МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ
Белорусский национальный технический университет

Кафедра «Геодезия и аэрокосмические геотехнологии»

А. Ю. Будо

ТЕОРИЯ ПОГРЕШНОСТЕЙ ИЗМЕРЕНИЙ

Учебно-методическое пособие
для студентов специальности
1-56 02 01 «Геодезия»

Рекомендовано учебно-методическим объединением по образованию в области горнодобывающей промышленности

Минск
БНТУ
2022

УДК 528.1.06:51(075.8)
ББК 26.1я7
Б90

Р е ц е н з е н т ы :
Д. В. Чадович, А. П. Романкевич

Будо, А. Ю.

Б90 Теория погрешностей измерений : учебно-методическое пособие для студентов специальности 1-56 02 01 «Геодезия» / А. Ю. Будо. – Минск : БНТУ, 2022. – 49 с.

ISBN 978-985-583-761-0.

Учебно-методическое пособие разработано в соответствии с типовой программой по дисциплине «Теория математической обработки геодезических измерений» и предназначено для закрепления теоретических знаний по разделу «Основы теории погрешностей измерений», приобретения исходных практических навыков математической обработки результатов геодезических измерений, применению прикладных компьютерных программ для расчетов и оформления графиков, отчетов, ведомостей.

Издание предназначено для студентов 2-го курса специальности 1-56 02 01 «Геодезия».

УДК 528.1.06:51(075.8)
ББК 26.1я7

ISBN 978-985-583-761-0

Лабораторная работа № 1

КЛАССИЧЕСКАЯ ОБРАБОТКА МНОГОКРАТНЫХ ИЗМЕРЕНИЙ ОДНОЙ ВЕЛИЧИНЫ

Задание: обработать результаты равноточных и неравноточных измерений, выполнить задачу эталонирования.

Исходными данными для работы является превышение h между двумя точками, измеренное N раз при разном количестве штативов n в каждом измерении. Значение вероятности принимается равным $P = 1 - q$, где уровень значимости q = № варианта / 100. При вычислениях удерживается на один десятичный знак больше, чем в исходных данных. Исходные данные выдаются преподавателем по вариантам.

1. Равноточные измерения одной величины

Предполагая, что условия теоремы Гаусса-Маркова и центральной предельной теоремы Ляпунова соблюдены, т. е. измерения независимы, их математическое ожидание и дисперсия постоянны, а закон распределения погрешностей измерений является нормальным, тогда последовательность обработки будет следующей.

Вычисляется среднее арифметическое в качестве оценки математического ожидания измеренных превышений:

$$\bar{h} = \frac{\sum_{i=1}^{N} h_i}{N}. \tag{1.1}$$

Вычисляется оценка стандарта в виде средней квадратической погрешности (СКП) по формуле Бесселя:

$$m = \sqrt{\frac{\sum_{i=1}^{N}\left(h_i - \bar{h}\right)^2}{N-1}}. \tag{1.2}$$

Вычисляется оценка среднего арифметического:

$$M = \frac{m}{\sqrt{N}}. \tag{1.3}$$

Рассмотренный выше способ оценки называется точечным. Более совершенным является так называемый способ доверительных интервалов. При нем определяется значение квантиля t распределения Стьюдента для доверительной вероятности P при количестве степеней свободы $r = N - 1$. Для этого можно воспользоваться статистическими таблицами (прил. А). Например, если дана выборка из 29 элементов (28 степеней свободы), то для вероятности $P = 0{,}95$ и $r = 28$ выбирается значение $t = 1{,}7011$. Также для нахождения квантиля t можно воспользоваться специализированными программами, например, в Microsoft Excel 2007 формула для вычисления квантиля при $q = 1 - P = 0{,}05$ и $r = 28$ будет иметь вид

=СТЬЮДРАСПОБР(0.05*2;28) (1.4)

Поскольку для нахождения истинного значения измеряемой величины используются интервальные оценки, то для вычисления нижней границы интервала истинного значения превышения необходимо воспользоваться формулой

$$h_{\min} = \overline{h} - t \cdot M, \tag{1.5}$$

а для верхней границы, соответственно,

$$h_{\min} = \overline{h} + t \cdot M, \tag{1.6}$$

и это свидетельствует о том, что с 90 % уверенностью (двусторонняя критическая область) мы находим истинное значение, лежащее на интервале $\left(h_{\min}, h_{\max}\right)$.

