неограниченном возрастании числа измерений сходится по вероятности к оцениваемому параметру.

Статистика считается достаточной, если в ней содержится вся необходимая информация о оцениваемом параметре. Среднее арифметическое и средняя квадратическая ошибка являются достаточными статистиками оценки математического ожидания стандарта.

### 6.2. Точечная оценка параметров

В том случае, когда осуществляется выборка большого объема ($n$>30), то для оценки параметров распределения пользуются статистиками, удовлетворяющими условию состоятельности. К таким статистикам относятся эмпирические моменты. Если дан ряд равноточных результатов измерений $x_1$, $x_2$, … , $x_n$ некоторой величины $X$, распределенной нормально, то оценка неизвестного параметра математического ожидания с помощью простой арифметической срелины.

Средняя квадратическая ошибка результата измерения, вычисляемая по формуле (4.25) является статистической оценкой стандарта $\sigma(x)$, т.е. $m \approx \sigma(x)$. В этом случае стандарт среднего квадратического отклонения составит согласно

$$\sigma(m) = \frac{\sigma(x)}{\sqrt{2n}} \approx m_m = \frac{m}{\sqrt{2n}}. \qquad (6.2)$$

Применение при оценке точности результатов измерения формулы Бесселя приводит к использованию следующей статистики

$$\sigma(m) = \frac{\sigma(x)}{\sqrt{2(n-1)}} \approx m_m = \frac{m}{\sqrt{2(n-1)}}. \qquad (6.3)$$

Приведенные формулы для вычисления $\sigma(m)$ используются в случае нормального распределения ошибок измерения. В случае иного распределения ошибок измерения применяются соответствующие формулы:

$$\sigma(m) \approx \frac{\sigma(x)}{\sqrt{2n}}\sqrt{1+0{,}5E_s(x)}; \qquad (6.4)$$

$$\sigma(m) \approx \frac{\sigma(x)}{\sqrt{2(n-1)}}\sqrt{1+0{,}5E_s\frac{n-1}{n}}\,, \qquad (6.5)$$

где $E_s(x)$ – величина эксцесса.

Нормальное распределение обладает свойством устойчивости. При этом результаты распределены нормально, а следовательно, и функция этих результатов также имеют нормальное или почти нормальное распределение.

Точечная оценка в любом случае является приближенной. Она имеет смысл лишь в том случае, если определены границы интервала, накрывающего действительный параметр распределения. При этом необходимо задаваться определенным уровнем надежности, который является доверительным.

### 6.3. Оценка параметров по малым выборкам. Понятие доверительного интервала

Ранее рассмотрены способы оценки с помощью выборочных числовых характеристик. Но такая оценка может быть оправдана только при больших выборках (при большом числе измерений), так как при большом числе испытаний среднее арифметическое сходится по вероятности к математическому ожиданию. То же самое можно сказать и об оценке дисперсий. При малых выборках значение оцениваемого параметра можно получить лишь приближенно. Ошибка в оценке будет тем больше, чем меньше выборка. Следовательно, при малых выборках оценка параметров без указания степени точности и надежности мало определена. Поэтому при малой выборке применяют метод оценивания неизвестных параметров с помощью доверительного интервала.

Чтобы получить представление о точности и надежности оценки $\overline{\Theta}$ для параметра Θ, можем для каждого малого $q > 0$ указать такое $\varepsilon$, что

$$P\left(\overline{\Theta} - \varepsilon < \Theta < \overline{\Theta} + \varepsilon\right) = 1 - q\ . \qquad (6.6)$$

Чем меньше $\varepsilon$, тем точнее наша оценка Θ. Из соотношения (3.19) следует, что вероятность того, что интервал $(\Theta - \varepsilon, \Theta + \varepsilon)$ покроет неизвестный параметр, равна $p$ = 1 - $q$. Такой интервал называют доверительным. Вероятность $p$ = 1 - $q$ принято называть доверительной вероятностью.

Остановимся на вопросе о том, какими критериями следует руководствоваться при выборе доверительной вероятности. Выбор доверительной вероятности не является математической задачей, а определяется конкретно решаемой проблемой. Приведем пример: пусть на двух предприятиях вероятность выпуска годных изделий $p$ = 1 - $q$ = 0,99, т.е. вероятность выпуска бракованных изделий $q$ = 0,01. Можно ли в рамках математической теории, т.е. не интересуясь характером выпускаемой продукции, решить вопрос о том, мала или велика вероятность $q$?

Пусть одно предприятие выпускает электролампы, а другое - парашюты. В первом случае можно мириться, что на сто электроламп одна бракованная, а во втором случае - нет. Следовательно, в первом случае вероятность брака $q$ приемлема, а во втором - нет. Поэтому выбор доверительной вероятности $p$ = 1 - $q$ следует производить, исходя из конкретных условий задачи.

### 6.4. Доверительные интервалы для математического ожидания

Пусть требуется по данным выборки оценить параметр $a$ - центр группирования генеральной совокупности. Установив закон распределения, нетрудно определить и вероятность $p = 1 – q$ предположения, что разность $\bar{x}$ - $a$ не выйдет по абсолютному значению за некоторые установленные пределы

$$P\left(\left|\bar{x}-a\right| \le \varepsilon\right) = 1 - q. \tag{6.7}$$

Пользуясь этим неравенством, можно решать и обратную задачу - по выборочному значению $\bar{x}$ установить интервал, который с вероятностью $p$ = 1 - $q$ перекроет “истинное” значение параметра $a$. Значит нам необходимо оценить параметр $a$ при помощи $\bar{x}$, если известно $\sigma(x)$.

Если результатам измерений сопутствуют случайные ошибки, обладающие нормальным распределением, то статистика $\bar{x}$ подчинена нормальному закону с параметрами распределения

$$M(x) = a; \qquad \sigma(\bar{x}) = \frac{\sigma(x)}{\sqrt{n}}. \tag{6.8}$$

Так как $\sigma(x)$ известно, то воспользуемся нормированной величиной

$$t = \frac{\bar{x} - a}{\sigma(x)}, \tag{6.9}$$

которое имеет распределение, не зависящее от каких-либо параметров, т.к. плотность этого распределения

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

Пользуясь таблицей нормированной функции Лапласа, определяем параметр $t$, а затем, учитывая

$$\bar{x} - a = t \cdot \sigma(\bar{x}) = \varepsilon ,$$

будем иметь

$$P\left\{\bar{x} - t \cdot \sigma(\bar{x}) < a < \bar{x} + t \cdot \sigma(\bar{x})\right\} = p = \Phi(t) \tag{6.10}$$

или

$$P\left\{\bar{x} - t\frac{\sigma(x)}{\sqrt{n}} < a < \bar{x} + t\frac{\sigma(x)}{\sqrt{n}}\right\} = p = \Phi(t). \tag{6.11}$$

Таким образом, интервал $\left(\bar{x} - t\frac{\sigma(x)}{\sqrt{n}}, \bar{x} + t\frac{\sigma(x)}{\sqrt{n}}\right)$ будет доверительным интервалом для оценки $a$, отвечающей доверительной вероятности $p = 1 - q$ . Обычно доверительную вероятность или уровень значимости выбирают равным 0,95; 0,99; 0,997 и т.д.

Рассмотрим случай, когда среднее квадратическое отклонение $\sigma(x)$ не известно. Требуется оценить неизвестное математическое ожидание $M(x) = a$. Используем вместо $\sigma(x)$ эмпирический стандарт или среднюю квадратическую ошибку $m$

$$m = \sqrt{\frac{\sum_{1}^{n}(x - \bar{x})^2}{n - 1}} ,$$

при этом

$$\bar{x} - t\frac{m}{\sqrt{n}} < a < \bar{x} + t\frac{m}{\sqrt{n}}, \qquad (6.12)$$

где $t = f(p, k)$; $k = n - 1$ – число степеней свободы.

Нормированная случайная величина $t_{p,k}$ распределена по закону Стьюдента и определяется по специальным таблицам распределения (табл. 2 приложения). Распределение Стьюдента является приближением к нормальному и предназначено для малого числа измерений.

Следовательно, доверительная вероятность для оцениваемого параметра составит

$$P\left(\bar{x} - t_{p,k}\frac{m}{\sqrt{n}} < a < \bar{x} + t_{p,k}\frac{m}{\sqrt{n}}\right) = 1 - q. \qquad (6.13)$$

Поэтому с вероятностью (надежностью) $p = 1 - q$ можно утверждать, что интервал

$$\left[\bar{x} - t_{p,k}\frac{m}{\sqrt{n}}, \bar{x} + t_{p,k}\frac{m}{\sqrt{n}}\right] \qquad (6.14)$$

является доверительным для оценки математического ожидания $a$.

Несмотря на кажущееся сходство формул (6.11) и (6.13), между ними суествует различие, заключающееся в том, что в формуле (6.13) параметр $t_{p,k}$ зависит не только от доверительной вероятности, но и от количества элементов в выборке. Особенно это различие заметно при малом количестве наблюдений.

**Пример 40.** Случайная величина $X$ имеет нормальное распределение с известным средним квадратическим отклонением $\sigma = 3$. Найти доверительный интервал для оценки неизвестного математического ожидания $M(x) = a$ по выборочной средней $\bar{x} = 4{,}1$, если объем выборки $n = 36$ и задана доверительная вероятность $p = 0{,}95$.

**Решение.** Из табл. 1 приложения для $\Phi(t) = 0{,}95$ находим параметр $t = 1{,}96$. Поскольку точность выполнения работ известна (выборка большого объема), для определения доверительного интервала воспользуемся соотношением (6.11), т.е.

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

тогда искомый интервал определится

$$4{,}1-\frac{1{,}96\cdot 3{,}0}{\sqrt{36}} < a < 4{,}1+\frac{1{,}96\cdot 3{,}0}{\sqrt{36}}\ ;$$

$$4{,}1 - 0{,}98 < a < 4{,}1 + 0{,}98\ ;$$

$$3{,}12 < a < 5{,}08\ .$$

Ответ: $3{,}12 < a < 5{,}08$ .

Так как $\overline{x}$ является случайной величиной, возможные значения которой меняются от выборки к выборке, то концы доверительного интервала $(\overline{x}-\varepsilon,\ \overline{x}+\varepsilon)$ также являются случайными величинами, меняющимися при изменении выборки.

В нашем примере выбрана доверительная вероятность $p = 0{,}95$. Это значит, что если произведено достаточно большое число наблюдений, то 95% из полученных результатов определяют такие доверительные границы, внутри которых будет находиться оцениваемый параметр $a$, и лишь 5% интервалов его не содержат. Поэтому было бы ошибкой утверждать, что математическое ожидание заключено в данном интервале с доверительной вероятностью $p = 0{,}95$. Математическое ожидание или попадает в интервал и тогда вероятность этого события равна единице, или находится вне его, вероятность в этом случае равна нулю.

Следовательно, доверительную вероятность не следует связывать с оцениваемым параметром - она связана лишь с границами интервала. определяющими случайность выборки.

**Пример 41.** Случайная величина $X$ наблюдалась 10 раз. Найти доверительный интервал для математического ожидания $M(x) = a$ с доверительной вероятностью $p = 0{,}99$. Вычисленные значения среднего арифметического из результатов измерений $\overline{x} = 10{,}78$, а среднее квадратическая ошибка $m = 0{,}25$.

**Решение**. Согласно табл. 2 приложения по доверительной вероятности $p = 0{,}95$ и степени свободы $k = 10 - 1 = 9$ находим параметр $t_{p,k} = 3{,}25$, тогда

$$\left(\overline{x}-t_{p,k}\frac{m}{\sqrt{n}} < a < \overline{x}+t_{p,k}\frac{m}{\sqrt{n}}\right)$$

или

$$\left(10{,}78-3{,}25\frac{0{,}25}{\sqrt{10}} < a < 10{,}78-3{,}25\frac{0{,}25}{\sqrt{10}}\right)\ ;$$

$$10{,}52 < a < 11{,}04.$$

Если воспользоваться нормированной функцией Лапласа, то по функции $\Phi(t) = 0{,}99$ находим параметр $t = 2{,}58$ и согласно (6.11) имеем

$$10{,}78 - 2{,}58 \cdot 0{,}079 < a < 10{,}78 + 2{,}58 \cdot 0{,}079;$$

$$10{,}58 < a < 10{,}98.$$

Ответ: $10{,}52 < a < 11{,}04$; $10{,}58 < a < 10{,}98$.

Как видно из приведенного примера распределение Стьюдента дает более широкий интервал, чем нормальное распределение. Это свидетельствует не о его слабости, а о малости выборки. Использование нормального распределения неоправданно суживает интервал, т.е. завышает точность измерений, занижая дисперсию.

### 6.5. Доверительная оценка среднего квадратического отклонения

Предполагая, что выборка произведена из нормальной совокупности, например, что $n$ независимых наблюдений $x_1, x_2, \ldots, x_n$ являются значениями нормально распределенной величины $X$, построим доверительный интервал для параметров $\sigma$ или $\sigma^2$.

Как известно, оценку параметра $\sigma^2$, когда $a$ неизвестно, вычисляют по формуле

$$m_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2, \tag{6.15}$$

а при известном $a$

$$m^2 = s^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - a\right)^2. \tag{6.16}$$

Построение доверительного интервала для дисперсии основано на том, что случайная величина $ns^2/\sigma^2$ имеет распределение $\chi^2$ с $k = n$ степенями свободы, а величина $n\,m_x^2/\sigma^2$ имеет распределение $\chi^2$ с $k = n - 1$ степенями свободы.

Подробно рассмотрим построение доверительного интервала для второго случая, как наиболее часто встречающегося в геодезической практике. Итак для выбранной доверительной вероятности $p$ = 1 - $q$, учитывая, что $n\,m_x^2/\sigma^2$ имеет распределение $\chi^2$ с $k$ = $n$ - 1 степенями свободы, можно определить

$$P\left(\chi_1^2 < \frac{nm_x^2}{\sigma^2} < \chi_2^2\right) = 1 - q \qquad (6.17)$$

или

$$P\left(\frac{nm_x^2}{\chi_2^2} < \sigma^2 < \frac{nm_x^2}{\chi_1^2}\right) = 1 - q\ . \qquad (6.18)$$

Доверительный интервал находится с доверительной вероятностью $p$ = 1 - $q$. Для заданного значения $q$ и степеней свободы $k$ = $n$ - 1 при определении величин $\chi_1^2$ и $\chi_2^2$ пользуются таблицами функции $\chi^2$ - распределения (табл.3 приложения). При этом значения $\chi_1^2$ находят для $p_1$ = 1 - 0,5$q$, а значения $\chi_2^2$- для $p_2$ = 0,5$q$.

При $n$ > 30 можно воспользоваться асимптотическим свойством $\chi^2$ – распределения, утверждающим, что величина

$$t = \sqrt{2\chi^2} - \sqrt{2n-1} \qquad (6.19)$$

подчиняется нормальному распределению с параметрами $M(t)$ = 0 и $\sigma^2(t)$ = 1. По формуле Лапласа имеем

$$P\big[|T| < t\big] = P(-t < T < t) = \Phi(t) = 1 - q\ , \qquad (6.20)$$

где $t$ определяется согласно (3.32), тогда

$$\chi_{1,2}^2 = \frac{1}{2}\left(\sqrt{2n-1} \pm t\right)^2\ . \qquad (6.21)$$

В дальнейшем для интервальной оценки воспользуемся выражением (6.18).

**Приме 42.** Построить доверительный интервал с вероятностью $p$ = 0,96 для дисперсии $\sigma^2$ случайной величины $X$, распределенной нормально, если $m^2$= 10,0 и $n$ = 20.

**Решение.** Поскольку доверительная вероятность $p$ = 1 - $q$ = 0,96, то $q$ = 0,04. Тогда для определения значений $\chi_1^2$ и $\chi_2^2$ установим вероятности $p_1$ = 1 - 0,5·0,04 = 0,98; $p_2$ = 0,5·0,04=0,02 и степень свободы $k$ = 20 - 1 = 19.

Согласно табл. 3 приложения $\chi_1^2 = 8{,}6$ и $\chi_2^2 = 33{,}7$. Подставляя численные значения в равенство (6.18), получим искомый доверительный интервал

$$\frac{20 \cdot 10{,}0}{33{,}7} < \sigma^2 < \frac{20 \cdot 10{,}0}{8{,}6}$$

или

$$5{,}935 < \sigma^2 < 23{,}256 .$$

Вычислим доверительный интервал для среднего квадратического отклонения

$$\sqrt{5{,}935} < \sigma < \sqrt{23{,}256}$$

или

$$2{,}43 < \sigma < 4{,}82 .$$

Ответ: $2{,}43 < \sigma < 4{,}82$ .

## 7. Проверка статистических гипотез

### 7.1. Понятие статистической гипотезы

В ряде случаев, возникающих при обработке результатов измерений, необходимо на основе известных реализаций установить, подчиняется та или иная случайная величина определенному закону распределения.

Статистической гипотезой называется всякое предположение как о законе распределения, так и об отдельных параметрах распределения случайной величины. В математической статистике разработаны различные приемы проверки статистических гипотез. В теории ошибок измерений их применяют для проверки:

1) случайного характера ошибок и выяснения действий систематических ошибок;

2) равноточности результатов измерений и их зависимости между собой или с другими параметрами, характеризующими условия измерений;

3) предполагаемого закона распределения случайных ошибок и нормальности их распределения.

Современная теория ошибок при оценке точности основывается на теории доверительных интервалов и способах проверки статистических гипотез. Вначале производится предварительная обработка результатов измерения какой-либо величины, а затем, исходя из теоретических предпосылок или приближенного анализа эмпирического материала, намечается определенная статистическая гипотеза, которая в дальнейшем сопоставляется с имеющимся эмпирическим материалом. Выдвинутую гипотезу обычно называют нулевой гипотезой и обозначают $H_0$. Если опытные данные не противоречат гипотезе, она принимается, а если она с ней не согласуется - гипотеза $H_0$ отвергается. Оценка соответствия статистической гипотезы опытным данным производится путем применения определенных правил, зависящих от характера выдвинутой гипотезы и называемых статистическими критериями. При этом используют различного рода критерии, основанные на вероятностной оценке имеющих место отклонений эмпирических данных от принятой гипотезы. Проверка состоит в определении значений статистических характеристик согласно результатов измерений и в сравнении их с тем, что было выведено на основе сделанных предположений. Такие характеристики называются критериями проверки статистических гипотез.

Для проверки нулевой гипотезы применяют статистику, распределение которой в условиях этой гипотезы известно. В зависимости от величины критериев принимают или отвергают нулевую гипотезу. При этом используют достаточно малую вероятность $q$ осуществления данного события. Вероятность $q$ называют уровнем значимости. Дополнением уровня значимости является область допустимых значений или доверительная вероятность $p = 1 - q$ (рис. 16).