Для оценки теоретического значения стандарта по статистическим таблицам (прил. Б) определяется величина χ^2. Для нижнего интервала χ_1^2 при $P_1 = q / 2$; и для верхнего χ_2^2 при $P_2 = 1 - q / 2$. Для примера, когда $r = 28$ и $q = 0{,}05$, формула в Excel 2007 для вычисления нижнего χ_1^2 имеет вид

=ХИ2ОБР(1–0.05/2;28) (1.7)

и для верхнего χ_2^2

$$=ХИ2ОБР(0.05/2;28) \tag{1.8}$$

Полученные значения – соответственно χ_1^2 = 15,3079 и χ_2^2 = = 44,4608. После этого вычисляют коэффициенты

$$v_1 = \sqrt{\frac{r}{\chi_2^2}} \tag{1.9}$$

и

$$v_2 = \sqrt{\frac{r}{\chi_1^2}}, \tag{1.10}$$

получая значения v_1 = 0,7936 и v_2 = 1,3525. После этого для истинного значения стандарта можно вычислить нижнюю границу

$$m_{\min} = v_1 \cdot m \tag{1.11}$$

и верхнюю

$$m_{\max} = v_2 \cdot m. \tag{1.12}$$

Аналогичным образом рассчитывается доверительный интервал для СКП среднего арифметического. Нижняя граница

$$M_{\min} = v_1 \cdot M \tag{1.13}$$

и верхняя

$$M_{\max} = v_2 \cdot M. \tag{1.14}$$

Для более глубокого понимания процесса обработки ряда равноточных измерений одной и той же величины можно воспользоваться дополнительной литературой [1, с. 97].

2. Неравноточные измерения одной величины

Наилучшей оценкой математического ожидания для неравноточных измерений является среднее взвешенное или общая арифметическая середина, которую можно вычислить по формуле

$$\bar{h}_2 = \frac{[p \cdot h]}{[p]} = \frac{e^T \cdot P \cdot h}{e^T \cdot P \cdot e}, \quad (1.15)$$

где e – вектор-столбец, состоящий из N единиц;

P – диагональная матрица $N \times N$ со значениями весов измерений на диагонали. Вес измерения при нивелировании может быть задан как $p_i = 1/n_i$, где n_i – количество штативов в i-й секции.

СКП единицы веса может быть определена по формуле Бесселя:

$$\mu = \sqrt{\frac{\sum_{i=1}^{N} p_i \left(h_i - \bar{h}_2\right)^2}{N-1}} = \sqrt{\frac{V^T P V}{N-1}}. \quad (1.16)$$

Величина $\bar{h}_2$ определяется с точностью (СКП средневзвешенного)

$$m_{h_2} = \sqrt{\frac{\sum_{i=1}^{N} p_i \left(h_i - \bar{h}_2\right)^2}{(N-1) \cdot \sum_{i=1}^{N} p_i}} = \sqrt{\frac{V^T P V}{(N-1) \cdot \left(e^T \cdot P \cdot e\right)}}, \quad (1.17)$$

где V – вектор, состоящий из поправок, вычисляемых по формуле

$$V_i = h_i - \bar{h}_2. \quad (1.18)$$

Истинные значения h_2 могут быть найдены для нижней границы как

$$h_{2_{\min}} = \bar{h}_2 - t \cdot m_{h_2}, \quad (1.19)$$

а для верхней границы, соответственно,

$$h_{2_{\max}} = \overline{h}_2 + t \cdot m_{h_2}. \quad (1.20)$$

Нижняя граница истинного значения $\mu_{ист}$ (СКП единицы веса) может быть найдена для нижней границы как

$$m_{\mu,\min} = v_1 \cdot \mu, \quad (1.21)$$

а для верхней границы

$$m_{\mu,\max} = v_2 \cdot \mu, \quad (1.22)$$

где значения v_1 и v_2 вычисляются по формулам (1.9), (1.10).

Нижняя граница истинного значения $m_{h_2,\min}$ (СКП средневзвешенного) может быть найдена как

$$m_{h_2,\min} = v_1 \cdot m_{h_2}, \quad (1.23)$$

и верхняя граница

$$m_{h_2,\max} = v_2 \cdot m_{h_2}, \quad (1.24)$$

где значения v_1 и v_2 вычисляются по формулам (1.9), (1.10).

Для более глубокого понимания процесса обработки ряда неравноточных измерений одной величины можно воспользоваться дополнительной литературой [1, с. 127].

3. Задача эталонирования

В случаях, когда необходимо определить точность прибора и есть эталон измеряемой величины (компаратор), производят N измерений эталона и вычисляют истинные погрешности, считая значение эталона равным истинному значению.

$$\Delta_i = h_i - h_{et}, \quad (1.25)$$

где h_{et} – принятое за истинное значение измеряемой величины. В работе в качестве эталонного принять значение, равное медиане[1] исходного ряда превышений.

Для оценки точности прибора используется формула Гаусса:

$$m = \sqrt{\frac{\sum_{i=1}^{N} \Delta_i^2}{N}}. \tag{1.26}$$

Полученная СКП характеризует точность прибора, если в измерениях отсутствовали значимые грубые и систематические погрешности.

Значимость грубых ошибок может быть оценена по правилу трех сигм (правило Райта): все измерения, выходящие за интервал

$$\left(h_{et} - 3 \cdot m\right) \le h_i \le \left(h_{et} + 3 \cdot m\right), \tag{1.27}$$

считаются грубыми и не включаются в дальнейшую обработку.

После удаления измерений с грубыми ошибками из ряда, проверяют наличие значимого систематического влияния.

Систематическое влияние считается значимым при невыполнении неравенства

$$-t \cdot m_{\overline{\Delta}} \le h_i \le t \cdot m_{\overline{\Delta}}, \tag{1.28}$$

где t – квантиль распределения Стьюдента для доверительной вероятности P (для каждого варианта своя) при количестве степеней свободы $r = N - 1$;

$\overline{\Delta}$ – среднее арифметическое из истинных погрешностей;

[1] *Медианой ряда чисел* называется число, стоящее посередине упорядоченного по возрастанию ряда чисел (в случае, если количество чисел нечетное). Если же количество чисел в ряду четно, то медианой ряда является полусумма двух стоящих посередине чисел упорядоченного по возрастанию ряда.

$m_{\bar{\Delta}}$ – оценка среднего арифметического $\bar{\Delta}$, вычисляемая по формуле

$$m_{\bar{\Delta}} = \frac{m}{\sqrt{N}}, \tag{1.29}$$

где m – СКП, вычисленная по формуле (1.26) для ряда без грубых ошибок.

При невыполнении неравенства (1.28) вычисляют новый ряд, свободный от систематического влияния $\Delta_i' = \Delta_i - \bar{\Delta}$, а оценку точности выполняют по формуле Бесселя:

$$m = \sqrt{\frac{\sum_{i=1}^{N} \Delta_i'^2}{N-1}}. \tag{1.30}$$

В работе выполнить расчет по формуле (1.30) даже при выполнении условия (1.28).

По результатам выполнения данной и всех последующих лабораторных работ студент:

1) проходит онлайн-тест для проверки расчетов своего варианта;

2) составляет отчет с кратким описанием выполненной работы.

Лабораторная работа № 2

АЛЬТЕРНАТИВНАЯ ОБРАБОТКА МНОГОКРАТНЫХ ИЗМЕРЕНИЙ ОДНОЙ ВЕЛИЧИНЫ

Задание: выполнить обработку непараметрическими методами (закон распределения неизвестен) многократно измеренного превышения, установить наличие систематического влияния и грубых ошибок в измеренном ряду данных.