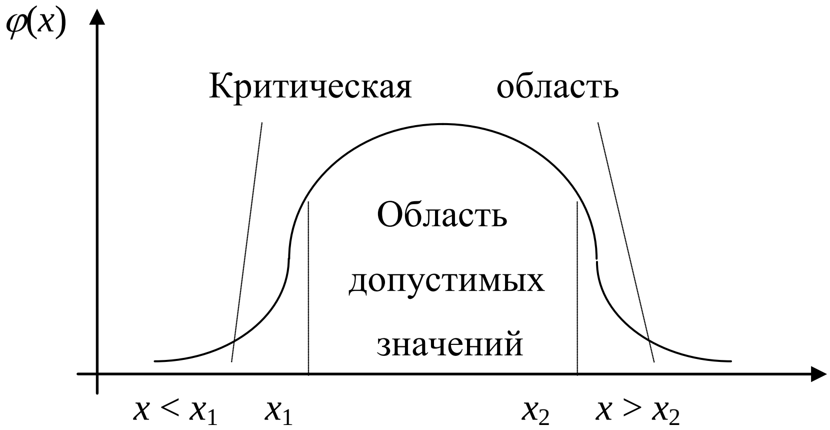


Рис. 16

Для установленного значения $q$ определяется критическая область значений критерия проверки (рис. 16). Если критерий попадает в данную область, то нулевая гипотеза $H_0$ отвергается. Если критерий попадает в область допустимых значений (или область нулевой гипотезы), то нулевая гипотеза принимается, вернее говоря, она не противоречит результатам измерений. Чем меньше уровень значимости, тем меньше вероятность отвергнуть (забраковать) принятую гипотезу.

### 7.2. Сравнение средних значений

Целью эксперимента является определение различий между значениями какого-либо параметра в различных объектах исследования. При этом необходимо установить, не является ли это различие следствием лишь случайных ошибок измерений. Для установления случайного или неслучайного характера расхождений производят обычно две серии наблюдений, для которых определяют эмпирические средние $\bar{x}_1$ и $x_2$. При этом возникает вопрос, считать ли расхождение между найденными параметрами достаточно большим, чтобы с уверенностью говорить о неслучайном характере этих различий. Часто в геодезии возникает вопрос о сравнении центров распределения двух случайных величин $X_1$ и $X_2$.

Пусть имеем две выборки объемов соответственно $n_1$ и $n_2$, при этом вычисляются средние значения $\bar{x}_1$ и $\bar{x}_2$ . Проверим нулевую гипотезу $H_0$ о равенстве математических ожиданий, т.е. $a_1 = a_2$ .

Рассмотрим два случая: первый заключается в том, что средние квадратические отклонения (стандарты) $\sigma_1^2$ и $\sigma_2^2$ известны. Проверим гипотезу $H_0$ , имея в виду, что значение $\overline{X}_1 - \overline{X}_2$ выборочных средних подчиняется нормальному распределению, тогда

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \ . \qquad (7.1)$$

В качестве критерия проверки используем нормированную величину $t$ разности $\left|\overline{x}_1 - \overline{x}_2\right|$

$$t = \frac{\overline{x}_1 - \overline{x}_2}{\sigma_{\overline{x}_1 - \overline{x}_2}} . \tag{7.2}$$

Критическая область определяется областью больших значений $|\,t\,| > t_q$, где $t_q$ соответствует $q$-уровню значимости, при этом должно выполняться условие

$$P\left(|\,t\,| > t_q\right) = q . \tag{7.3}$$

Рассмотрим случай, когда значения дисперсий $\sigma_1^{\,2}$ и $\sigma_2^{\,2}$ распределений неизвестны. Описанный выше способ не применим. В качестве критерия используют нормированный параметр $t_{p,k}$ распределения Стьюдента

$$t = \frac{\left|\overline{x}_1 - \overline{x}_2\right|}{\sqrt{n_1 s_1^2 + n_2 s_2^2}} \sqrt{\frac{n_1 \cdot n_2 \left(n_1 + n_2 - 2\right)}{n_1 + n_2}} , \tag{7.4}$$

где $n_1 + n_2 - 2 = k$ - степень свободы.

При этом устанавливается выполнение следующих условий:

$$\begin{aligned} &|\,t\,| < t_{p,k} ; \\ &|\,t\,| > t_{p,k} . \end{aligned} \tag{7.5}$$

В первом случае можно утверждать, что нулевая гипотеза $H_0$ не противоречит результатам измерений, а во втором - наблюдаемое расхождение считаем существенным и нулевую гипотезу $H_0$ отвергаем.

**Пример 43.** При исследовании точности измерения угла на пункте триангуляции способом круговых приемов выполнены две серии наблюдений с разных площадок (одна расположена на высоте 15,0 м , а вторая - на высоте 1,2 м): $n_1$ = 112 и $n_2$ = 106. Выборочные средние арифметические значения соответственно составили $\overline{x}_1$ = 24,65″; $\overline{x}_2$ = 25,27″ и эмпирические дисперсии: $s_1^2 = 1{,}21$ ; $s_2^2 = 1{,}15$. Проверить гипотезу о равенстве центров распределения, т.е. $H_0:\ a_1 = a_2$.

**Решение.** Учитывая, что имеем дело с выборками большого объема , можно эмпирические значения дисперсий принять в качестве теоретических. Необходимо проверить гипотезу о равенстве центров распределения при известной точности. При уровне значимости $q$ = 0,05 по функции Лапласа $\Phi(t)$ = $p$ = 1 - 0,05 = 0,95. Согласно табл. 1 приложения параметр $t$ = 1,96. Затем вычисляем среднее квадратическое отклонение разности $\bar{x}_1$ - $\bar{x}_2$

$$\sigma_{\bar{x}_1-\bar{x}_2}=\sqrt{\frac{1{,}21}{112}+\frac{1{,}15}{106}}=0{,}147''.$$

Определим величину разности $|\bar{x}_1-\bar{x}_2|=0{,}62''$, а затем вычислим $t_q\cdot\sigma_{\bar{x}_1-\bar{x}_2}=1{,}96\cdot0{,}147=0{,}29''$. Полученные значения сравним между собой

$$|\bar{x}_1-\bar{x}_2|=0{,}62''>t_q\cdot\sigma_{\bar{x}_1-\bar{x}_2}=0{,}29''.$$

Следовательно, величина нашей разности превышает допустимое значение и попадает в критическую область. Нулевая гипотеза отвергается.

Увеличим область допустимых значений, т.е. примем доверительную вероятность $p$ = 0,99. При этом нормированный параметр $t_q$ = 2,58, тогда

$$t_q\cdot\sigma_{\bar{x}_1-\bar{x}_2}=2{,}58\cdot0{,}147=0{,}38''.$$

И в этом случае нулевая гипотеза $H_0$ также отвергается.

Несмотря на то, что углы измерялись одним и тем же методом и инструментами одинаковой точности, значения углов существенно расходятся между собой. Это объясняется различием действий боковой рефракции на процесс измерения горизонтальных углов, если они расположены на разной высоте.

**Пример 44.** Проверить гипотезу о равенстве двух центров распределения $a_1$ и $a_2$ результатов измерений горизонтального угла по выборкам объемов $n_1$ = 9 и $n_2$ = 16, выполненных в разное время (утром и вечером). Вычисленные значения вероятнейших углов составили: $\bar{x}_1$ = 44,08″ и $\bar{x}_2$ = 44,82″. Дисперсия соответственно: $s_1^2=0{,}22$ ; $s_2^2=0{,}39$.

**Решение.** При небольших объемах выборки воспользуемся нормированным параметром t, вычисляемом по формуле (7.4)

$$t=\frac{|44{,}08-44{,}82|}{\sqrt{9\cdot0{,}22+16\cdot0{,}39}}\sqrt{\frac{9\cdot16(9+16-2)}{9+16}}=2{,}97.$$

Согласно табл. 2 приложения по уровню значимости $q$ = 1 -0,95 = 0,05 и степени свободы $k$ = 9 + 16 - 2 = 23 находим параметр $t_{p,k}$ = 2,07. Так как выполняется условие $t$ = 2,97 > $t_{p,k}$ = 2,07, то нулевая гипотеза отвергается.

### 7.3. Проверка гипотез о дисперсиях

Статистические гипотезы о дисперсиях играют большую роль, поскольку дисперсия характеризует точность как самих работ , так и методов и средств измерений. Для проверки гипотез о равенстве дисперсий в двух генеральных совокупностях по независимым выборкам необходимо определить критерий, который не зависел от каких-либо неизвестных параметров. Этому требованию удовлетворяет отношение двух эмпирических дисперсий, вычисленных по этим выборкам

$$\frac{s_1^2}{s_2^2} = F > 1 \ , \qquad (7.6)$$

причем в качестве числителя берется наибольшая из двух дисперсия.

$F$- распределение (Фишера-Снедекора) зависит только от степеней свободы $k_1 = n_1 - 1$ и $k_2 = n_2 - 1$ , если объем выборок соответственно $n_1$ и $n_2$.

Чтобы проверить гипотезу о равенстве дисперсий нужно построить критическую область для критерия $F$ (рис. 17), т.е.

$$F > F_2;$$

$$0 < F < F_1$$

или

$$P(F > F_2) = 0{,}5q;$$

$$P(F < F_1) = 0{,}5q. \qquad (7.7)$$

Такой выбор критической области позволяет обеспечить наилучшую чувствительность нашего критерия $F$. В том случае, если вычисленное

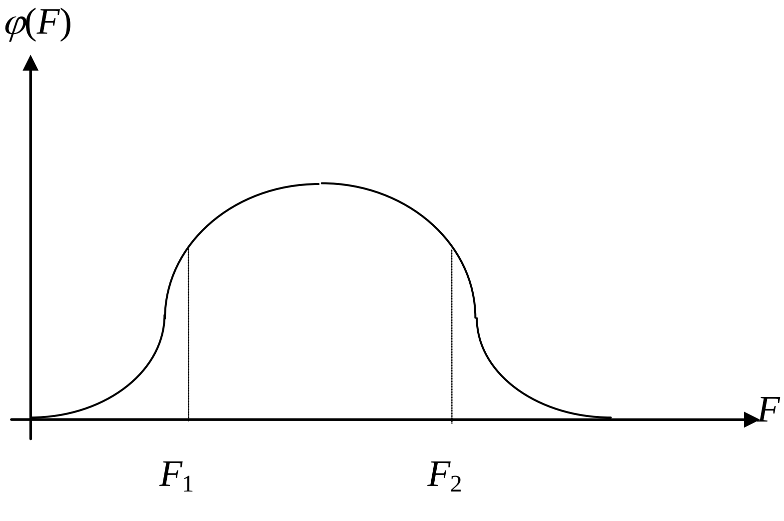


Рис. 17

значение критерия $F$ попадает в критическую область, т.е. оказывается вне допустимых значений $F_1$ и $F_2$, то гипотеза о равенстве дисперсий отвергается.

**Пример 45.** При испытании опытного образца оптического теодолита был измерен горизонтальный угол двумя сериями п$_1$ = 18 и п$_2$ = 15 приемов.

Во второй серии использовалась для уменьшения зрительного параллакса окулярная диафрагма. Выборочные средние квадратические отклонения соответственно составили $s_1 = 1{,}12''$ и $s_2 = 0{,}86''$. С доверительной вероятностью $p = 0{,}90$ проверить гипотезу о равенстве дисперсий $\sigma_1^2$ и $\sigma_2^2$ .

**Решение.** Определяем уровень значимости $q = 1 - 0{,}90 = 0{,}10$. Согласно (7.7) попадание в критическую область возможно при уровне $0{,}5 \cdot q$, т.е. при $0{,}5 \cdot 0{,}10 = 0{,}05$. При этом значении и степеням свободы $k_1 = 18 - 1 = 17$ и $k_2 = 15 - 1 = 14$ по табл. 5 приложения находим $F_q = 2{,}43$, которое сравниваем с вычисленным (7.7)

$$F = \frac{s_1^2}{s_2^2} = \frac{1{,}25}{0{,}74} = 1{,}69 .$$

Так как $F = 1{,}69 < F_q = 2{,}43$, то выдвинутая гипотеза о равенстве дисперсий не противоречит результатам наблюдений. В том и другом случаях результаты получены с одинаковой степенью точности.

### 7.4. Проверка гипотез о законе распределения

В общем случае сравнивают теоретическое и эмпирическое распределения, начиная с графических построений эмпирической кривой распределения (рис. 1-3). Проводя на графиках сглаженную плавную кривую, выдвигают гипотезу о возможном теоретическом распределении. Соответствие проверяют с помощью критериев согласия.

Для приближенной проверки гипотезы нормальности можно использовать эмпирические значения асимметрии и эксцесса, которые вычисляются по формулам (3.48) и (3.52). Средние квадратические отклонения вычисляются согласно (3.51) и (3.54). Большие значения асимметрии и эксцесса по сравнению со средними квадратическими отклонениями служат основанием для браковки гипотезы нормальности распределения. Если соблюдаются условия

$$|A_s| < 3\sigma_A$$

и (7.8)

$$|E_s| < 3\sigma_E ,$$

то асимметрия и эксцесс считаются несущественными.

**Пример 46.** Из выборки объема $n = 64$ угловых невязок звеньев триангуляции I класса получены эмпирические значения асимметрии и эксцесса $A_s = 0{,}44$

и $E_s$ = 0,58. Проверить нулевую гипотезу $H_0$ нормальности распределения при помощи асимметрии и эксцесса.

**Решение.** Вычислить средние квадратические отклонения найденных параметров

$$\sigma_A = \sqrt{\frac{6}{64}} = 0{,}31; \quad \sigma_E = \sqrt{\frac{24}{64}} = 0{,}61.$$

Так как $A_s$ = 0,44 < $3\sigma_A$ = 0,93 и $E_s$ = 0,58 < $3\sigma_E$ = 1,83, то нет основания отвергать гипотезу о нормальности распределения угловой невязки.

В результате независимых измерений получено эмпирическое распределение значений случайной величины $x_1$, $x_2$, ... , $x_n$. Предполагается, что случайная величина $X$ имеет теоретическую непрерывную функцию распределения $F(x)$. По выборочным данным находится эмпирическая функция $F_n(x)$. Критерий согласия Колмогорова основан на сопоставлении эмпирической функции $F_n(x)$ с теоретическими значениями $F(x)$. Для этого определим разности $|F_n(x) - F(x)|$.

Максимальное значение этой разности

$$D_n^0 = \max|F_n(x) - F(x)| \qquad (7.9)$$

представляет собой величину, предельное значение которой было установлено Колмогоровым. Распределение Колмогорова выражает предельную вероятность того, что значение $D_n$, умноженное на $\sqrt{n}$, не превосходит заданного числа $\lambda$

$$\lim P(D \cdot \sqrt{n} \le \lambda) = K(\lambda) \qquad (7.10)$$

какой бы ни была непрерывная функция $F(x)$.

На основании предельного соотношения (7.10) можно написать при больших $n$ следующее приближенное равенство:

$$P\left\{F(x) - \frac{\lambda}{\sqrt{n}} \le F_n(x) \le F(x) + \frac{\lambda}{\sqrt{n}}\right\} \approx K(\lambda)\ . \qquad (7.11)$$

Значения функции

$$P\left\{D_n\sqrt{n} > \lambda\right\} = 1 - K(\lambda) \qquad (7.12)$$

представлены в табл. 6 приложения. Критическая область

$$D > \frac{\lambda_q}{\sqrt{n}}\ . \qquad (7.13)$$

Сравнивая эмпирические значения функции частостей $F_n(x)$ со значениями теоретической функции $F(x)$ и выбрав наибольшую величину разности (по абсолютному значению) $D_n^0$, находим

$$\lambda_0 = D_n^0 \cdot \sqrt{n}. \qquad (7.14)$$

Для этого значения $\lambda_0$ по табл. 6 приложения выбирают значение 1 - $K(\lambda_0)$. Если табличное значение 1 - $K(\lambda)$ при данном уровне значимости $q$ окажется мало, т.е. не более $q$, то это значит, что осуществилось маловероятное событие и расхождение между эмпирическим и теоретическим распределениями можно считать существенным. При этом выполняется условие $\lambda_0 > \lambda_q$ .

В противном случае, когда табличное значение не мало, тогда расхождение между распределениями можно признать случайным, следовательно, гипотезу о нормальности распределения следует считать согласованной. При этом имеет место неравенство $\lambda_0 < \lambda_q$ .

Определим нормальную функцию распределения

$$F(x) = 0{,}5 + \Phi(t)\ , \qquad (7.15)$$

где $\Phi(t)$ находят по табл. 1 приложения согласно вычисленного нормированного параметра

$$t = \frac{\beta_i - \overline{x}}{\sigma}\ ,$$

где $\beta_i$ - правая граница интервала или класса.

**Пример 47.** Произведено 500 измерений. После группировки вычислены $\overline{x}$ = 6,26 и $s_x$ = 0,70. Выдвигается гипотеза нормальности эмпирического распределения. Вычисления выполнены в табл. 11.

**Решение.** Наибольшим значением является разность для четвертого интервала, которая составила

$$D_n^0 = \max\left|F_n(x) - F(x)\right| = 0{,}0257\,.$$

Вычислим параметр $\lambda_0$

$$\lambda_0 = D_n^0 \cdot \sqrt{n} = 0{,}575\ \cdot$$

Тогда значение 1 - K($\lambda$) = 0,90 (табл. 6 приложения), что соответствует 10% уровню значимости, т.е. q = 0,10.

Таблица 11

| Границы интервалов | Частота | Накопленная частость | $t = \dfrac{\beta_i - \bar{x}}{\sigma}$ | $\Phi(t)$ | $F(x)$=0,5+$\Phi(t)$ | $\lvert F_n(x)$-$F(x)\rvert$ |
|---|---|---|---|---|---|---|
| 4,0 - 4,5 | 4 | 0,008 | -2,51 | -0,4939 | 0,0061 | 0,0019 |
| 4,5 - 5,0 | 14 | 0,036 | -1,80 | -0,4641 | 0,0359 | 0,0001 |
| 5,0 - 5,5 | 55 | 0,146 | -1,08 | -0,3599 | 0,1401 | 0,0059 |
| 5,5 - 6,0 | 92 | 0,330 | -0,37 | -0,1443 | 0,3557 | 0,0257 |
| 6,0 - 6,5 | 160 | 0,650 | 0,34 | 0,1331 | 0,6331 | 0,0169 |
| 6,5 - 7,0 | 96 | 0,842 | 1,06 | 0,3554 | 0,8554 | 0,0134 |
| 7,0 - 7,5 | 66 | 0,974 | 1,77 | 0,4616 | 0,9616 | 0,0124 |
| 7,5 - 8,0 | 11 | 0,996 | 2,48 | 0,4934 | 0,9934 | 0,0026 |
| 8,0 - 8,5 | 2 | 1,000 | 3,20 | 0,4993 | 0,9993 | 0,0007 |
| | 500 | | | | | |

Как видим найденное значение 1 - K($\lambda$) не мало, а следовательно, эмпирический ряд согласуется с теоретическим распределением. Это подтверждается следующим неравенством

$$\lambda_0\,(0{,}575) < \lambda_q\,(1{,}224)\,.$$

Так как при использовании выборок большого объема обычно выполняют группирование и нормирование значений случайной величины, то целесообразнее применить для проверки гипотезы нормальности распределения статистику, которая не зависела бы от неизвестных параметров. Такой статистикой является $\chi^2$. Поэтому в соответствии с выдвинутой гипотезой о нормальности распределения вычисляют вероятности $p_1, p_2, \ldots, p_r$ попадания значений в соответствующий интервал (3.42)

$$p_j = \Phi(t_\beta) - \Phi(t_\alpha)\,,$$

где $j = 1, 2, \ldots, r$ - число классов или интервалов.