Исходные данные выдаются преподавателем по вариантам.

При количестве измерений, меньшем пятидесяти, сложно выявить закон распределения погрешностей, поэтому в таких случаях находят несколько альтернативных оценок, которые сравниваются между собой.

1. Выявление мешающих параметров непараметрическими методами

Наличие *систематического влияния* в измерениях можно определить, построив линию тренда, т. е. аппроксимировав исходные данные функций вида

$$h = a \cdot i + b, \tag{2.1}$$

где i – номер измерения по порядку от 1 до n (переставлять результаты измерений нельзя);

a – показатель систематического влияния;

b – оценка наиболее надежного значения.

Найти неизвестные коэффициенты линии a, b можно, решив систему уравнений, которая в матричном виде может быть представлена как

$$\begin{pmatrix} a \\ b \end{pmatrix} = N^{-1} \cdot B, \tag{2.2}$$

где N и B вычисляются по формулам:

$$N = \begin{pmatrix} \sum\left(i^2\right) & \sum i \\ \sum i & n \end{pmatrix}, \tag{2.3}$$

$$B = \begin{pmatrix} \sum\left(h_i \cdot i\right) \\ \sum\left(h_i\right) \end{pmatrix}, \tag{2.4}$$

где n – число всех измерений.

После вычисления коэффициентов прямой линии необходимо ее построить вместе с графиком функции $h = f(i)$.

Далее необходимо рассчитать погрешность модели по формуле

$$\mu = \sqrt{\frac{\sum_{i=1}^{N}\left(h_i^{\text{calc}} - h_i\right)^2}{n-k}}, \tag{2.5}$$

где $h_i^{\text{calc}} = a \cdot i + b$;

n – количество измерений;

k – количество неизвестных параметров, в нашем случае $k = 2$.

Систематическое влияние на измерения можно определить по правилу «трех сигм» или правилу Райта, согласно которому систематическое влияние считается значимым, если не выполняется условие

$$-3 \cdot m_a \le a \le 3 \cdot m_a, \tag{2.6}$$

где погрешность m_a может быть определена по формуле

$$m_a = \mu\sqrt{\left(N^{-1}\right)_{1,1}}, \tag{2.7}$$

в которой $\left(N^{-1}\right)_{1,1}$ – первый диагональный элемент обратной нормальной матрицы.

При более точном подходе необходимо вычислить t-статистику Стьюдента:

$$t = \frac{a}{m_a}. \tag{2.8}$$

Если модуль вычисленного значения окажется меньше табличного, определенного для вероятности P и числа степеней свободы $r = N - k$, то делается вывод об отсутствии значимого систематического влияния с этой вероятностью. В работе принять P = 95 %.

Для нахождения измерений с *грубыми ошибками* может быть использован **критерий Хэмпэла**, согласно которому грубым считается измерение, лежащее вне интервала

$$\text{AMO}_{\text{low}} \le h \le \text{AMO}_{\text{high}}, \tag{2.9}$$

где АМО – абсолютное медианное отклонение, вычисляемое по формуле

$$\text{AMO} = \text{med}\left(\left|h_i - \text{med}(h)\right|\right), \tag{2.10}$$

где $\text{med}(h)$ – медиана, вычисляемая из вариационного ряда измерений h_i.

Нижняя граница АМО вычисляется по формуле

$$\text{AMO}_{\text{low}} = \text{med}(h) - 5{,}2 \cdot \text{AMO}, \tag{2.11}$$

верхняя граница –

$$\text{AMO}_{\text{high}} = \text{med}(h) + 5{,}2 \cdot \text{AMO}. \tag{2.12}$$

2. Альтернативные оценки результатов измерений

Перед получением альтернативных оценок должны быть найдены:

– среднее арифметическое:

$$\bar{h} = \frac{\sum_{i=1}^{N} h_i}{n}; \tag{2.13}$$

– средняя квадратическая погрешность:

$$m = \sqrt{\frac{\sum_{i=1}^{N}\left(h_i - \bar{h}\right)^2}{n-1}}; \tag{2.14}$$