В том случае, если частота в интервале (классе) равна $n_j$ , то статистика

$$\chi^2 = \sum_{j=1}^{r} \frac{(n_j - n_j^0)^2}{n_j^0}\,, \tag{7.16}$$

где $n_j^0 = n \cdot p_j$.

При нормальном распределении эмпирического ряда эта статистика имеет почти $\chi^2$- распределение с числом степеней свободы

$$k = r - l, \qquad (7.17)$$

где $l$ - число определяемых параметров распределения.

В нашем случае, когда при вычислении частот использовали среднее значение, среднее квадратическое отклонение и объем выборки, параметр $l$ будет равен 3, т.е. $k = r - 3$.

Критическая область

$$\chi^2 > \chi_q^2. \qquad (7.18)$$

Если нулевая гипотеза неверна, то $\chi^2$ оказывается настолько большим, что даже при сравнительно малых $q$ попадает в критическую область.

В.И. Романовский предложил более простое решение, упрощающее применение критерия согласия $\chi^2$. Если

$$\frac{|\chi^2 - k|}{\sqrt{2k}} \geq 3, \qquad (7.19)$$

то расхождение считается существенным и гипотеза о нормальности распределения отвергается. В противном случае, когда

$$\frac{|\chi^2 - k|}{\sqrt{2k}} < 3, \qquad (7.20)$$

расхождение можно считать случайным и гипотезу о нормальности эмпирического ряда распределения можно считать не противоречащей результатам измерения.

При числе измерений $n > 30$ $\chi^2$ - распределение приближается к нормальному закону, а следовательно, можно воспользоваться нормированным параметром интегральной функции

$$t = \sqrt{2\chi^2} - \sqrt{2k-1}. \qquad (7.21)$$

В качестве границы между случайным и существенным расхождением берется уровень значимости $q = 0{,}05$. Если $P(\chi^2) < 0{,}05$, то эмпирическое значение $\chi^2$ считается неслучайным, поскольку событие с такой малой вероятностью считается невозможным. В таком случае расхождение между гипотезой и опытными данными также следует считать не случайным, а существенным. Следовательно, малая величина $P(\chi^2)$ указывает на недостаточное согласие между гипотезой и наблюдениями. Если же вероятность $P(\chi^2) \geq 0{,}05$, расхождения между гипотезой и наблюдениями можно считать случайными, а гипотезу более или менее согласующейся с опытными данными.

**Пример 48.** В условиях предыдущего примера проверить гипотезу о нормальности распределения при помощи критерия согласия Пирсона (табл. 12).

**Решение.** Определяем $P(\chi^2 > \chi_q^2) = q$ при степени свободы $k$ = 5 - 3 = 2 и доверительной вероятности $p$ = 0,95, согласно табл. 3 приложения. Значение $\chi_q^2$ = 12,60 > $\chi^2$ = 12,45.

При $p$ = 0,99 и той же степени свободы $\chi_q^2$ = 16,80 > $\chi^2$ = 12,45.

Воспользуемся упрощенным правилом В.И Романовского.

$$\frac{|\chi^2 - k|}{\sqrt{2k}} = \frac{12{,}45 - 6{,}00}{\sqrt{2 \cdot 6}} = 1{,}86 < 3{,}0 \ .$$

Вывод: Расхождение можно считать случайным и гипотеза о нормальности эмпирического ряда распределения не противоречит опытным данным.

Таблица 12

| Границы интервалов | Частота $n_j$ | Вероятность $p_j$ | $n_j^0 = np_j$ | $n_j - n_j^0$ | $(n_j - n_j^0)^2$ | $\frac{(n_j - n_j^0)^2}{n_j^0}$ |
|---|---|---|---|---|---|---|
| 4,0 - 4,5 | 4 | 0,006 | 3 | +1 | 1 | 0,33 |
| 4,5 - 5,0 | 14 | 0,030 | 15 | -1 | 1 | 0,07 |
| 5,0 - 5,5 | 55 | 0,140 | 52 | +3 | 9 | 0,17 |
| 5,5 - 6,0 | 92 | 0,216 | 108 | -16 | 256 | 2,37 |
| 6,0 - 6,5 | 160 | 0,278 | 139 | +21 | 441 | 3,17 |
| 6,5 - 7,0 | 96 | 0,220 | 110 | -14 | 196 | 1,78 |
| 7,0 - 7,5 | 66 | 0,108 | 54 | +12 | 144 | 2,67 |
| 7,5 - 8,0 | 11 | 0,032 | 16 | -5 | 25 | 1,56 |
| 8,0 - 8,5 | 2 | 0,006 | 3 | -1 | 1 | 0,33 |
| Суммы | 500 | 1,000 | 500 | | | 12,45 |

## 7.5. Проверка гипотез различия выборок

В случае, когда имеются две выборки объемов $n_1$ и $n_2$, возникает вопрос о принадлежности этих выборок к одной генеральной совокупности. При решение данной задачи можно воспользоваться критерием Колмогорова $\lambda$, вычисляемым по формуле

$$\lambda = D_{\max}\sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}}\,. \tag{7.22}$$

Критерий $\lambda$ удобен в том отношении, что он не зависит от вида распределения в генеральной совокупности и может использоваться для малых выборок.

**Пример 49.** Даны распределения значений угла двумя различными теодолитами. Проверить гипотезу о принадлежности распределений к одной генеральной совокупности при q = 0,05. Результаты обработки представлены в табл. 13.

**Решение.** Наибольшее значение разности $D_{max}$ = 0,094 , тогда значение

$$\lambda = 0{,}094\sqrt{\frac{101\cdot 99}{101+99}} = 0{,}66\ .$$

Значение 1- $K(\lambda)$ = 0,776 является достаточно большим и гипотеза об однородности выборок не противоречит результатам измерений.

Таблица 13

| $x_i$ | Частоты | | Частости | | Накоплен. частость | | Разность накоплен. частости |
|---|---|---|---|---|---|---|---|
| | I | II | I | II | I | II | |
| 0,0 - 0,1 | 1 | - | 0,010 | - | 0,010 | - | 0,010 |
| 0,1 - 0,2 | 4 | 4 | 0,040 | 0,040 | 0,050 | 0,040 | 0,010 |
| 0,2 - 0,3 | 8 | 10 | 0,079 | 0,101 | 0,129 | 0,141 | 0,012 |
| 0,3 - 0,4 | 18 | 10 | 0,178 | 0,101 | 0,307 | 0,242 | 0,065 |
| 0,4 - 0,5 | 6 | 12 | 0,059 | 0,121 | 0,366 | 0,363 | 0,003 |
| 0,5 - 0,6 | 7 | 12 | 0,069 | 0,121 | 0,435 | 0,484 | 0,049 |
| 0,6 - 0,7 | 5 | 5 | 0,050 | 0,050 | 0,485 | 0,534 | 0,049 |
| 0,7 - 0,8 | 14 | 13 | 0,139 | 0,131 | 0,624 | 0,665 | 0,041 |
| 0,8 - 0,9 | 2 | 6 | 0,020 | 0,061 | 0,644 | 0,726 | 0,082 |
| 0,9 - 1,0 | 6 | 7 | 0,059 | 0,071 | 0,703 | 0,797 | 0,094 |
| 1,0 - 1,1 | 11 | 7 | 0,109 | 0,071 | 0,812 | 0,868 | 0,056 |
| 1,1 - 1,2 | 9 | 4 | 0,089 | 0,040 | 0,901 | 0,908 | 0,007 |
| 1,2 - 1,3 | 4 | 3 | 0,040 | 0,030 | 0,941 | 0,938 | 0,003 |
| 1,3 - 1,4 | 4 | 1 | 0,040 | 0,010 | 0,981 | 0,948 | 0,033 |
| 1,4 - 1,5 | 2 | 5 | 0,020 | 0,050 | 1,001 | 0,998 | 0,030 |
| Суммы | 101 | 99 | 1,001 | 0,998 | | | |

## 7.6. Критерий Кочрена. Проверка гипотезы об однородности ряда дисперсий

В случае, когда имеют место $k$ выборок одинакового объема, можно воспользоваться для проверки гипотезы об однородности дисперсий методом, предложенным Кочреном. Этот метод основан на рассмотрении последовательности величин $G_i$:

$$G_i = \frac{s_i^2}{s_1^2 + s_2^2 + \ldots + s_k^2}, \qquad (7.23)$$

где $i$ - номер выборки.

Для максимального значения этой последовательности $G_{max}$, который отвечает наибольшей дисперсии, установлен закон распределения. Значения $G_q$ представлены в табл. 7 приложения.

Если $G_{max} > G_q$ при $n - 1$ и $k$, то вероятность получить такое или большее значение оказывается меньше того уровня значимости, для которого составлена таблица, и поэтому гипотеза об однородности ряда дисперсий с тем же уровнем значимости отвергается.

Таблица 14

| № Прием | *A* | | *B* | | *C* | | *D* | | *E* | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *v*,c | *vv* | *v*,c | *vv* | *v*,c | *vv* | *v*, c | *vv* | *v*, c | *vv* |
| 1 | -1,25 | 1,56 | +1,63 | 2,66 | -1,70 | 2,89 | -0,63 | 0,40 | -0,83 | 0,69 |
| 2 | -1,95 | 3,80 | -0,47 | 0,22 | -0,30 | 0,09 | -1,93 | 3,72 | -1,53 | 2,34 |
| 3 | +0,95 | 0,90 | -0,07 | 0 | +2,10 | 4,41 | -0,13 | 0,02 | +0,57 | 0,32 |
| 4 | -3,25 | 10,56 | -1,77 | 3,13 | -2,70 | 7,29 | -1,73 | 2,99 | -2,03 | 4,12 |
| 5 | -0,25 | 0,06 | +0,03 | 0 | +1,60 | 2,56 | -0,53 | 0,28 | -0,43 | 0,18 |
| 6 | +2,05 | 4,20 | +0,23 | 0,05 | +1,40 | 1,96 | +2,97 | 8,82 | +1,37 | 1,88 |
| 7 | +0,15 | 0,02 | -0,87 | 0,76 | -0,90 | 0,81 | +0,47 | 0,22 | +0,17 | 0,03 |
| 8 | +0,35 | 0,12 | +0,23 | 0,05 | +0,50 | 0,25 | +1,57 | 2,46 | +0,77 | 0,59 |
| 9 | -0,05 | 0 | -1,37 | 1,88 | -1,20 | 1,44 | -0,33 | 0,11 | +1,07 | 1,14 |
| 10 | +2,75 | 7,56 | +0,43 | 0,18 | +2,00 | 4,00 | +0,07 | 0 | -0,73 | 0,53 |
| 11 | +0,45 | 0,20 | -0,97 | 0,44 | +0,30 | 0,09 | +0,77 | 0,59 | +0,67 | 0,45 |
| 12 | -1,05 | 1,10 | +1,13 | 1,28 | -1,20 | 1,44 | -1,33 | 1,77 | +1,07 | 1,14 |
| 13 | -1,15 | 1,32 | +1,23 | 1,51 | -0,70 | 0,49 | -0,13 | 0,22 | -0,63 | 0,40 |
| 14 | +2,15 | 4,62 | +0,03 | 0 | +0,40 | 0,16 | +0,47 | 0,22 | -0,13 | 0,02 |
| 15 | +0,15 | 0,02 | +0,63 | 0,40 | +0,40 | 0,16 | +0,47 | 0,22 | +0,67 | 0,45 |
| Суммы | +9,00<br>-8,95 | 36,04 | +5,57<br>-5,52 | 13,06 | +8,70<br>-8,70 | 28,14 | +6,79<br>-6,74 | 21,84 | +6,36<br>-6,31 | 14,28 |
| $s^2$ | | 2,57 | | 0,93 | | 2,01 | | 1,56 | | 1,02 |

**Пример 50.** На геодезическом пункте 2 класса измерены углы способом круговых приемов теодолитом Т2 15-ю приемами. Значения отклонений $v_i$ измеренных углов от средних значений приведены в табл. 14.

**Решение.** Выбираем значение наибольшей дисперсии: $s_1^2 = 2{,}57$, для которого вычисляем $G_1$

$$G_i = \frac{2{,}57}{2{,}57 + 0{,}93 + 2{,}01 + 1{,}56 + 1{,}02} = 0{,}318 .$$

По степеням свободы $n$ -1 = 15 - 1 = 14 и $k$ = 5 при уровне значимости $q$ = 0,05 по табл.7 приложения получаем $G_q$ = 0,364, которое сравниваем с вычисленным. При этом выполняется следующее неравенство $G_1$ = 0,318 < $G_q$ = 0,364.

Вывод. Можно считать, что все эмпирические дисперсии измерения горизонтальных углов по точности можно отнести к однородным.

## 7.7. $\chi^2$ как критерий однородности распределений

Критерий $\chi^2$ используется для проверки гипотезы однородности распределений в рассматриваемых совокупностях. Пусть имеем две независимые выборки объемов $n_1$ и $n_2$, разбитых на две группы с частостями $k_1'$; $k_2'$; . . . ; $k_r'$ и $k_1''$; $k_2''$; . . . ; $k_r''$, при этом выполняется условие

$$\begin{aligned} k_1' + k_2' + \ . \ . \ . \ + k_r' = n_1 \ ; \\ k_1'' + k_2'' + \ . \ . \ . \ + k_r'' = n_2 \ . \end{aligned} \qquad (7.24)$$

Требуется проверить гипотезу о том, что распределение случайной величины в двух генеральных совокупностях, из которых взяты выборки, одинаковы. Для этого используем критерий $\chi^2$ в виде

$$\chi^2 = n_1 \cdot n_2 \sum_{i=1}^{r} \frac{1}{k_i' + k_i''} \left( \frac{k_i'}{n_1} - \frac{k_i''}{n_2} \right)^2 . \qquad (7.25)$$

Этот критерий при больших $n_1$ и $n_2$ распределен по закону $\chi^2$ с $r$ - 1 степенями свободы.

**Пример 51.** В процессе исследования точности угловых измерений в районе Цимлянского водохранилища выполнялись одновременно инструментальные наблюдения теодолитами Т2 со столика и со штатива на трех пунктах $A$, $B$, $C$. Разности значений измеренных углов “сигнал” - ”штатив” представляют собой эмпирическое распределение. Необходимо проверить однородность указанных распределений, полученных на пунктах *A* и *C*. Значения разностей представлены в табл. 15.

**Решение.** Вычислив в табл. 15 числовые значения суммы произведений

$$\frac{1}{k_i' + k_i''}\left(\frac{k_i'}{n_1} - \frac{k_i''}{n_2}\right)^2,$$

определяем $\chi^2 = 57\cdot 99\cdot 417\cdot 10^{-5} = 23{,}53$.

По табл. 6 приложения согласно принятому уровню значимости $q = 0{,}05$ и степени свободы $r - 1 = 5 - 1 = 4$ находим $\chi^2 = 9{,}5$. Так как

$$\chi^2 = 23{,}53 > \chi_q^2 = 9{,}5,$$

то гипотеза об однородности разностей результатов измерения горизонтальных углов со столика "сигнала" и со "штатива" на пунктах *A* и *C* отвергается. Здесь факторами, влияющими на измерения, могут быть рефракция (визирование над поверхностью воды и над степью) и различие температур.

Таблица 15

| Границы интервалов | Частоты | | $k_i' + k_i''$ | $\frac{1}{k_i' + k_i''}$ | $\frac{k_i'}{n_1}$ | $\frac{k_i''}{n_2}$ | $\left[\frac{k_i'}{n_1} - \frac{k_i''}{n_2}\right]^2$ | 5x8 |
|---|---|---|---|---|---|---|---|---|
| | *A* | *C* | | | | | | |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| -3 - -2 | 9 | - | 9 | $111\cdot 10^{-3}$ | 0,158 | - | $250\cdot 10^{-4}$ | $277{,}5\cdot 10^{-5}$ |
| -2 - -1 | 13 | 13 | 26 | $38{,}5\cdot 10^{-3}$ | 0,228 | 0,130 | $96\cdot 10^{-4}$ | $37{,}0\cdot 10^{-5}$ |
| -1 - 0 | 20 | 35 | 55 | $18{,}2\cdot 10^{-3}$ | 0,351 | 0,354 | 0 | 0 |
| 0 - +1 | 12 | 37 | 49 | $20{,}4\cdot 10^{-3}$ | 0,210 | 0,374 | $269\cdot 10^{-4}$ | $54{,}9\cdot 10^{-5}$ |
| +1 - +2 | 3 | 14 | 17 | $58{,}8\cdot 10^{-3}$ | 0,053 | 0,142 | $79\cdot 10^{-4}$ | $47{,}6\cdot 10^{-5}$ |
| Суммы | 57 | 99 | 156 | | 1,000 | 1,000 | | $417{,}6\cdot 10^{-5}$ |

## 7.8. Критерий Аббе

Пусть $x_1, x_2, \ldots, x_n$ - результаты равноточных независимых измерений какой-либо величины с одной и той же дисперсией $\sigma^2$, распределенных нормально. Критерий Аббе предназначен для проверки гипотезы об отсутствии в результатах измерений переменных систематических ошибок, т.е. для проверки равенства математических ожиданий $a_1 = a_2 = \ldots = a_n = a$ против альтернативы $|a_{i+1} - a_i| > 0$ для всех значений $i = 1, 2, \ldots, n - 1$.