– средняя абсолютная погрешность:

$$v = \frac{\sum_{i=1}^{N}\left|h_i - \bar{h}\right|}{n-1}. \qquad (2.15)$$

Данные величины являются оценками математического ожидания и стандарта для двух крайних законов распределения (закон Гаусса и закон Лапласа). При этом первая пара оценок весьма чувствительна к отклонению результатов измерений от нормальности и к влиянию мешающих параметров. Вторая пара оценок нечувствительна к этим отклонениям (робастна). Поэтому степень отличия среднего арифметического от медианы может сказать о значимости посторонних влияний. Если отличия не значимы, то используется первая пара оценок, если значима, то вторая.

Другой подход в определенной выше ситуации заключается в вычислении непараметрических оценок, которые по определению свободны от закона распределения. Наиболее распространенные оценки такого рода – это **L-оценки** и **R-оценки**.

В работе предлагается вычислить следующие наиболее часто встречающиеся **L-оценки** (оценки в линейных комбинациях):

1. **Усеченное среднее** (α-усеченное среднее). Для ее нахождения в вариационном ряду необходимо отбросить с левой и правой стороны α % значений, а из оставшихся взять обычное среднее арифметическое;

2. **Винзоризованное среднее** (α-винзоризованное среднее). Для его нахождения необходимо в вариационном ряду α % крайних значений присвоить значения: слева – $\alpha + 1$ значение, а справа – $(n - \alpha - 1)$ значение. Другими словами, необходимо последним $k = (n \cdot \alpha)$ значениям вариационного ряда присвоить значение предыдущего для них элемента, а первым $k = (n \cdot \alpha)$ значениям присвоить значение следующего после них элемента. Из преобразованного ряда берется обычное среднее арифметическое.

Пусть α = 20 % и имеется набор данных (отсортированных по возрастанию):

2, 3, 4, 5, 7, 9, 10, 12, 14, 30.

Расчет 20 % винзоризованного среднего в нашем примере предполагает перед вычислением среднего арифметического замену первых двух и последних двух значений в ряду данных (2, 3 и 14, 30):

$$4, 4, 4, 5, 7, 9, 10, 12, 12, 12.$$

После замены и расчета среднего результат равняется 7,9.

Сравнение полученных оценок с обычным средним арифметическим также может сказать (по определенному выше правилу) какую величину взять в качестве конечной.

В работе для вычисления усеченного и винзоризованного среднего значение α принять равным 10 %, количество крайних значений в вариационном ряду округлять в большую сторону.

Из **R-оценок** (оценки в ранговых критериях) предлагается вычислить следующие:

1. **Оценка Бикела-Ходжеса**. Находится как медиана из ряда, полученного из средних арифметических двух значений из вариационного ряда: первое – последнее, второе – предпоследнее и т. д. В случае нечетного количества чисел значение, для которого нет пары, в расчет не берется:

$$\theta_{B-X} = \text{med}\left(\frac{h_{i:n} + h_{n+1-i:n}}{2}\right). \tag{2.16}$$

2. **Оценка Лемана-Хождеса**. Ее получают как медиану из всех возможных пар средних в ряду измерений. В работе можно использовать упрощенную оценку, когда в комбинациях для формирования средних значений номер первого слагаемого j всегда меньше номера второго слагаемого k:

$$\bar{h} = \theta_{L-X} = \text{med}\left(\frac{h_{(j)} + h_{(k)}}{2}\right). \tag{2.17}$$

Наряду с этими оценками большое распространение в условиях неопределенности и малом количестве измерений получила **адаптивная оценка Хогга**, когда по величине индикатора выбирается та или иная формула вычисления оценки. Для ее получения используется следующий подход:

$$\bar{h} = \begin{cases} S(0,25;n), & k \le 2; \\ C_t(0;n), & 2 < k \le 4; \\ C_t(0,25;n), & 4 < k \le 5,5; \\ C_t(0,5;n), & 5,5 < k, \end{cases} \tag{2.18}$$

где $S(0,25;n)$ – среднее из первых 25 % и последних 25 % значений вариационного ряда;

$C_t(\alpha;n)$ – α-урезанное среднее.