В современной форме статистики критерий Аббе представляет собой отношение

$$\tau = \frac{d^2}{s^2}, \tag{7.26}$$

где $d^2$ и $s^2$ - оценки дисперсии $\sigma^2$, вычисленные по формулам

$$\tau = \frac{d^2}{s^2} = \frac{\frac{1}{2(n-1)}\sum_{i=1}^{n-1}(x_{i+1}-x_i)^2}{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\overline{x})^2} = \frac{1}{2}\frac{\sum_{i=1}^{n-1}(x_{i+1}-x_i)^2}{\sum_{i=1}^{n}(x_i-\overline{x})^2}. \qquad (7.27)$$

Если нулевая гипотеза верна, то $M(\tau) = 1$; $D(\tau) = \frac{n-2}{n^2-1}$. Если центр группирования величины $X$ смещался, то знаменатель $s^2$, будет намного больше числителя $d^2$, поэтому значения статистики $\tau$ будут меньше тех значений, которые наблюдаются, когда справедлива нулевая гипотеза.

Распределение значений $\tau$ табулированы (табл. 8 приложения). При фиксированных параметрах $n$ и $p$ функция $\tau_{p,n}$ представляет собой

$$P(\tau < \tau_{p,n}) = p. \qquad (7.28)$$

Если оказалось, что $\tau < \tau_{p,n}$, то основная гипотеза о равенстве математических ожиданий, а следовательно, об отсутствии переменных систематических ошибок отвергается.

**Пример 52.** Географическая широта Ташкентской обсерватории определялась из наблюдений звезд, причем каждая звезда определялась дважды при двух положениях инструмента. Проверить гипотезу о равенстве математических ожиданий $a_1 = a_2 = ... = a_{28}$ результатов измерений [14]. Результаты обработки представлены в табл. 16.

**Решение.** В табл. 12 вычислены суммы квадратов последовательных разностей $\sum(x_{i+1}-x_i)^2$ и отклонений от арифметической средины $\sum(x_i-\overline{x})^2$ на основании формулы (7.27) находим

$$\tau = \frac{87{,}63}{2\cdot 96{,}64} = 0{,}453.$$

Для проверки гипотезы об отсутствии систематических ошибок в результатах наблюдений воспользуемся табл. 8 приложения. В этой таблице по степени свободы $n = 28$ и $q = 0{,}001$ определяем

$$\tau_{p,n} = 0{,}467.$$

При этом выполняется следующее неравенство

$$\tau = 0{,}453 < \tau_{p,n} = 0{,}467.$$

Следовательно, гипотеза об отсутствии смещения центра группирования в процессе наблюдений отвергается. Поэтому результаты наблюдений $x_i = \varphi$ ($i$ = 1, 2, . . . , 28) широты Ташкентской обсерватории содержат значительные переменные систематические ошибки.

Таблица 16

| № Прием | Широта $\varphi$ | $d_i$ | $d_i d_i$ | $x_i - \bar{x}$ | $(x_i - \bar{x})^2$ |
|---|---|---|---|---|---|
| 1 | 41° 19′ 33,10″ | -0,26″ | 0,06 | +1,64 | 2,69 |
| 2 | 32,84 | +0,12 | 0,01 | +1,38 | 1,90 |
| 3 | 32,96 | +1,00 | 1,00 | +1,50 | 2,25 |
| 4 | 33,96 | -5,01 | 25,10 | +2,50 | 6,25 |
| 5 | 28,95 | +2,22 | 4,93 | -2,51 | 6,30 |
| 6 | 31,17 | -1,09 | 1,19 | -0,29 | 0,08 |
| 7 | 30,08 | +0,42 | 0,18 | -1,38 | 1,90 |
| 8 | 30,50 | -1,69 | 2,86 | -0,96 | 0,92 |
| 9 | 28,81 | -0,11 | 0,01 | -2,65 | 7,02 |
| 10 | 28,70 | +2,46 | 6,05 | -2,76 | 7,62 |
| 11 | 31,16 | -2,84 | 8,07 | -0,30 | 0,09 |
| 12 | 28,32 | +3,06 | 9,36 | -3,14 | 9,86 |
| 13 | 31,38 | -0,60 | 0,36 | -0,08 | 0,01 |
| 14 | 30,78 | +1,37 | 1,88 | -0,68 | 0,46 |
| 15 | 32,15 | +0,25 | 0,06 | +0,69 | 0,48 |
| 16 | 32,40 | +2,49 | 6,20 | +0,94 | 0,88 |
| 17 | 34,89 | -1,44 | 2,07 | +3,43 | 11,76 |
| 18 | 33,45 | -0,11 | 0,01 | +1,99 | 3,96 |
| 19 | 33,34 | +0,21 | 0,04 | +1,88 | 3,53 |
| 20 | 33,55 | -1,16 | 1,35 | +2,09 | 4,37 |
| 21 | 32,39 | +0,76 | 0,52 | +0,93 | 0,86 |
| 22 | 33,11 | -1,60 | 2,56 | +1,65 | 2,72 |
| 23 | 31,51 | +0,59 | 0,35 | +0,05 | 0 |
| 24 | 32,10 | -0,75 | 0,56 | +0,64 | 0,41 |
| 25 | 31,35 | -0,12 | 0,01 | -0,11 | 0,01 |
| 26 | 31,23 | -3,43 | 11,76 | -0,23 | 0,05 |
| 27 | 27,80 | +1,04 | 1,08 | -3,66 | 13,40 |
| 28 | 28,84 | | | -2,62 | 6,86 |
| | 41° 19′ 31,46″ | | 87,63 | -0,06 | 96,64 |

## 8. Дисперсионный анализ
### 8.1. Задачи дисперсионного анализа

Дисперсионный анализ - это статистический метод анализа результатов наблюдений, которые зависят от различных, одновременно действующих факторов, выбор наиболее важных факторов и оценка их влияния.

В каждом эксперименте средние значения наблюдаемых величин меняются в связи с изменением основных факторов (качественных и количественных), определяющих условия опыта. Исследование влияния тех или иных факторов на изменчивость средних значений наблюдаемых величин является задачей дисперсионного анализа.

С этой целью производится разложение общей дисперсии случайной величины на независимые случайные слагаемые, каждое из которых характеризует влияние того или иного фактора или их взаимодействия. Каждое из этих независимых случайных слагаемых дает оценку дисперсии в общей совокупности. Проверка значимости этих оценок дисперсии производится при помощи таблиц значений статистики $F$.

Если наблюденное значение статистики $F$ окажется меньше табличного $F_q$, то это указывает на то, что не имеется оснований утверждать о влиянии какого-либо фактора на изменчивость средних значений случайной величины; если же наблюденное значение $F$ окажется больше табличного $F_q$, то рассматриваемый фактор влияет на изменчивость средних значений.

После того как при помощи дисперсионного анализа будет произведена оценка влияния факторов на изменчивость средних значений случайной величины в целом и эта оценка дисперсии будет признана значимой, необходимо перейти к подробному исследованию отдельных факторов.

Если исследуется влияние одного фактора на исследуемую величину, то речь идет об однофакторном дисперсионном анализе. Если изучается влияние двух факторов, то речь идет о двухфакторном анализе.

## 8.2. Однофакторный дисперсионный анализ

Рассмотрим действие единичного фактора $A$ (количественного или качественного), который принимает $k$ различных значений (уровней фактора). На $i$-м уровне производится $n_i$ наблюдений:

$$\begin{matrix} y_{11}, \ y_{12}, \ \ldots\ldots\ldots\ldots\ y_{k1}; \\ y_{12}, \ y_{22}, \ \ldots\ldots\ldots\ldots\ y_{k2}; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_{1n}, \ y_{2n}, \ \ldots\ldots\ldots\ldots\ y_{kn}. \end{matrix} \qquad (8.1)$$

Общее число опытов

$$N = n_1 + n_2 + . \, . \, . + n_k .$$

Необходимо проверить нулевую гипотезу о равенстве средних значений на различных уровнях фактора $A$.

Наиболее простые расчеты получаются при $n_1 = n_2 = . \, . \, . n_k = n$. При этом $N = k \cdot n$.

Обозначим через $\bar{y}_i$ среднее значение наблюдений на $i$-м уровне (табл. 17)

$$\bar{y}_i = \frac{1}{n}\sum_{j=1}^{n} y_{ij} = \frac{A_i}{n}, \tag{8.2}$$

а общее среднее значение для всей выборки из $N$ наблюдений

$$\bar{y} = \frac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{n} y_{ij} = \frac{1}{k}\sum_{i=1}^{k} \bar{y}_i . \tag{8.3}$$

Таблица 17

| Номера наблюдений | Уровни фактора $A$ | | | |
|---|---|---|---|---|
| | $a_1$ | $a_2$ | . . . . | $a_n$ |
| 1<br>2<br>.<br>.<br>.<br>.<br>$n$ | $y_{11}$<br>$y_{12}$<br>.<br>.<br>.<br>.<br>$y_{1n}$ | $y_{21}$<br>$y_{22}$<br>.<br>.<br>.<br>.<br>$y_{2n}$ | . . .<br>. . .<br>.<br>.<br>.<br>.<br>. . . | $y_{k1}$<br>$y_{k2}$<br>.<br>.<br>.<br>.<br>$y_{kn}$ |
| | $A_1 = \sum_{j=1}^{n} y_{1j}$ | $A_2 = \sum_{j=1}^{n} y_{2j}$ | . . . . | $A_k = \sum_{j=1}^{n} y_{kj}$ |

Для проведения дисперсионного анализа необходимо произвести оценку общей дисперсии $S(x) = s^2$, которую находят следующим образом:

$$s^2 = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n}(y_{ij} - \bar{y})^2}{N-1} = \frac{1}{N-1}\left\{\sum_{i=1}^{k}\sum_{j=1}^{n} y_{ij}^2 - \frac{\left(\sum_{i=1}^{k}\sum_{j=1}^{n} y_{ij}\right)^2}{N}\right\}, \tag{8.4}$$

и разложить на составляющие, характеризующие влияние фактора $A$ и фактора случайности. Фактор случайности легко оценить благодаря наличию повторных опытов на каждом уровне. Определим выборочную дисперсию на каждом уровне

$$s_i^2 = \frac{1}{n-1}\sum_{j=1}^{n}(y_{ij} - \overline{y}_i)^2 = \frac{1}{n-1}\left\{\sum_{j=1}^{n} y_{ij}^2 - \frac{\left(\sum_{j=1}^{n} yy\right)^2}{n}\right\}. \qquad (8.5)$$

Если нет уверенности в равноточности экспериментов, однородность дисперсий $s_1^2, s_2^2, \ldots, s_k^2$ можно проверить по критерию Кочрена (табл.6 приложения).

Если между дисперсиями нет значимых различий, для оценки генеральной дисперсии $\sigma^2$, характеризующей фактор случайности, используют выборочную дисперсию $s_0^2$

$$s_0^2 = \frac{1}{k}\sum_{i=1}^{k} s_i^2 = \frac{1}{k(n-1)}\left\{\sum_{i=1}^{k}\sum_{j=1}^{n} y_{ij}^2 - \frac{1}{n}\sum_{i=1}^{k}\left(\sum_{j=1}^{n} y_{ij}\right)^2\right\}. \qquad (8.6)$$

Число степеней свободы дисперсии $s_0^2$ равно: $k(n-1) = N - k$ . Приближенную оценку для дисперсии фактора $A$ можно получить следующим образом:

$$\sigma_A^2 \approx s^2 - s_0^2. \qquad (8.7)$$

Более точную оценку для $\sigma_A^2$ можно получить, рассматривая отклонения средних $\overline{y}_i$ на отдельных уровнях от общего среднего всей выборки $\overline{y}$. Действительно,

$$\frac{1}{k-1}\sum_{i=1}^{k}\left(\overline{y}_i - \overline{y}\right)^2 \approx \sigma_A^2 + \frac{\sigma^2}{n} \approx \sigma_A^2 + \frac{s_0^2}{n}, \qquad (8.8)$$

отсюда

$$\sigma_A^2 \approx \frac{1}{k-1}\sum_{i=1}^{k}\left(\overline{y}_i - \overline{y}\right)^2 - \frac{s_0^2}{n}. \qquad (8.9)$$

Дисперсия фактора $A$ для модели с фиксированными уровнями $\sigma_A^2$ не связана ни с какой случайной величиной, это условное название для математического ожидания среднего квадрата отклонений, обусловленного влиянием фактора $A$.

Введем обозначение

$$s_A^2 = \frac{n}{k-1}\sum_{j=1}^{n}\left(y_i - \bar{y}\right)^2 \approx n\sigma_A^2 + s_0^2. \qquad (8.10)$$

В теории дисперсионного анализа $s_0^2$ называют оценкой дисперсий по факторам, а $s_\Delta^2$ - оценкой остаточной дисперсии. Можно составить соотношение

$$F = \frac{s_A^2}{s_0^2}. \qquad (8.11)$$

Величина $F$ имеет $F$ - распределение с числом степеней свободы $k-1$ и $k(n-1) = N-k$.

Из выражения (8.11) вытекает: чем меньше величина $F$, то есть основания принять нулевую гипотезу о равенстве средних значений, и наоборот, чем больше величина $F$, тем больше оснований отвергнуть гипотезу о равенстве средних значений.

Распределение величины $F$ известно, поэтому можно выбрать некоторую критическую величину $F_q$ с критической вероятностью или уровнем значимости $q$ из таблиц $F$ - распределения по $q$ и числу степеней свободы $k-1$ и $k(n-1)$.

Гипотезу о равенстве средних значений отвергают, если $F > F_q$, т.е. влияние фактора $A$ считают значимым, и принимают, если $F < F_q$.

Таким образом, при однофакторном дисперсионном анализе выявляется влияние одного фактора на среднее значение величины. В геодезической практике это может быть связано, например, с выявлением систематической составляющей в измеряемых величинах в зависимости от времени суток (влияние рефракции). Также с помощью дисперсионного анализа можно сравнивать и точностные характеристики (с точки зрения наличия систематических составляющих) различных измерительных приборов, имеющих одинаковую точность, например, измерять различными светодальномерами один и тот же базис.

**Пример 53.** В табл. 18 представлены результаты пятикратного ($n$ = 5) измерения расстояния $D$. Измерения выполнялись тремя светодальномерами ($m$ = 3) в различных условиях.

Таблица 18

| N Прибор | Отклонение $D$, данное в сотых долях см | | | | | | | | | | $\sum_j y_{ij}$ | $(\sum_j y_{ij})^2$ | $\sum_j y_{ij}^2$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | 2 | | 3 | | 4 | | 5 | | | | |
| | $y_{i1}$ | $y_{i1}^2$ | $y_{i2}$ | $y_{i2}^2$ | $y_{i3}$ | $y_{i3}^2$ | $y_{i4}$ | $y_{i4}^2$ | $y_{i5}$ | $y_{i5}^2$ | | | |
| 1 | -4 | 16 | -2 | 4 | -21 | 441 | -4 | 16 | -4 | 16 | -35 | 1225 | 493 |
| 2 | +7 | 49 | +11 | 121 | +30 | 900 | +28 | 784 | +27 | 729 | +103 | 10609 | 2583 |
| 3 | +19 | 361 | +2 | 4 | -13 | 169 | -9 | 81 | +2 | 4 | 1 | 1 | 619 |
| $\sum_{i=1}$ | +22 | 426 | +11 | 129 | -4 | 1510 | +15 | 881 | +25 | 749 | +69 | 11835 | 3695 |

Требуется выяснить, существенно ли влияние условий измерений на величину расстояния $D$.

**Решение.** Определим полную сумму квадратов отклонений $Q$ с $k$ = ($mn$ - 1) степенями свободы:

$$Q = \sum_{i=1}^{n}\sum_{j=1}^{m}(y_{ij} - \bar{y})^2; \qquad (8.12)$$

сумму квадратов отклонений между приборами $Q_1$ с $k_1$ = ($m$ - 1) степенями свободы:

$$Q_1 = n \cdot \sum_{i=1}^{m}(\bar{y}_i - \bar{y})^2; \qquad (8.13)$$

сумму квадратов отклонений внутри приборов $Q_2$ с $k_2$ = ($mn$ - $m$) степенями свободы:

$$Q_2 = \sum_{i=1}^{m}\sum_{j=1}^{n}(\bar{y}_{ij} - \bar{y}_i)^2. \qquad (8.14)$$

Подставляя значения, получим

$$Q = \sum_{1}^{N} y_{ij}^2 - \frac{1}{N}\left(\sum y_{ij}\right)^2 = 3695 - \frac{1}{3\cdot 5}69^2 = 3695 - 317 = 3378; \qquad k = 15 - 1 = 14;$$

где $N = m \cdot n$.

$$Q_1 = n \cdot \sum_{i=1}^{m}(\bar{y}_i - \bar{y})^2 = \frac{\sum_{i=1}^{m}\left(\sum_{j=1}^{n} y_{ij}\right)^2}{n} - \frac{\left(\sum_{ij} y_{ij}\right)^2}{N} = \frac{11835}{5} - \frac{69^2}{5\cdot 3} = 2367 - 317 = 2050; \quad k = 3 - 1 = 2;$$

$$Q_2 = \sum_{i=!}^{m}\sum_{j=1}^{n}(y_{ij} - \overline{y}_i)^2 = \sum_{ij} y_{ij}^2 - \frac{\sum_{i=1}^{m}\left(\sum_{j=1}^{n} y_{ij}\right)^2}{n} = 3695 - \frac{11835}{5} = 3695 - 2367 = 1328; \quad k = 15 - 3 = 12.$$

Для нашего примера таблица однофакторного дисперсионного анализа будет иметь следующий вид (табл. 19)

Таблица 19

| Компонента дисперсии | Сумма квадратов | Число степеней свободы | Средний квадрат |
|---|---|---|---|
| Между приборами | 2050 | 2 | 1025,0 |
| Внутри приборов | 1328 | 12 | 110,7 |
| Полная | 3378 | 14 | 241,3 |

Производя проверку нулевой гипотезы с помощью $F$ - распределения, найдем отношение среднего квадрата между приборами к среднему квадрату внутри приборов, т.е.

$$F = \frac{1025{,}0}{110{,}7} = 9{,}26.$$

По табл. 5 приложения находим критические границы для $F_q$ при $k = 2$ большей дисперсии и $k = 12$ меньшей дисперсии, равные при уровне значимости $q_1 = 0{,}05$ - 3,88 и $q_2 = 0{,}01$ - 6,93. Полученное из наблюдений значение $F$ превышает указанные границы, т.е.

$$F = 9{,}26 > F_q \begin{cases} 3{,}88 \\ 6{,}93 \end{cases}.$$

Вывод: нулевая гипотеза отвергается, т.е. приборы имеют различные систематические ошибки.

### 8.3. Двухфакторный дисперсионный анализ

Когда исследуется действие двух, трех и т.д. факторов, то структура дисперсионного анализа остается такой же, что и при однофакторном анализе, но усложняются вычисления. Рассмотрим задачу оценки действия двух одновременно действующих факторов. Допустим, что в примере со светодальномерами измерения производились различными наблюдателями и требуется оценить, обуславливается ли рассеивание полученных значений расстояния $D$ и средних его значений в группах различием между приборами или различием между наблюдателями, производившими измерения.

В данном случае идея дисперсионного анализа состоит в разложении суммы квадратов отклонений общего среднего на компоненты, характеризующие факторы изменчивости.

Допустим, что имеются два фактора $A$ и $B$, по которым можно данные наблюдения классифицировать. Пусть по признаку $A$ все наблюдения разбиваются на $r$ групп ($A_1, A_2, \ldots, A_r$), а по признаку $B$ - на $v$ групп ($B_1, B_2, \ldots, B_v$) таким образом, что все результаты наблюдений разбиваются на $rv$ групп, т.е. общее число наблюдений $N = rv$. В данном случае $y_{ij}$ - наблюдение, попавшее в группу $A_i$ по признаку $A$ и в группу $B_i$ по признаку $B$. Тогда можно записать

$$\begin{array}{l} y_{11},\ y_{12},\ \ldots\ldots,\ y_{1v}; \\ y_{21},\ y_{22},\ \ldots\ldots,\ y_{2v}; \\ \ldots\ldots\ldots\ldots \\ y_{r1},\ y_{r2},\ \ldots\ldots,\ y_{rv}. \end{array}$$

Среднее значение по каждому столбцу и строке, а также общее среднее будут вычисляться следующим образом:

$$\bar{y}_i = \frac{1}{v}\sum_{j=1}^{v} y_{ij}, \qquad \bar{y}_j = \frac{1}{r}\sum_{i=1}^{r} y_{ij},$$

$$\bar{y} = \frac{1}{N}\sum_{i=1}^{r}\sum_{j=1}^{v} y_{ij}\ .$$

Полная сумма квадратов отклонений отдельных наблюдений от общей средней с $k = (rv - 1)$ степенями свободы

$$\begin{aligned} Q &= \sum_{i=1}^{r}\sum_{j=1}^{v}(y_{ij} - \bar{y})^2 = \sum_{i=1}^{r}\sum_{j=1}^{v}(y_{ij} - \bar{y}_i - \bar{y}_j + \bar{y} + \bar{y}_i - \bar{y} + \bar{y}_j - \bar{y})^2 = \\ &= v\sum_{i=1}^{r}(\bar{y}_i - \bar{y})^2 + r\sum_{j=1}^{v}(\bar{y}_j - \bar{y})^2 + \sum_{i=1}^{r}\sum_{j=1}^{v}(y_{ij} - \bar{y}_i - \bar{y}_j + \bar{y})^2 = Q_1 + Q_2 + Q_3, \end{aligned} \tag{8.15}$$

где слагаемое $Q_1$ - сумма квадратов разностей между средними по строкам и общим средним (характеризует влияние фактора $A$) с $k_1 = 1$ степенями свободы; слагаемое $Q_2$ - сумма квадратов разностей между средними по столбцам и общим средним (характеризует влияние фактора $B$) с $k_2 = (v - 1)$ степенями свободы;

слагаемое $Q_3$ - остаточная сумма квадратов (характеризует влияние неучтенных факторов) с $k_3 = (r - 1)(v - 1)$ степенями свободы.

Произведем оценку дисперсий:

$$\tag{8.16}$$

$$s_1^2 = \frac{Q_1}{r-1} = \frac{v\sum_{i=1}^{r}(\overline{y}_i - \overline{y})^2}{r-1}; \tag{8.17}$$

$$\tag{8.18}$$

$$s_3^2 = \frac{Q_3}{(r-1)(v-1)} = \frac{\sum_i \sum_j (y_{ij} - \overline{y}_i - \overline{y}_j + \overline{y})^2}{(r-1)(v-1)}. \tag{8.19}$$

В двухфакторном анализе для выяснения значимости влияния факторов $A$ и $B$ на исследуемый признак сравнивают дисперсии по факторам с остаточной дисперсией. Известно, если случайная величина распределена нормально, то отношение дисперсий имеет $F$ - распределение. Таким образом,

$$F_A = \frac{s_1^2}{s_3^2}, \qquad F_B = \frac{s_2^2}{s_3^2}. \tag{8.20}$$

Полученные значения $F_A$ и $F_B$ сравнивают с табличными при выбранном уровне значимости $q$. Если $F_A < F_q$ и $F_B < F_q$, то нулевую гипотезу о равенстве средних можно принять, т.е. влияние факторов $A$ и $B$ на исследуемый признак не значительно.

## 9. Элементы теории корреляции

### 9.1. Понятие о стохастической связи

При математической обработке результатов измерений, исследовании новых инструментов и методов работ и т.д. часто приходится устанавливать зависимость полученных результатов от каких-либо факторов. Иногда один параметр влияет на другой.

Если зависимость между результатами измерений или определенными параметрами установлена и выражена формулой, то ее можно использовать при предвычислении ожидаемой точности исследуемого прибора или для организации наблюдений и обработки результатов измерений. При этом встречаются две формы связи: функциональная и стохастическая.

Существо функциональной зависимости заключается в том, что какая-либо величина определяется как однозначная функция одной или нескольких величин. Эта зависимость между двумя переменными величинами характеризуется тем, что каждому значению одной величины соответствует вполне определенное значение другой.

Функциональная связь может быть и между случайными величинами, некоторые были рассмотрены нами ранее. Но между случайными величинами может быть связь и другого рода, проявляющаяся в том, что одна случайная переменная реагирует на изменение другой измерением своего закона распределения. Такую связь назовем стохастической (вероятностной).

Стохастическая связь между двумя случайными величинами появляется обычно тогда, когда имеются общие случайные факторы, влияющие как на одну, так и на другую величины наряду с другими неодинаковыми для обеих величин случайными факторами. Так, например, если $X$ представляет некоторую функцию от случайных величин $Z_1, Z_2, \ldots, Z_m; V_1, V_2, \ldots, V_k$:

$$X = f\,(Z_1, Z_2, \ldots, Z_m; V_1, V_2, \ldots, V_k),$$

а $Y$ представляет функцию от тех же случайных величин $Z_1, Z_2, \ldots, Z_m$ и некоторой совокупности других $U_1, U_2, \ldots U_l$, т.е.

$$Y = \varphi\,(Z_1, Z_2, \ldots Z_m; U_1, U_2, \ldots, U_l),$$

то величины $X$ и $Y$ будут между собой стохастически связаны.

В практике статистических исследований часто рассматривается частный случай стохастической связи, называемый статистической связью. Об этой связи говорят тогда, когда условное математическое ожидание одной случайной переменной является функцией значения, принимаемого другой случайной переменной, т.е.

$$M\,(Y/X) = f\,(X). \qquad (9.1)$$

Таким образом, изучение статистических зависимостей основывается на исследовании таких связей между случайными переменными, при которых значение одной случайной переменной изменяется в зависимости от того, какие значения принимает другая случайная переменная.

Чтобы изучить статистическую зависимость, как видно из формулы (9.1), необходимо знать условное математическое ожидание случайной переменной. Для его оценки нужно установить вид распределения ($X$, $Y$). Однако определенный вид распределения по отдельной ограниченной по объему выборке, может привести к серьезным ошибкам. Поэтому переходят от условного математического ожидания случайной переменной к условному среднему значению, т.е. принимают

$$M\,(Y/ = x) = \overline{y}\,(x), \qquad (9.2)$$

где $M$ ($Y/X = x$) - математическое ожидание случайной переменной $Y$ при условии, что случайная величина $X$ приняла значение $x$.

Зависимость между одной случайной переменной и условным средним значением другой случайной переменной называется корреляционной зависимостью. Она характеризуется формой и теснотой связи.

Основное применение, которое находит теория корреляции, относится к решению такой задачи как указание пределов, в которых с заранее заданной надежностью (вероятностью) будет находиться интересующая нас величина, если другие связанные с ней величины получают определенные значения. Так, например, нас может интересовать корреляционная связь между невязками в приращениях координат по сторонам триангуляции (полигонометрии, трилатерации) в зависимости от направления сторон относительно осей координат; в какой зависимости находятся результаты измерений зенитных расстояний от времени наблюдений и т.д.

## 9.2. Понятие о коэффициенте корреляции

Простейшей характеристикой связи между случайными величинами $X$ и $Y$ служит математическое ожидание произведения отклонений $X$ и $Y$ от их центров группирования, т.е.

$$cov(X,Y) = \mu_{xy} = M[(x - a)(y - b)]\,, \qquad (9.3)$$

где $a = M(X)$; $b = M(Y)$.

Числовая $\mu_{xy}$ получила название ковариации (совместного варьирования случайных величин). Эту характеристику называют также корреляционным моментом.

Если $X$ и $Y$ независимы, то

$$cov(X,Y) = \mu_{xy} = 0. \qquad (9.4)$$

Величина $cov(X,Y)$ зависит от единицы измерения, в котрой выражены $X$ и $Y$, поэтому она сама по себе не может служить показателем связи (тем более, что величины могут быть неоднородными). Чтобы иметь дело с безразмерным показателем, рассматривают ковариации нормированных отклонений.

$$t_x = \frac{x-a}{\sigma_x}; \quad t_y = \frac{y-b}{\sigma_y}, \qquad (9.5)$$

где $\sigma_x, \sigma_y$ - дисперсии величин $x, y$.

Тогда

$$\rho_{xy} = \mathrm{cov}(t_x, t_y) = M\left(\frac{x-a}{\sigma_x} \cdot \frac{y-b}{\sigma_y}\right) = \frac{M[(x-a)(y-b)]}{\sigma_x \sigma_y} = \frac{\mu_{xy}}{\sigma_x \sigma_y}. \qquad (9.6)$$

Показатель $\rho_{xy}$ называется коэффициентом корреляции. Коэффициент корреляции характеризует тесноту связи.

Случайные величины $x$ и $y$, для которых коэффициент корреляции $\rho_{xy} = 0$, называются некоррелированными. Если $\rho_{xy} \neq 0$, то случайные величины $x$ и $y$ находятся в коррелированной связи.

Коэффициент корреляции обладает следующими свойствами:

1. Коэффициент корреляции изменяется в пределах от -1 до +1, т.е.

$$-1 \le \rho_{xy} \le +1; \quad (9.7)$$

2. Если коэффициент корреляции равен +1 и -1, между $x$ и $y$ существует строгая прямолинейная зависимость, т.е.

$$y=Rx+l\,. \quad (9.8)$$

В случае, когда $\rho_{xy}$=+1, с увеличением или уменьшением $x$ увеличивается или уменьшается $y$. Если $\rho_{xy}$=-1, то с увеличением $x$ уменьшается $y$.

3. Есл $\rho_{xy}$= 0, то между $x$ и $y$ прямолинейной связи не существует (нелинейная связь существовать может).

Таким образом, значение $\rho_{xy}$, близкое к единице, указывает на близость связи между величинами $x$ и $y$ к функциональной. В отдельном случае, когда $|\rho_{xy}| = 1$, все двумерное распределение сосредоточено на прямой

$$\frac{x-a}{\sigma_x} = \frac{y-b}{\sigma_y}\,. \quad (9.9)$$

Для экспериментального изучения зависимости между двумя случайными величинами $X$ и $Y$ выполняют некоторое количество $n$ независимых наблюдений. Результат $i$-го наблюдения дает пару значений $x_i$, $y_i$ (где $i$ = 1, 2, ... , $n$). По этим значениям определяются оценки как средних значений, так и коэффициента корреляции.

Оценками теоретических значений $a$ и $b$ служат эмпирические средние значения

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i\,, \qquad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i. \quad (9.10)$$

Соответственно оценками дисперсий $\sigma_x^2$, $\sigma_y^2$ служат эмпирические дисперсии

$$s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2;$$
$$s_y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i-\bar{y})^2. \quad (9.11)$$

Оценкой корреляционного момента служит эмпирический корреляционный момент

$$\mu_{xy} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \quad (9.12)$$

и соответственно по этим оценкам определяют эмпирический коэффициент корреляции

$$r = \frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)}{\left(n-1\right)s_x s_y}\ , \quad (9.13)$$

где $s_x$, $s_y$ - соответственно эмпирические средние квадратические отклонения случайных величин $X$ и $Y$.

Эмпирический коэффициент корреляции, так же как и теоретический $\rho_{xy}$, не превосходит по абсолютной величине единицы, т.е.

$$|r| \leq 1,$$

и обладает теми же свойствами.

## 9.3. Доверительная оценка коэффициента корреляции

Как было сказано выше оценкой коэффициента корреляции $\rho$ является эмпирический коэффициент корреляции $r$. Однако коэффициент корреляции $r$ является надежной оценкой $\rho$ лишь в том случае, когда распределение величин $X$ и $Y$ достаточно близко к нормальному. В этом случае эмпирический коэффициент корреляции имеет приближенное нормальное распределение с параметрами $\rho$ и $\sigma_r$. Зная закон распределения, можно найти доверительные интервалы для параметра $\rho$

$$r - t_q\sigma_r < \rho < r + t_q\sigma_r\,, \quad (9.14)$$

где $r$ - эмпирический коэффициент корреляции;

$t_q$ - параметр нормированной функции Лапласа, определяемый по табл. 1 приложения;

$\sigma_r$ - среднее квадратическое отклонение $r$.

Для больших выборок ($n \geq 30$) можно использовать следующую оценку для среднего квадратического отклонения $r$ от $\rho$:

$$\sigma_r \approx s_r = \frac{1-r^2}{\sqrt{n}}. \qquad (9.15)$$

Подставив выражение (9.15) в неравенство (9.14), получим

$$r - t_q \frac{1-r^2}{\sqrt{n}} < \rho < r + t_q \frac{1-r^2}{\sqrt{n}} \ . \qquad (9.16)$$

**Пример 54.** Оценить коэффициент корреляции с доверительной вероятностью $p = 0{,}95$, если $s_x = 16{,}3''$, $s_y = 11{,}0''$, $r = 0{,}914$ и $n = 140$.

**Решение.** Принимая $\sigma_x \approx s_x$ и $\sigma_y \approx s_y$; $\rho \approx r$, вычисляем среднее квадратическое отклонение коэффициента корреляции

$$s_r = \frac{1-0{,}914^2}{\sqrt{140}} = 0{,}014 \, .$$

Для доверительной вероятности $p = 0{,}95$ по табл. 1 приложения находим $t_q = 1{,}96$. Таким образом, имеем

$$0{,}914 - 1{,}96 \cdot 0{,}014 < \rho < 0{,}914 + 1{,}96 \cdot 0{,}014;$$

$$0{,}887 < \rho < 0{,}941.$$

Вывод: с вероятностью $p = 0{,}95$ коэффициент корреляции будет находиться в указанных границах.

При небольшом числе наблюдений ($n < 30$) распределение эмпирического коэффициента корреляции существенно отличается от нормального. В этом случае для оценки надежности коэффициента корреляции пользуются специальной функцией, так называемым критерием Фишера

$$z = \frac{1}{2} \ln \frac{1+r}{1-r}. \qquad (9.17)$$

Распределение величины $z$ не зависит от значений $\rho$ и $n$. С ростом числа наблюдений распределение быстро приближается к нормальному.

Среднее квадратическое отклонение величины $z$ вычисляется по формуле

$$\sigma_z = \frac{1}{\sqrt{n-3}}. \qquad (9.18)$$

Значения $z$ вычисляются непосредственно по формуле (9.17) или с помощью таблиц (табл. 4 приложения).

**Пример 55.** Оценить с доверительной вероятностью $p = 0{,}95$ коэффициент корреляции $\rho$ по эмпирическому его значению $r = 0{,}79$, если последний определен по выборке объема $n = 15$.

По табл. 4 приложения по значению $r$ = 0,79 находим z = 1,071. По формуле (9.18) определим

$$\sigma_z = \frac{1}{\sqrt{15-3}} = 0{,}29,$$

а затем получим

$$1{,}071 - 1{,}96 \cdot 0{,}29 < z < 1{,}071 + 1{,}96 \cdot 0{,}29;$$

$$0{,}504 < z < 1{,}638.$$

Для $\rho$:

$$0{,}465 < \rho < 0{,}922.$$

Ответ: $0{,}465 < \rho < 0{,}922$ .

## 9.4. Проверка гипотезы об отсутствии корреляционной связи

При корреляционном анализе возникает вопрос о достоверности связи между переменными, т.е. о том , является ли полученный коэффициент корреляции существенным и не объясняется ли его получение случайностями выборки.

Таким образом, возникает необходимость в проверке гипотезы $H_0$, о том, что $\rho = 0$. В качестве критерия в этом случае служит эмпирический коэффициент корреляции $r$ .

При больших объемах выборок согласно формуле (9.15) определяют $\sigma_r$ и строят критическую область

$$|r| > t_q \cdot \sigma_r \qquad (9.19)$$

с заданным уровнем значимости $q$. Если полученное по данным выборки $r$ окажется в данной области, то гипотезу $\rho = 0$ отвергают.

При малых выборках используют критерий Фишера z.

**Пример 56**. Требуется проверить реальность корреляционной связи $r = 0{,}914$ при $q = 0{,}01$, $\sigma_r \approx S_r = 0{,}04$.

**Решение**. Для доверительной вероятности $p$ = 0,99 по табл. 1 приложений находим $t_q = 2{,}58$. Таким образом, подставляя значения в формулу (9.19) имеем

$$|r|_{0{,}914} > 2{,}58 \cdot 0{,}014 = 0{,}036.$$

Вывод: с вероятностью не менее $p$ = 0,99 между переменными действительно существует корреляционная зависимость.

**Пример 57**. Горизонтальный угол измерен с сигнала и со штатива 10 сериями. Коэффициент корреляции $r = 0{,}334$. Проверить существенность корреляционной связи при $q = 0{,}01$.

**Решение**. По формуле (9.18) определим

$$\sigma_z = \frac{1}{\sqrt{10-3}} = 0{,}378.$$

Для доверительной вероятности $P$ = 0,99 по табл. 1приложения находим $t_q$ = 2,58. Таким образом, подставляя значения в формулу (9.19), имеем

$$0{,}347 < 2{,}58 \cdot 0{,}38;$$

$$0{,}347 < 0{,}975.$$

Вывод: коэффициент корреляции считать существенным нельзя и его появление объясняется случайностями выборки.

### 9.5. Понятие регрессии

Наиболее важные особенности стохастической связи находят выражение в тех измерениях, какие испытывает центр условного распределения одной величины при изменении другой.