Если α = 0, то получают стандартное среднее арифметическое;

при α = 0,25 из вариационного ряда удаляется 25 % наименьших и 25 % наибольших значений, а из оставшихся берется среднее арифметическое;

при α = 0,5 удаляется по 50 % слева и справа – стандартная медиана

Для оценки коэффициента *k* используется два подхода.

1. В качестве индикатора *k* берется значение оценки не центрированного эксцесса

$$k = E = \frac{\sum_{i=1}^{N}\left(h_i - \bar{h}\right)^4}{n \cdot m^4}. \tag{2.19}$$

2. Значение коэффициента, обозначенного t_N, вычисляют по формуле

$$k = t_N = \frac{a_N(0,05) - b_N(0,05)}{a_N(0,5) - b_N(0,5)}, \tag{2.20}$$

где $a_N(\beta)$, $b_N(\beta)$ – среднее по $(100 \cdot \beta)\,\%$ наибольших и наименьших элементов вариационного ряда соответственно. В случае дробных значений $a_N(\beta)$ и $b_N(\beta)$ округляются в большую сторону.

Вычисления выполнить при коэффициенте *k*, который рассчитан с использованием первого и второго подхода. Сделать выводы.

Лабораторная работа № 3

МЕТОДЫ ВЫЯВЛЕНИЯ ДОПОЛНИТЕЛЬНЫХ МЕШАЮЩИХ ПАРАМЕТРОВ

Задание: выполнить обработку многократно измеренного превышения, выявить наличие *гетероскедастичности*, систематического влияния и *автокорреляции* в измеренном ряду данных.

Исходные данные выдаются преподавателем по вариантам.

Как известно, к **основным мешающим параметрам** относят значимые грубые и систематические погрешности, незнание влияния которых может испортить эффективность используемой оценки. Способы их выявления делят на:

1) параметрические (закон распределения известен или может быть определен);

2) непараметрические (закон распределения определить невозможно).

К **дополнительным мешающим параметрам** можно отнести эффект *гетероскедастичности* (неравноточность результатов измерений) и эффект *автокорреляции* (зависимость элементов в одном ряду между собой).

1. Выявление эффектов гетероскедастичности

Наиболее простым тестом выявления степени неравноточности групп результатов измерений является критерий *ранговой корреляции Спирмена*. Критерий выявляет корреляцию между номером измерения i и поправкой

$$v_i = h_i - \overline{h}, \qquad (3.1)$$

которая при отсутствии неравенства дисперсий измерений должна быть статистически незначимой. Для этого находят ранги исследуемого ряда следующим образом:

1. Присваивают каждому измерению номер i от 1 до N.

2. Вычисляют среднее арифметическое и отклонение от него v_i для всех элементов ряда.

3. Выстраивают ряд отклонений в вариационный ряд (по возрастанию).

4. Получают ранг n_i отклонения v_i, который равен номеру i каждого элемента исходного ряда отклонений в вариационном ряду. Т. е. если элемент в исходном ряду измерений имел порядковый № 6, а в построенном вариационном ряду данный элемент получил № 15, то принимаются следующие значения: $i = 6$; $n_i = 15$.

5. И так далее для каждого элемента ряда.