Геометрическое место центров условных распределений $\bar{y}(x)$ случайной величины $Y$, соответствующих заданным значениям $X = x$ величины $X$, называют линией корреляции или регрессии $Y$ по $X$, т.е.

$$M(Y/X=x) = \bar{y}(x) \text{ - регрессия } Y \text{ по } X$$

и (9.20)

$$M(X/Y=y) = \bar{x}(y) \text{ регрессия } X \text{ по } Y.$$

Функции регрессии выражают математическое ожидание величины $Y$ (или $X$), когда другая величина принимает определенное числовое значение, т.е. функция $M$ ($Y/X = x$) показывает, какое будет в среднем значение случайной величины $Y$, если величина $X$ принимает значение $x$. Это будет справедливо и для функции $M$ ($X/Y = y$).

Таким образом, функция регрессии может быть использована для прогнозирования одной из случайных величин, если известно значение другой случайной величины. Точность такого прогноза определяется дисперсией условного распределения

$$\sigma^2(X/Y=y) = M[(X/Y=y) - M(X/Y=y)]^2;$$
$$\sigma^2(Y/X=x) = M[(Y/X=x) - M(Y/X=x)]^2. \qquad (9.21)$$

Степень стохастической связи можно показать диаграммами рассеивания, изображенными на рис. 18. Координатами точек являются наблюденные значения случайных величин $X$ и $Y$. На рис. 18, *а* - показана тесная стохастическая связь, *б* - слабая стохастическая связь, *в* - диаграмма рассеивания при отсутствии связи между величинами $X$ и $Y$.

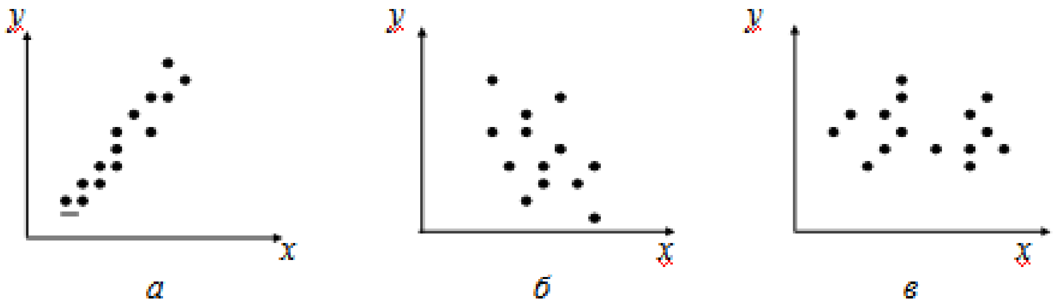


Рис. 18

Возможности практического применения функции регрессии ограничены, так как для ее оценки необходимо знать вид распределения переменных $X$, $Y$. Зная вид этого распределения, можно точно определить вид функции регрессии, а затем оценить его параметры. Однако для этой цели чаще всего располагают лишь выборкой ограниченного объема, что может привести к значительным ошибкам.

Для характеристики формы связи при изучении корреляционной зависимости пользуются понятием кривой регрессии. Кривой регрессии $Y$ по $X$ называется условное среднее значение случайной переменной $Y$, рассматриваемой как функция от $x$, т.е.

$$\bar{y}(x) = f(x).$$

Аналогично и для кривой регрессии $X$ на $Y$. Для определения кривой регрессии условным средним $\bar{y}(x)$ пользуются потому, что эта функция дает наименьшую среднюю ошибку оценки прогноза.

Если функции регрессии $Y$ на $X$ и $X$ на $Y$ линейны, то линии регрессии превращаются в прямые регрессии. Угловые коэффициенты этих прямых выражаются через коэффициент корреляции, который служит также мерой линейной зависимости между величинами.

Прямая регрессия $Y$ на $X$ имеет уравнение

$$Y - b = \rho\frac{\sigma_y}{\sigma_x}(x - a)\,. \tag{9.22}$$

Прямая регрессия $X$ на $Y$ имеет уравнение

$$x - a = \rho\frac{\sigma_x}{\sigma_y}(y - b). \tag{9.23}$$

Коэффициенты

$$\beta_{Y/X} = \rho\frac{\sigma_y}{\sigma_{xx}}; \qquad \beta_{X/Y} = \rho\frac{\sigma_x}{\sigma_y} \tag{9.24}$$

называются коэффициентами регрессии $Y$ на $X$ и $X$ на $Y$. Оба коэффициента имеют тот же знак, что и коэффициент корреляции. Обе прямые проходят через точку ($a$, $b$) - центр распределения. Прямые регрессии $Y$ на $X$ и $X$ на $Y$ совпадают тогда, когда $|\rho| = 1$, т.е. в случае линейной функциональной зависимости между величинами $X$ на $Y$.

По результатам эксперимента параметры уравнений (9.22), (9.23) оцениваются с помощью формул

$$\bar{x} = \frac{1}{n}\sum_1^n x_i; \qquad \bar{y} = \frac{1}{n}\sum_1^n y_i;$$

$$s_x^2 = \frac{1}{n-1}\sum(x - \bar{x})^2; \qquad s_y^2 = \frac{1}{n-1}\sum(y - \bar{y})^2;$$

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{(n-1)\cdot s_x \cdot s_y}.$$

Эти оценки определяют эмпирические прямые регрессии

$$\left.\begin{aligned} Y - \overline{Y} &= r\frac{s_y}{s_x}(x - \overline{x}); \\ X - \overline{X} &= r\frac{s_y}{s_x}(y - \overline{y}), \end{aligned}\right\} \tag{9.25}$$

причем

$$b_{Y/X} = r\frac{s_y}{s_x}; \qquad b_{X/Y} = r\frac{s_x}{s_y} \tag{9.26}$$

эмпирические коэффициенты регрессии имеют тот же знак, что и эмпирический коэффициент корреляции $r$ и связаны соотношением

$$b_{X/Y} \cdot b_{Y/X} = r^2. \tag{9.27}$$

Для коэффициентов регрессии $\beta$ производится интервальная оценка с заданным уровнем значимости

$$b_{X/Y} - tg\frac{s_y}{s_x}\frac{1-r^2}{\sqrt{n}} < \beta_{Y/X} < b_{Y/X} + tg\frac{s_y}{s_x}\frac{1-r^2}{\sqrt{n}}. \tag{9.28}$$

Аналогично и для коэффициентов регрессии $\beta_{X/Y}$.

Кроме того, среднее квадратическое отклонение коэффициентов регрессии при большом объеме выборок можно вычислить по формулам

$$\left.\begin{aligned} \sigma_{\beta_{Y/X}} &= \frac{\sigma_y}{\sigma_x}\sqrt{\frac{1-r^2}{n-3}}; \\ \sigma_{\beta_{X/Y}} &= \frac{\sigma_x}{\sigma_y}\sqrt{\frac{1-r^2}{n-3}}. \end{aligned}\right\} \tag{9.29}$$

**Пример 58.** В табл. 20 приведены расстояния $D$, измеренные светодальномером, и ошибки этих сторон $\Delta_i$ . По данным табл. 20 вычислить коэффициенты корреляции, регрессии, оценить их точность с вероятностью не менее 0,90. Составить уравнение регрессии.

**Решение.** Вычислим эмпирические значения дисперсий

$$s_D = \sqrt{\frac{61,47}{19}} = 1,80; \qquad s_\Delta = \sqrt{\frac{39,20}{19}} = 1,44.$$

Коэффициент корреляции

$$r = \frac{+38,54}{19 \cdot 1,80 \cdot 1,44} = +0,78.$$

Оценим надежность коэффициента корреляции $\rho$. Так как $n$ = 20, то используем критерий Фишера $z$. Поскольку значение $r$ = +0,78 по табл. 4 приложения найдем $z$ = 1,0454.

Вычислим среднее квадратическое отклонение

$$\sigma_z = \frac{1}{\sqrt{20-3}} = 0,243.$$

Для доверительной вероятности $p$ = 0,90 по табл. 1 приложения находим $t_q$ = 1.645. Следовательно, для величины $Z$ определяется следующий интервал:

$$1,0454 - 1,645 \cdot 0,243 < Z < 1,0454 + 1,645 \cdot 0,243;$$

$$1,0454 - 0,3997 < Z < 1,0454 + 0,3997;$$

$$0{,}646 < Z < 1{,}4451.$$

Таблица 20

| N | Результаты вычислений | | Вычисления | | | | |
|---|---|---|---|---|---|---|---|
| | $D$, км | $\lvert\Delta\rvert$, см | $v_D$ | $v_\Delta$ | $v_D \cdot v_D$ | $v_\Delta \cdot v_\Delta$ | $v_D \cdot v_\Delta$ |
| 1 | 8,7 | 7,0 | +3,8 | +3,2 | 14,44 | 10,24 | +12,16 |
| 2 | 3,7 | 3,0 | -1,2 | -0,8 | 1,44 | 0,64 | +0,96 |
| 3 | 6,0 | 4,0 | +1,1 | +0,2 | 1,21 | 0,04 | +0,22 |
| 4 | 3,3 | 3,0 | -1,6 | -0,8 | 2,56 | 0,64 | +1,28 |
| 5 | 5,1 | 4,0 | +0,2 | +0,2 | 0,04 | 0,04 | +0,04 |
| 6 | 6,1 | 4,0 | +1,2 | +0,2 | 1,44 | 0,04 | +0,24 |
| 7 | 2,7 | 3,0 | -2,2 | -0,8 | 4,84 | 0,64 | +1,76 |
| 8 | 4,9 | 4,0 | 0,00 | +0,2 | 0,00 | 0,04 | 0,00 |
| 9 | 3,1 | 4,0 | -1,8 | +0,2 | 3,24 | 0,04 | -0,36 |
| 10 | 3,7 | 2,0 | -1,2 | -1,8 | 1,44 | 3,24 | +2,16 |
| 11 | 5,7 | 6,0 | +0,8 | +2,2 | 0,64 | 4,84 | +1,76 |
| 12 | 4,9 | 5,0 | 0,00 | +1,2 | 0,00 | 1,44 | 0,00 |
| 13 | 5,6 | 3,0 | +0,7 | -0,8 | 0,49 | 0,64 | -0,56 |
| 14 | 7,6 | 4,0 | +2,7 | +0,2 | 7,29 | 0,04 | +0,54 |
| 15 | 4,2 | 3,0 | -0,7 | -0,8 | 0,49 | 0,64 | +0,56 |
| 16 | 2,0 | 2,0 | -2,9 | -1,8 | 8,41 | 3,24 | +5,22 |
| 17 | 4,0 | 2,0 | -0,9 | -1,8 | 0,81 | 3,24 | +1,62 |
| 18 | 6,5 | 5,0 | +1,6 | +1,2 | 2,56 | 1,44 | +1,92 |
| 19 | 7,2 | 6,0 | +2,3 | +2,2 | 5,29 | 4,84 | +5,06 |
| 20 | 2,7 | 2,0 | -2,2 | -1,8 | 4,84 | 3,24 | +3,96 |
| Ср. | 4,9 | 3,8 | +14,4<br>-14,7<br>Σ -0,3 | +11,2<br>-11,2<br>0,0 | 61,47 | 39,20 | +38,54 |

Тогда интервал для неизвестного параметра $\rho$

$$0{,}569 < \rho < 0{,}896\,.$$

Примем нулевую гипотезу $H_0$: $\rho = 0$. Для проверки этой гипотезы воспользуемся критерием Фишера

$$|Z| > t_q \cdot \sigma_Z = 0{,}3997 \approx 0;$$

$$1{,}0454 > 0{,}3997.$$

Таким образом, гипотеза $H_0$: $\rho = 0$ отвергается.

Составим уравнения регрессии $\Delta$ на $D$:

$$\Delta_i - \overline{\Delta} = b_{\Delta/D}(D_i - \overline{D});$$

$$\Delta_i = b_{\Delta/D} D_i + (\overline{\Delta} - b_{\Delta/D}\overline{D}).$$

Подставляя значения, получим

$$\Delta_i = +0{,}78\frac{1{,}44}{1{,}80}D_i + (3{,}8 - 0{,}78\frac{1{,}44}{1{,}80}\cdot 4{,}9);$$

$$\Delta_i = (0{,}62 \cdot D_i + 0{,}70)\text{ см},$$

где $D_i$ - расстояние в км.

Оценим коэффициент регрессии

$$b_{\Delta/D} - t \cdot \frac{s_\Delta}{s_D}\frac{1-r^2}{\sqrt{n}} < \beta_{\Delta/D} < b_{\Delta/D} + t \cdot \frac{s_\Delta}{s_D}\frac{1-r^2}{\sqrt{n}};$$

$$0{,}62 - 0{,}12 < \beta_{\Delta/D} < 0{,}62 + 0{,}12;$$

$$0{,}40 < \beta_{\Delta/D} < 0{,}74.$$

Вычислим среднее квадратическое отклонение для коэффициента регрессии

$$\sigma_{\Delta/D} = \frac{s_\Delta}{s_D}\sqrt{\frac{1-r^2}{n-3}};$$

$$\sigma_{\Delta/D} = \frac{1{,}44}{1{,}80}\sqrt{\frac{1-0{,}78^2}{20-3}} = 0{,}12.$$

Следовательно,

$$\beta_{\Delta/D} \pm \sigma_{b\Delta/D} = 0{,}62 \pm 0{,}12.$$

### 9.6. Понятие о множественной корреляционной связи

При изучении сложных явлений возникает необходимость исследовать влияние нескольких случайных величин (больше двух) друг на друга или на одну из них. Наиболее правильное представление о связи между этими величинами можно получить в том случае, если исследовать одновременно рассматриваемые случайные величины.

Совместное изучение трех или более случайных величин позволяет в то же время выделить влияние на одну из этих случайных величин какой-нибудь другой при некоторых постоянных значениях остальных величин. Таким образом, можно установить более или менее обоснованные предположения о причинных зависимостях между изучаемыми явлениями. Для решения этих вопросов применяется множественная корреляция.

Пусть случайная величина $Y$ зависит от $x_1$, $x_2$, ... , $x_n$. Тогда для изучения связи между переменными составляют уравнение связи между зависимой случайной величиной $Y$ и $n$ переменными $x$. Условимся, что связь эта линейна и форма ее определяется следующим уравнением регрессии:

$$\overline{y}(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ . \ . \ . \ + \beta_n x_n, \qquad (9.30)$$

где $x_1, x_2, \ldots, x_n$ - переменные;

$\beta_0, \beta_1, \ldots, \beta_n$ - неизвестные параметры, называемые коэффициентами регрессии.

Метод, позволяющий по выборке, которая содержит отдельно наблюдавшиеся значения переменных $y$ и $x_1, x_2, \ldots, x_n$, оценить значение неизвестных параметров $\beta_0, \beta_1, \ldots, \beta_n$, называется множественной регрессией. Выражение (9.30) называется уравнением множественной линейной регрессии.

Определение коэффициентов уравнения регрессии получают, исходя из принципа наименьших квадратов

$$F = \sum \left[ y - (\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n) \right]^2 = \min. \qquad (9.31)$$

Дифференцируя это выражение по $\beta_0, \beta_1, \ldots, \beta_n$ и приравнивая каждое из полученных выражений нулю, получают систему нормальных уравнений

$$\left.\begin{array}{l} n\beta_0 + \beta_1 \sum x_1 + \beta_2 \sum x_2 + \ldots + \beta_n \sum x_n = \sum y; \\ \beta_0 \sum x_1 + \beta_1 \sum x_1 x_1 + \beta_2 \sum x_1 x_2 + \ldots + \beta_n \sum x_1 x_n = \sum x_1 y; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \beta_0 \sum x_n + \beta_1 \sum x_1 x_n + \beta_2 \sum x_2 x_n + \ldots + \beta_n \sum x_n x_n = \sum x_n y. \end{array}\right\} \qquad (9.32)$$

Систему (9.32) затем решают любым известным способом. Все переменные нормируют, т.е. вместо $y, x_1, x_2, \ldots, x_n$ запишем

$$t_0 = \frac{y - \bar{y}}{\sigma_0},\ t_1 = \frac{x_1 - \bar{x}_1}{\sigma_1},\ \ldots,\ t_n = \frac{x_n - \bar{x}_n}{\sigma_n}, \qquad (9.33)$$

где $\bar{y}, \bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n$ - средние значения переменных $y, x_1, x_2, \ldots, x_n$;

$\sigma_0, \sigma_1, \ldots, \sigma_n$ - средние квадратические отклонения величин $y, x_1, x_2, \ldots, x_n$.

Так как среднее значение признака равно нулю, а среднее квадратическое отклонение равно единице, то можно записать

$$\bar{t}_1 = \frac{1}{N} \sum_1^N \frac{x_1 - \bar{x}_1}{\sigma_1} = \frac{1}{\sigma_1} \left( \frac{1}{N} \sum_1^N x_1 - \frac{1}{N} \sum_1^N \bar{x}_1 \right) = 0, \qquad (9.34)$$

где $N$ - число значений в выборке;

$$\sigma_{t_1}^2 = \frac{1}{N-1}\sum_1^N \frac{(x_1-\bar{x}_1)^2}{\sigma_1^2} = \frac{1}{\sigma_1^2}\left[\frac{1}{N-1}\sum_1^N (x_1-\bar{x}_1)^2\right], \qquad (9.35)$$

где

$$\frac{1}{N-1}\sum_1^N (x_1-\bar{x}_1)^2 = \sigma_1^2.$$

Таким образом, имеем

$$\sigma_{t_1}^2 = \frac{\sigma_1^2}{\sigma_1^2} = 1 .$$

Коэффициент корреляции

$$r_{12} = \frac{1}{N}\sum_1^N \frac{(t_1-\bar{t}_1)(t_2-\bar{t}_2)}{\sigma_{t_1}\sigma_{t_2}} = \frac{1}{N}\sum \frac{(t_1-0)(t_2-0)}{1\cdot 1} = \frac{1}{N}\sum_1^N t_1 t_2. \qquad (9.36)$$

Уравнение множественной линейной регрессии в нормированном виде выразится следующим образом:

$$\bar{t}_y = a_1 t_1 + a_2 t_2 + \ . \ . \ . \ + a_n t_n, \qquad (9.37)$$

где $t_1, t_2, \ldots t_n$ - нормированные значения переменных $x_1, x_2, \ldots, x_n$;

$\bar{t}_y$ - среднее значение нормированной зависимой переменной;

$a_1, a_2, \ldots, a_n$ - нормированные коэффициенты множественной регрессии, которые находятся из условия

$$\sum (t_y - \bar{t}_y) = \min.$$

С учетом этого условия можно записать следующую систему нормальных уравнений:

$$\left.\begin{aligned}
r_{10} &= a_1 \ + a_2 r_{12} + \ldots + a_n r_{1n};\\
r_{20} &= a_1 r_{21} + a_2 r_{22} + \ldots + a_n r_{2n};\\
&. \ . \ . \ . \ . \ . \ . \ . \ . \ . \ . \ . \ . \ . \ . \ . \ . \ .\\
r_{i0} &= a_1 r_{i1} + a_2 r_{i2} + \ldots + a_j r_{ij} + \ldots + a_n r_{in};\\
&. \ . \ . \ . \ . \ . \ . \ . \ . \ . \ . \ . \ . \ . \ . \ . \ . \ . \ . \ . \ .\\
r_{n0} &= a_1 r_{n1} + a_2 r_{n2} + . \ . . + a_n r_{nn},
\end{aligned}\right\} \qquad (9.38)$$

где $r_{ij}$ - коэффициент корреляции между $i$-й и $j$-й независимыми переменными;

$r_{io}$ - коэффициент корреляции между зависимой переменной $y$ и $i$-й независимой переменной $x$.

Решая систему уравнений (9.38), определим искомые значения коэффициентов $a_1$, $a_2$, ... , $a_n$. Тогда оценки коэффициентов регрессии можно выразить следующим образом:

$$\begin{aligned} \beta_i &= a_i \sigma_y / \sigma_i, \qquad (i = 1,\ 2,\ \ldots,\ n); \\ \beta_0 &= \overline{y} - \beta_1 \overline{x}_1 - \beta_2 \overline{x}_2 - \ \ldots\ - \beta_n \overline{x}_n, \end{aligned} \tag{9.39}$$

где $a_i$ - оценки коэффициентов регрессии в нормированном виде, которые получают из решения системы (9.38).

## Библиографический список


1. Большаков В.Д., Маркузе Ю.И. Практикум по теории математической обработки геодезических измерений.- М.: Недра, 1984.- 352 с.
2. Большаков В.Д., Маркузе Ю.И., Голубев В.В. Уравнивание геодезических построений: Справочное пособие.- М.: Недра, 1989.- 413 с.
3. Губеладзе А.Р. Основы теории ошибок измерений: уч. пособие.- Ростов н/Д: РГСУ, 1998.- 106 с.
4. Губеладзе А.Р. ТМОГИ. Обработка результатов измерений и уравнивание полигонометрических ходов: учебное пособие. Ростов н/Д: Рост. гос. строит. ун-т, 2013. 93 с.
5. Губеладзе А.Р., Яговкина Е.Н. Статистическая обработка результатов измерений: уч. Пособие.- Ростов н/Д: РГСУ, 1999.- 130 с.
6. Губеладзе О.А. Геодезия. Уравнивание нивелирной сети III класса: учебное пособие. Ростов н/Д: Рост. гос. строит. ун-т, 2013. 81 с.
7. Гудков В.М., Хлебников А.В. Математическая обработка маркшейдерско-геодезических измерений: Учебник для вузов.- М.: Недра, 1990.- 335 с.
8. Лесных Н.Б. Теория математической обработки геодезических измерений. Теория ошибок измерений: учеб. пособие. Новосибирск.: СГГА, 2010. 43 с.
9. Маркузе Ю.И., Бойко Е.Г., Голубев В.В. Геодезия. Вычисление и уравнивание геодезических сетей: Справочное пособие.- М.: Картогеоцентр-Геодезиздат, 1994.- 431 с.

ПРИЛОЖЕНИЕ

Таблица 1

Значение функции $\Phi(t) = \dfrac{2}{\sqrt{2\pi}} \int\limits_0^t e^{-\frac{t^2}{2}} dt$

| *t* | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0,0 | 0000 | 0080 | 0160 | 0239 | 0319 | 0399 | 0478 | 0558 | 0638 | 0717 |
| 0,1 | 0797 | 0876 | 0955 | 1034 | 1113 | 1192 | 1271 | 1350 | 1428 | 1507 |
| 0,2 | 1585 | 1663 | 1741 | 1819 | 1897 | 1974 | 2051 | 2128 | 2205 | 2282 |
| 0,3 | 2358 | 2434 | 2510 | 2586 | 2661 | 2737 | 2812 | 2886 | 2960 | 3035 |
| 0,4 | 3108 | 3182 | 3255 | 3328 | 3401 | 3473 | 3545 | 3616 | 3688 | 3759 |
| 0,5 | 3829 | 3900 | 3969 | 4039 | 4108 | 4177 | 4245 | 4313 | 4381 | 4448 |
| 0,6 | 4515 | 4581 | 4647 | 4713 | 4778 | 4843 | 4908 | 4971 | 5035 | 5098 |
| 0,7 | 5161 | 5223 | 5285 | 5346 | 5407 | 5468 | 5528 | 5587 | 5646 | 5705 |
| 0,8 | 5763 | 5821 | 5878 | 5935 | 5991 | 6047 | 6102 | 6157 | 6211 | 6265 |
| 0,9 | 6319 | 6372 | 6424 | 6476 | 6528 | 6579 | 6629 | 6680 | 6729 | 6778 |
| 1,0 | 6827 | 6875 | 6923 | 6970 | 7017 | 7063 | 7109 | 7154 | 7199 | 7243 |
| 1,1 | 7287 | 7330 | 7373 | 7415 | 7457 | 7499 | 7540 | 7580 | 7620 | 7960 |
| 1,2 | 7699 | 7737 | 7775 | 7813 | 7850 | 7887 | 7923 | 7959 | 7994 | 8030 |
| 1,3 | 8064 | 8098 | 8132 | 8165 | 8198 | 8230 | 8262 | 8293 | 8324 | 8355 |
| 1,4 | 8385 | 8415 | 8444 | 8473 | 8501 | 8529 | 8557 | 8584 | 8611 | 8638 |
| 1,5 | 8664 | 8690 | 8715 | 8740 | 8764 | 8789 | 8812 | 8836 | 8859 | 8882 |
| 1,6 | 8904 | 8926 | 8948 | 8969 | 8990 | 9011 | 9031 | 9051 | 9070 | 9090 |
| 1,7 | 9109 | 9127 | 9146 | 9164 | 9181 | 9199 | 9216 | 9233 | 9249 | 9266 |
| 1,8 | 9281 | 9297 | 9312 | 9328 | 9342 | 9357 | 9371 | 9385 | 9399 | 9412 |
| 1,9 | 9426 | 9439 | 9451 | 9464 | 9476 | 9488 | 9500 | 9512 | 9523 | 9534 |
| 2,0 | 9545 | 9556 | 9566 | 9576 | 9586 | 9596 | 9606 | 9616 | 9625 | 9634 |
| 2,1 | 9643 | 9651 | 9660 | 9658 | 9676 | 9684 | 9692 | 9700 | 9707 | 9715 |
| 2,2 | 9722 | 9729 | 9736 | 9742 | 9749 | 9756 | 9762 | 9768 | 9774 | 9780 |
| 2,3 | 9786 | 9791 | 9797 | 9802 | 9807 | 9812 | 9817 | 9822 | 9827 | 9832 |
| 2,4 | 9836 | 9840 | 9845 | 9849 | 9853 | 9857 | 9861 | 9865 | 9869 | 9872 |
| 2,5 | 9876 | 9879 | 9883 | 9886 | 9889 | 9892 | 9895 | 9898 | 9901 | 9904 |
| 2,6 | 9907 | 9910 | 9912 | 9915 | 9917 | 9920 | 9922 | 9924 | 9926 | 9928 |
| 2,7 | 9931 | 9933 | 9935 | 9937 | 9939 | 9940 | 9942 | 9944 | 9946 | 9947 |
| 2,8 | 9949 | 9950 | 9952 | 9954 | 9955 | 9956 | 9958 | 9959 | 9960 | 9962 |
| 2,9 | 9963 | 9964 | 9965 | 9966 | 9967 | 9968 | 9969 | 9970 | 9971 | 9972 |
| 3,0 | 9973 | 9974 | 9975 | 9976 | 9976 | 9977 | 9978 | 9979 | 9979 | 9980 |
| 3,1 | 9981 | 9981 | 9982 | 9982 | 9983 | 9984 | 9984 | 9985 | 9985 | 9986 |
| 3,2 | 9986 | 9987 | 9987 | 9988 | 9988 | 9988 | 9989 | 9989 | 9990 | 9990 |
| 3,3 | 9990 | 9991 | 9991 | 9991 | 9992 | 9992 | 9992 | 9992 | 9993 | 9993 |
| 3,4 | 9993 | 9994 | 9994 | 9994 | 9994 | 9994 | 9995 | 9995 | 9995 | 9995 |
| 3,5 | 9995 | 9996 | 9996 | 9996 | 9996 | 9996 | 9996 | 9996 | 9997 | 9997 |
| 3,6 | 9997 | 9997 | 9997 | 9997 | 9997 | 9997 | 9998 | 9998 | 9998 | 9998 |
| 3,7 | 9998 | 9998 | 9998 | 9998 | 9998 | 9998 | 9998 | 9998 | 9998 | 9998 |
| 3,8 | 9999 | 9999 | 9999 | 9999 | 9999 | 9999 | 9999 | 9999 | 9999 | 9999 |
| 3,9 | 9999 | 9999 | 9999 | 9999 | 9999 | 9999 | 9999 | 9999 | 9999 | 9999 |
| 4,0 | 9999 | - | - | - | - | - | - | - | - | - |

Продолжение прил.

Таблица 2

Квантили распределения Стьюдента

| $K$ | $q$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 0,2 | 0,1 | 0,05 | 0,02 | 0,01 | 0,005 | 0,002 | 0,001 |
| 1 | 3,08 | 6,31 | 12,71 | 31,82 | 63,66 | 127,32 | 318,30 | 636,61 |
| 2 | 1,89 | 2,92 | 4,30 | 6,96 | 9,92 | 14,09 | 22,33 | 31,60 |
| 3 | 1,64 | 2,35 | 3,18 | 4,54 | 5,84 | 7,45 | 10,21 | 12,92 |
| 4 | 1,53 | 2,13 | 2,78 | 3,75 | 4,60 | 5,60 | 7,17 | 8,61 |
| 5 | 1,48 | 2,02 | 2,57 | 3,36 | 4,03 | 4,77 | 5,89 | 6,87 |
| 6 | 1,44 | 1,94 | 2,45 | 3,14 | 3,71 | 4,32 | 5,21 | 5,96 |
| 7 | 1,41 | 1,89 | 2,36 | 3,00 | 3,50 | 4,03 | 4,79 | 5,41 |
| 8 | 1,40 | 1,86 | 2,31 | 2,90 | 3,36 | 3,83 | 4,50 | 5,04 |
| 9 | 1,38 | 1,83 | 2,26 | 2,82 | 3,25 | 3,69 | 4,30 | 4,78 |
| 10 | 1,37 | 1,81 | 2,23 | 2,76 | 3,17 | 3,58 | 4,14 | 4,59 |
| 11 | 1,36 | 1,80 | 2,20 | 2,72 | 3,11 | 3,50 | 4,02 | 4,44 |
| 12 | 1,36 | 1,78 | 2,18 | 2,68 | 3,05 | 3,43 | 3,93 | 4,32 |
| 13 | 1,35 | 1,77 | 2,16 | 2,65 | 3,01 | 3,37 | 3,85 | 4,22 |
| 14 | 1,34 | 1,76 | 2,14 | 2,62 | 2,98 | 3,33 | 3,79 | 4,14 |
| 15 | 1,34 | 1,75 | 2,13 | 2,60 | 2,95 | 3,29 | 3,73 | 4,07 |
| 18 | 1,33 | 1,73 | 2,10 | 2,55 | 2,88 | 3,20 | 3,61 | 3,92 |
| 20 | 1,33 | 1,72 | 2,09 | 2,53 | 2,85 | 3,15 | 3,55 | 3,85 |
| 23 | 1,32 | 1,71 | 2,07 | 2,50 | 2,81 | 3,10 | 3,48 | 3,77 |
| 25 | 1,32 | 1,71 | 2,06 | 2,49 | 2,79 | 3,08 | 3,45 | 3,73 |
| 28 | 1,31 | 1,70 | 2,05 | 2,47 | 2,76 | 3,05 | 3,41 | 3,67 |
| 30 | 1,31 | 1,70 | 2,04 | 2,46 | 2,75 | 3,03 | 3,39 | 3,65 |
| 120 | 1,29 | 1,66 | 1,98 | 2,36 | 2,62 | 2,86 | 3,16 | 3,37 |
| $\infty$ | 1,28 | 1,64 | 1,96 | 2,33 | 2,58 | 2,81 | 3,09 | 3,29 |

Таблица 3

Квантили распределения $\chi^2$

| $K$ | $q$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 0,990 | 0,975 | 0,950 | 0,900 | 0,100 | 0,050 | 0,025 | 0,010 |
| 1 | 0,0002 | 0,0010 | 0,0039 | 0,0158 | 2,71 | 3,84 | 5,02 | 6,63 |
| 2 | 0,0201 | 0,0506 | 0,103 | 0,211 | 4,61 | 5,99 | 7,38 | 9,21 |
| 3 | 0,115 | 0,216 | 0,352 | 0,548 | 6,25 | 7,81 | 9,35 | 11,34 |
| 4 | 0,297 | 0,484 | 0,711 | 1,06 | 7,78 | 9,49 | 11,14 | 13,28 |
| 5 | 0,554 | 0,831 | 1,15 | 1,61 | 9,24 | 11,07 | 12,83 | 15,09 |
| 6 | 0,872 | 1,24 | 1,64 | 2,20 | 10,64 | 12,59 | 14,45 | 16,81 |
| 7 | 1,24 | 1,69 | 2,17 | 2,83 | 12,02 | 14,07 | 16,01 | 18,48 |
| 8 | 1,65 | 2,18 | 2,73 | 3,49 | 13,36 | 15,51 | 17,53 | 20,09 |
| 9 | 2,09 | 2,70 | 3,33 | 4,17 | 14,68 | 16,92 | 19,02 | 21,67 |
| 10 | 2,56 | 3,25 | 3,94 | 4,87 | 15,99 | 18,31 | 20,48 | 23,21 |
| 11 | 3,05 | 3,82 | 4,57 | 5,58 | 17,28 | 19,68 | 21,92 | 24,73 |
| 12 | 3,57 | 4,40 | 5,23 | 6,30 | 18,55 | 21,03 | 23,34 | 26,22 |
| 13 | 4,11 | 5,01 | 5,89 | 7,04 | 19,81 | 22,36 | 24,74 | 27,69 |
| 14 | 4,66 | 5,63 | 6,57 | 7,79 | 21,06 | 23,68 | 26,12 | 29,14 |

Продолжение прил.

Окончание табл. 3

| *K* | *q* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 0,990 | 0,975 | 0,950 | 0,900 | 0,100 | 0,050 | 0,025 | 0,010 |
| 15 | 5,23 | 6,26 | 7,26 | 8,55 | 22,31 | 25,00 | 27,49 | 30,58 |
| 16 | 5,81 | 6,91 | 7,96 | 9,31 | 23,54 | 26,30 | 28,85 | 32,00 |
| 17 | 6,41 | 7,56 | 8,67 | 10,08 | 24,77 | 27,59 | 30,19 | 33,41 |
| 18 | 7,01 | 8,23 | 9,39 | 10,86 | 25,99 | 28,87 | 31,53 | 34,81 |
| 19 | 7,63 | 8,91 | 10,12 | 11,65 | 27,20 | 30,14 | 32,82 | 36,19 |
| 20 | 8,26 | 9,59 | 10,85 | 12,44 | 28,41 | 31,41 | 34,17 | 37,57 |
| 21 | 8,90 | 10,28 | 11,59 | 13,24 | 29,62 | 32,67 | 35,48 | 38,93 |
| 22 | 9,54 | 10,98 | 12,34 | 14,04 | 30,81 | 33,92 | 36,78 | 40,29 |
| 23 | 10,20 | 11,69 | 13,09 | 14,85 | 32,01 | 35,17 | 38,08 | 41,64 |
| 24 | 10,86 | 12,40 | 13,85 | 15,66 | 33,20 | 36,42 | 39,36 | 42,98 |
| 25 | 11,52 | 13,12 | 14,61 | 16,47 | 34,38 | 37,65 | 40,65 | 44,31 |
| 26 | 12,20 | 13,84 | 15,38 | 17,29 | 35,56 | 38,88 | 41,92 | 45,64 |
| 27 | 12,88 | 14,57 | 16,15 | 18,11 | 36,74 | 40,11 | 43,19 | 46,96 |
| 28 | 13,56 | 15,31 | 16,93 | 18,94 | 37,92 | 41,34 | 44,46 | 48,28 |
| 29 | 14,26 | 16,05 | 17,71 | 19,77 | 39,09 | 42,56 | 45,72 | 49,59 |
| 30 | 14,95 | 16,79 | 18,49 | 20,60 | 40,26 | 43,77 | 46,98 | 50,89 |
| 40 | 22,16 | 24,43 | 26,51 | 29,05 | 51,81 | 55,76 | 59,34 | 63,69 |
| 60 | 37,48 | 40,48 | 43,19 | 46,46 | 74,40 | 79,08 | 83,30 | 88,91 |
| 120 | 86,92 | 91,58 | 95,70 | 100,62 | 140,23 | 146,57 | 152,21 | 158,95 |

Таблица 4

Значения $Z=\frac{1}{2}\ln\frac{1+r}{1-r}$, отвечающие значениям $r$

| *r* | Сотые доли[1] | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 0,0 | 0,0000 | 0,0100 | 0,0200 | 0,0300 | 0,0400 | 0,0500 | 0,0601 | 0,0701 | 0,0802 | 0,0902 |
| 0,1 | 0,1003 | 0,1104 | 0,1206 | 0,1307 | 0,1409 | 0,1511 | 0,1614 | 0,1717 | 0,1820 | 0,1923 |
| 0,2 | 0,2027 | 0,2132 | 0,2237 | 0,2342 | 0,2448 | 0,2554 | 0,2661 | 0,2769 | 0,2877 | 0,2986 |
| 0,3 | 0,3095 | 0,3205 | 0,3316 | 0,3428 | 0,3541 | 0,3654 | 0,3769 | 0,3884 | 0,4001 | 0,4118 |
| 0,4 | 0,4236 | 0,4356 | 0,4477 | 0,4599 | 0,4722 | 0,4847 | 0,4973 | 0,5101 | 0,5230 | 0,5361 |
| 0,5 | 0,5493 | 0,5627 | 0,5763 | 0,5901 | 0,6042 | 0,6184 | 0,6328 | 0,6475 | 0,6625 | 0,6777 |
| 0,6 | 0,6931 | 0,7089 | 0,7250 | 0,7414 | 0,7582 | 0,7753 | 0,7928 | 0,8107 | 0,8291 | 0,8480 |
| 0,7 | 0,8673 | 0,8872 | 0,9076 | 0,9287 | 0,9505 | 0,9730 | 0,9962 | 1,0203 | 1,0454 | 1,0714 |
| 0,8 | 1,0986 | 1,1270 | 1,1568 | 1,1881 | 1,2212 | 1,2562 | 1,2933 | 1,3331 | 1,3758 | 1,4219 |
| 0,9 | 1,4722 | 1,5275 | 1,5890 | 1,6584 | 1,7380 | 1,8318 | 1,9459 | 2,0923 | 2,2976 | 2,6466 |
| 0,99 | 2,6466 | 2,6996 | 2,7587 | 2,8257 | 2,9031 | 2,9945 | 3,1063 | 3,2504 | 3,4534 | 3,8002 |

[1] Для последней строки - это тысячные доли.

Продолжение прил.

Таблица 5

Квантили $F$- распределения

| $k_2$ \ $k_1$ | 1 | 2 | 3 | 4 | 5 | 6 | 8 | 12 | 24 | ∞ |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | При | $q =$ | 0,05 | | | | |
| 1 | 161,4 | 199,5 | 215,7 | 224,6 | 230,2 | 234,0 | 238,9 | 243,9 | 249,0 | 254,3 |
| 2 | 18,51 | 19,00 | 19,16 | 19,25 | 19,30 | 19,33 | 19,37 | 19,41 | 19,45 | 19,50 |
| 3 | 10,13 | 9,55 | 9,28 | 9,12 | 9,01 | 8,94 | 8,84 | 8,74 | 8,64 | 8,53 |
| 4 | 7,71 | 6,94 | 6,59 | 6,39 | 6,26 | 6,16 | 6,06 | 5,91 | 5,77 | 6,63 |
| 5 | 6,61 | 5,79 | 5,41 | 5,19 | 5,05 | 4,95 | 4,82 | 4,68 | 4,53 | 4,36 |
| 10 | 4,96 | 4,10 | 3,71 | 3,48 | 3,33 | 3,22 | 3,07 | 2,91 | 2,74 | 2,54 |
| 15 | 4,54 | 3,68 | 3,29 | 3,06 | 2,90 | 2,79 | 2,64 | 2,48 | 2,29 | 2,07 |
| 20 | 4,35 | 3,49 | 3,10 | 2,87 | 2,71 | 2,60 | 2,45 | 2,28 | 2,08 | 1,84 |
| 25 | 4,24 | 3,38 | 2,99 | 2,76 | 2,60 | 2,49 | 2,34 | 2,16 | 1,96 | 1,71 |
| 30 | 4,17 | 3,32 | 2,92 | 2,69 | 2,53 | 2,42 | 2,27 | 2,09 | 1,89 | 1,62 |
| 40 | 4,08 | 3,23 | 2,84 | 2,61 | 2,45 | 2,34 | 2,18 | 2,00 | 1,79 | 1,52 |
| 60 | 4,00 | 3,15 | 2,76 | 2,52 | 2,37 | 2,25 | 2,10 | 1,92 | 1,70 | 1,39 |
| 120 | 3,92 | 3,07 | 2,68 | 2,45 | 2,29 | 2,17 | 2,02 | 1,83 | 1,61 | 1,25 |
| ∞ | 3,84 | 2,99 | 2,60 | 2,37 | 2,21 | 2,09 | 1,94 | 1,75 | 1,52 | 1,00 |
| | | | | При | $q=$ | 0,025 | | | | |
| 1 | 64,78 | 799,5 | 864,2 | 899,6 | 921,8 | 937,1 | 956,7 | 976,7 | 997,2 | 1018,3 |
| 2 | 38,51 | 39,00 | 39,16 | 39,25 | 39,30 | 39,33 | 39,37 | 39,42 | 39,46 | 39,50 |
| 3 | 17,43 | 16,04 | 15,44 | 15,10 | 14,80 | 14,74 | 14,54 | 14,34 | 14,12 | 13,90 |
| 4 | 12,22 | 10,65 | 9,98 | 9,60 | 9,36 | 9,20 | 8,98 | 8,75 | 8,51 | 8,26 |
| 5 | 10,01 | 8,43 | 7,76 | 7,39 | 7,15 | 6,98 | 6,76 | 6,52 | 6,28 | 6,02 |
| 10 | 6,94 | 5,46 | 4,83 | 4,47 | 4,24 | 4,07 | 3,86 | 3,62 | 3,36 | 3,08 |
| 15 | 6,20 | 4,76 | 4,15 | 3,80 | 3,58 | 3,42 | 3,20 | 2,96 | 2,70 | 2,40 |
| 20 | 5,87 | 4,46 | 3,86 | 3,52 | 3,29 | 3,13 | 2,91 | 2,68 | 2,41 | 2,08 |
| 25 | 5,69 | 4,29 | 3,69 | 3,35 | 3,13 | 2,97 | 2,75 | 2,52 | 2,21 | 1,91 |
| 30 | 5,57 | 4,18 | 3,59 | 3,25 | 3,03 | 2,87 | 2,65 | 2,41 | 3,14 | 1,79 |
| 40 | 5,42 | 4,05 | 3,46 | 3,13 | 2,90 | 2,74 | 2,53 | 2,29 | 2,01 | 1,64 |
| 60 | 5,29 | 3,92 | 3,34 | 3,01 | 2,79 | 2,63 | 2,41 | 2,17 | 1,88 | 1,48 |
| 120 | 5,15 | 3,80 | 3,23 | 2,89 | 2,67 | 2,52 | 2,30 | 2,06 | 1,76 | 1,31 |
| ∞ | 5,02 | 3,69 | 3,12 | 2,79 | 2,57 | 2,41 | 2,19 | 1,94 | 1,64 | 1,00 |
| | | | | При | $q=$ | 0,01 | | | | |
| 1 | 4052 | 4999 | 5403 | 5625 | 5764 | 5859 | 5981 | 6106 | 6234 | 6366 |
| 2 | 98,49 | 99,00 | 99,17 | 99,25 | 99,30 | 99,33 | 99,36 | 99,42 | 99,46 | 99,60 |
| 3 | 34,12 | 30,81 | 29,46 | 28,71 | 28,24 | 27,91 | 27,49 | 27,05 | 26,60 | 26,42 |
| 4 | 21,20 | 18,00 | 16,69 | 15,98 | 15,52 | 15,21 | 14,80 | 14,37 | 13,93 | 13,45 |
| 5 | 16,26 | 13,27 | 12,05 | 11,39 | 10,97 | 10,67 | 10,29 | 9,89 | 9,47 | 9,02 |
| 10 | 10,04 | 7,56 | 6,55 | 5,99 | 5,64 | 5,39 | 5,06 | 4,71 | 4,33 | 3,91 |
| 15 | 8,68 | 6,36 | 5,42 | 4,89 | 4,56 | 4,32 | 4,00 | 3,67 | 3,29 | 2,87 |
| 20 | 8,10 | 5,85 | 4,94 | 4,43 | 4,10 | 3,87 | 3,56 | 3,23 | 2,86 | 2,42 |
| 25 | 7,77 | 5,57 | 4,68 | 4,18 | 3,86 | 3,63 | 3,32 | 2,99 | 2,62 | 2,17 |

Продолжение прил.

Окончание табл. 5

| $k_2$ \ $k_1$ | 1 | 2 | 3 | 4 | 5 | 6 | 8 | 12 | 24 | ∞ |
|---|---|---|---|---|---|---|---|---|---|---|
| 30 | 7,56 | 5,39 | 4,51 | 4,02 | 3,70 | 3,47 | 3,17 | 2,84 | 2,47 | 2,01 |
| 40 | 7,31 | 5,18 | 4,31 | 3,83 | 3,51 | 3,29 | 2,99 | 2,66 | 2,29 | 1,80 |
| 60 | 7,08 | 4,98 | 4,13 | 3,65 | 3,34 | 3,12 | 2,82 | 2,50 | 2,12 | 2,60 |
| 120 | 6,85 | 4,76 | 3,95 | 3,48 | 3,17 | 2,96 | 2,66 | 3,34 | 1,95 | 1,38 |
| ∞ | 6,64 | 4,60 | 3,78 | 3,32 | 3,02 | 2,80 | 2,51 | 2,18 | 1,79 | 1,00 |
| | | | | При | $q$= | 0,005 | | | | |
| 1 | 16211 | 20000 | 21615 | 22500 | 23056 | 23437 | 23925 | 24426 | 24940 | 25465 |
| 2 | 198,5 | 199,0 | 199,2 | 199,3 | 199,34 | 199,3 | 199,4 | 199,4 | 199,5 | 199,5 |
| 3 | 55,55 | 49,80 | 47,47 | 46,20 | 45,39 | 44,84 | 44,13 | 43,39 | 42,62 | 41,88 |
| 4 | 31,33 | 26,28 | 24,26 | 23,16 | 22,46 | 21,98 | 21,35 | 20,70 | 20,03 | 19,32 |
| 5 | 22,78 | 18,31 | 16,53 | 15,56 | 14,94 | 14,51 | 13,96 | 13,38 | 12,78 | 12,14 |
| 10 | 12,83 | 9,43 | 8,08 | 7,34 | 6,87 | 6,54 | 6,12 | 5,66 | 5,17 | 4,64 |
| 15 | 10,80 | 7,70 | 6,48 | 5,80 | 5,37 | 5,07 | 4,67 | 4,25 | 3,79 | 3,26 |
| 20 | 9,94 | 6,99 | 5,82 | 5,17 | 4,76 | 4,47 | 4,09 | 3,68 | 3,22 | 2,69 |
| 25 | 9,48 | 6,60 | 5,46 | 4,84 | 4,43 | 4,15 | 3,78 | 3,37 | 2,92 | 2,38 |
| 30 | 9,18 | 6,36 | 5,24 | 4,62 | 4,23 | 3,95 | 3,58 | 3,18 | 2,73 | 2,18 |
| 40 | 8,83 | 6,07 | 4,98 | 4,37 | 3,99 | 3,71 | 3,35 | 2,95 | 2,50 | 1,93 |
| 60 | 8,50 | 5,80 | 4,73 | 4,14 | 3,76 | 3,49 | 3,13 | 2,74 | 2,29 | 1,69 |
| 120 | 8,18 | 5,54 | 4,50 | 3,92 | 3,55 | 3,28 | 2,93 | 2,54 | 2,09 | 1,43 |
| ∞ | 7,88 | 5,30 | 4,28 | 3,72 | 3,35 | 3,09 | 2,74 | 2,36 | 1,90 | 1,00 |

Таблица 6

Квантили распределения $\lambda$ (Колмогорова)

| $\lambda$ | 1-k($\lambda$) | $\lambda$ | 1-k($\lambda$) | $\lambda$ | 1-k ($\lambda$) | $\lambda$ | 1-k($\lambda$) |
|---|---|---|---|---|---|---|---|
| 0,30 | 1,000 | 0,65 | 0,792 | 1,00 | 0,270 | 1,70 | 0,0062 |
| 0,35 | 0,999 | 0,70 | 711 | 1,10 | 178 | 1,80 | 0032 |
| 0,40 | 997 | 0,75 | 627 | 1,20 | 112 | 1,90 | 0015 |
| 0,45 | 984 | 0,80 | 544 | 1,30 | 068 | 2,00 | 0007 |
| 0,50 | 964 | 0,85 | 465 | 1,40 | 040 | 2,10 | 0003 |
| 0,55 | 923 | 0,90 | 393 | 1,50 | 022 | 2,30 | 0001 |
| 0,60 | 864 | 0,95 | 328 | 1,60 | 012 | 2,50 | 0000 |

Продолжение прил.

Таблица 7

Квантили распределения Кочрена

| $k$ | $n$ - 1 | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 8 | 10 | 16 | 36 | 144 | $\infty$ |
| При $q$ = 0,05 | | | | | | | | | | | |
| 2 | 0,9985 | 9750 | 9392 | 9057 | 8772 | 8159 | 7880 | 7341 | 6602 | 5813 | 5000 |
| 3 | 9969 | 8709 | 7977 | 7457 | 7071 | 6333 | 6025 | 5466 | 4748 | 4031 | 3333 |
| 4 | 9065 | 7679 | 6841 | 6287 | 5895 | 5175 | 4884 | 4366 | 3720 | 3093 | 2500 |
| 5 | 8412 | 6838 | 5981 | 5441 | 5065 | 4387 | 4118 | 3645 | 3066 | 2513 | 2000 |
| 6 | 7808 | 6161 | 5321 | 4803 | 4447 | 3817 | 3568 | 3135 | 2612 | 2119 | 1667 |
| 7 | 7271 | 5612 | 4800 | 4307 | 3974 | 3384 | 3154 | 2756 | 2278 | 1833 | 1420 |
| 8 | 6798 | 5157 | 4377 | 3910 | 3595 | 3043 | 2829 | 2462 | 2022 | 1616 | 1250 |
| 9 | 6385 | 4775 | 4027 | 3584 | 3286 | 2768 | 2568 | 2226 | 1820 | 1446 | 1111 |
| 10 | 6020 | 4450 | 3733 | 3311 | 3029 | 2541 | 2353 | 2032 | 1655 | 1308 | 1000 |
| 15 | 4709 | 3346 | 2758 | 2419 | 2195 | 1815 | 1671 | 1429 | 1144 | 0889 | 0667 |
| 20 | 3894 | 2705 | 2205 | 1921 | 1735 | 1422 | 1303 | 1108 | 0879 | 0675 | 0500 |
| 30 | 2929 | 1980 | 1593 | 1377 | 1237 | 1002 | 0921 | 0771 | 0604 | 0457 | 0333 |
| 40 | 2370 | 1576 | 1259 | 1082 | 0968 | 0780 | 0713 | 0595 | 0462 | 0347 | 0250 |
| 60 | 1737 | 1131 | 0895 | 0765 | 0682 | 0552 | 0497 | 0411 | 0316 | 0234 | 0167 |
| 120 | 0998 | 0632 | 0295 | 0419 | 0371 | 0292 | 0266 | 0218 | 0165 | 0120 | 0083 |
| $\infty$ | 0000 | 0000 | 0000 | 0000 | 0000 | 0000 | 0000 | 0000 | 0000 | 0000 | 0000 |
| При $q$ = 0,01 | | | | | | | | | | | |
| 2 | 0,9999 | 9950 | 9794 | 9586 | 9373 | 8823 | 8539 | 7949 | 7067 | 6062 | 5000 |
| 3 | 9933 | 9423 | 8831 | 8335 | 7933 | 7107 | 6743 | 6059 | 5153 | 4230 | 3333 |
| 4 | 9676 | 8643 | 7814 | 7212 | 6761 | 5897 | 5536 | 4884 | 4057 | 3252 | 2500 |
| 5 | 9279 | 7885 | 6957 | 6329 | 5875 | 5037 | 4697 | 4094 | 3351 | 2644 | 2000 |
| 6 | 8828 | 7218 | 6258 | 5635 | 5195 | 4401 | 4084 | 3529 | 2858 | 2229 | 1667 |
| 7 | 8376 | 6644 | 5685 | 5080 | 4659 | 3911 | 3616 | 3105 | 2494 | 1929 | 1429 |
| 8 | 7945 | 6152 | 5209 | 4627 | 4226 | 3522 | 3248 | 2779 | 2214 | 1700 | 1250 |
| 9 | 7544 | 5727 | 4810 | 4251 | 3870 | 3207 | 2950 | 2514 | 1992 | 1521 | 1111 |
| 10 | 7175 | 5358 | 4469 | 3934 | 3572 | 2945 | 2704 | 2297 | 1811 | 1376 | 1000 |
| 15 | 5747 | 4069 | 3317 | 2882 | 2593 | 2104 | 1918 | 1612 | 1251 | 0934 | 0667 |
| 20 | 4799 | 3297 | 2654 | 2288 | 2048 | 1646 | 1501 | 1248 | 0960 | 0709 | 0500 |
| 30 | 3632 | 2412 | 1913 | 1635 | 1454 | 1157 | 1054 | 0867 | 0658 | 0480 | 0333 |
| 40 | 2940 | 1915 | 1508 | 1281 | 1135 | 0898 | 0816 | 0668 | 0503 | 0363 | 0250 |
| 60 | 1251 | 1371 | 1069 | 0902 | 0796 | 0625 | 0567 | 0461 | 0344 | 0245 | 0157 |
| 120 | 1225 | 0759 | 0585 | 0489 | 0429 | 0334 | 0302 | 0242 | 0178 | 0125 | 0083 |
| $\infty$ | 0000 | 0000 | 0000 | 0000 | 0000 | 0000 | 0000 | 0000 | 0000 | 0000 | 00000 |

Продолжение прил.

Таблица 8

Критерий Аббе

| *n* | *q* | | | *n* | *q* | | |
|---|---|---|---|---|---|---|---|
| | 0,001 | 0,01 | 0,05 | | 0,001 | 0,05 | 0,01 |
| 4 | 0,2919 | 0,3128 | 0,3902 | 32 | 0,4963 | 0,6089 | 0,7177 |
| 6 | 1817 | 2808 | 4451 | 34 | 5090 | 6193 | 7256 |
| 8 | 2018 | 3314 | 4912 | 36 | 5208 | 6290 | 7323 |
| 10 | 2408 | 3759 | 5311 | 38 | 5319 | 6381 | 7396 |
| 12 | 2778 | 4140 | 5638 | 40 | 5425 | 6467 | 7461 |
| 14 | 3112 | 4466 | 5908 | 42 | 5524 | 6548 | 7521 |
| 16 | 3413 | 4746 | 6137 | 44 | 5616 | 6622 | 7576 |
| 18 | 3684 | 4989 | 6330 | 46 | 5701 | 6693 | 7628 |
| 20 | 3926 | 5203 | 6498 | 48 | 5781 | 6757 | 7676 |
| 22 | 4142 | 5393 | 6645 | 50 | 5853 | 6814 | 7718 |
| 24 | 4334 | 5562 | 6776 | 52 | 5922 | 6869 | 7759 |
| 26 | 4509 | 5713 | 6893 | 54 | 5989 | 6924 | 7799 |
| 28 | 4670 | 5850 | 6996 | 56 | 6051 | 6974 | 7836 |
| 30 | 4822 | 5975 | 7091 | 60 | 6174 | 7071 | 7906 |

**О г л а в л е н и е**

Учебное издание

Губеладзе Автандил Рубенович

Яговкина Елена Николаевна

Губеладзе Иван Олегович

**Введение в теорию ошибок измерений**

Учебное пособие

Редактор *М.А. Цыганова*
*Компьютерная верстка и макет А.Р. Губеладзе, И.О. Губеладзе*
Тем. план 2015 г., поз.

---

Подписано в печать 5.03.2015г. Формат 60х84/16. Бумага писчая. Ризограф.
Уч.-изд. л. 1,2 Тираж 3 экз. Заказ

---

Редакционно-издательский центр
Ростовского государственного строительного университета
344022, Ростов н/Д., ул. Социалистическая, 162