Коэффициент ранговой корреляции Спирмена вычисляется по формуле:

$$r_{i,v} = 1 - 6 \cdot \frac{\sum_{i=1}^{N}\left(\left(n_i - i\right)^2\right)}{N \cdot \left(N^2 - 1\right)}. \tag{3.2}$$

Полученный коэффициент корреляции исследуется на значимость при помощи квантиля t распределения Стьюдента, который сравнивается с рассчитанным значением

$$t = \frac{\left|r_{i,v}\right| \cdot \sqrt{N-2}}{\sqrt{1 - r_{i,v}^2}}. \tag{3.3}$$

Полученное значение сравнивается с эталонным, выбираемым из статистических таблиц по модифицированной вероятности для двухстороннего интервала $(1 + P) / 2$ и числу степеней свободы $N - 2$. При выполнении неравенства $t > t_{\text{эт}}$ исходная гипотеза об отсутствии гетероскедастичности отвергается с вероятностью ошибки $(1 - P) / 2$. При вычислениях принять значение $P = 0{,}95$. Тогда, например, при числе измерений $N = 20$ число степеней свободы $N - 2 = 18$, и значение квантиля для модифицированной вероятности $(1 + P) / 2$ будет равно $t_{\text{эт}} = 2{,}10$.

Второй распространенный тест на наличие гетероскедастичности в результатах измерений называется *тестом Голдфелда-Квандта*. Его суть: если в вариационном ряду группа первых результатов и группа последних имеет достаточно похожую меру рассеивания, то эффект неравноточности результатов измерений незначителен. Для практической реализации теста поступают следующим образом:

1) делят вариационный ряд на три примерно равных части (меньшая в середине, равные по краям). Два крайних подряда аппроксимируют по методу наименьших квадратов в зависимости от номера i (задача выявления тренда);

2) вычисляют суммы квадратов отклонений для первой $\left[v^2\right]_1$ и второй $\left[v^2\right]_2$ регрессий, а также F – статистику Фишера – по формуле

$$F = \frac{\left[v^2\right]_2}{\left[v^2\right]_1}. \tag{3.4}$$

При этом в числителе должна быть большая величина. Критическое значение $F_{кр}$ выбираются из таблиц распределения Фишера по уровню значимости $\alpha = 1 - P$ и числу степеней свободы

$$r_1 = r_2 = k - t - 1,$$

где k – число элементов в крайнем ряду,

t – число неизвестных параметров (объясняющих переменные) в принятой модели регрессии, т. е. $t = 2$.

Если выполняется неравенство $F < F_{кр}$, то с вероятностью P гипотеза об отсутствии гетероскедастичности принимается.

При вычислениях принять значение P = 0,95. Тогда, например, при числе степеней свободы $r_1 = r_2 = 5$ значение квантиля F-распределения вероятности будет равно $F_{0,95;5;5}$ = 5,05. Для нахождения квантиля можно воспользоваться статистическими таблицами (прил. В) либо функцией Excel:

$$\text{=FРАСПОБР(0.05;5;5)} \tag{3.5}$$

2. Выявление систематического влияния непараметрическими способами

Наиболее часто используемые критерии выявления систематических влияний в результатах измерений при условии, что закон распределения неизвестен и количество измерений невелико, – это *критерий серий* и *критерий «восходящих» и «нисходящих» серий*.

Критерий серий относят к критериям, выявляющим значимость систематического влияния только монотонного характера, (сдвиг или тренд) на основе проверки вероятностной независимости среди элементов исследуемого ряда. Для этого производят вычисления следующим образом:

1. Строят вариационный ряд и находят медиану med(h).

2. Формируют знаковый ряд из плюсов и минусов по правилу: если значение исходного (невариационного) ряда больше медианы, то вместо i-го числа записывают знак «+», если меньше, то знак «–». Элементы ряда, равные med(h), пропускают.

3. Находят количество серий (последовательностей подряд идущих знаков) $v(N)$.

4. Находят число элементов в наибольшей серии $\tau(N)$.

Для стохастической независимости и, следовательно, отсутствия значимого систематического влияния монотонного характера должны одновременно выполняться два неравенства:

$$v(N) > 0{,}5 \cdot \left(N + 1 - 1{,}96 \cdot \sqrt{N - 1}\right), \tag{3.6}$$

$$\tau(N) > 3{,}3 \cdot (N + 1). \tag{3.7}$$

Если хотя бы одно из неравенств не выполняется, то гипотеза об отсутствии систематического влияния в исходных измерениях отвергается с вероятностью ошибки, заключенной в пределах от 0,05 до 0,0975, т. е. с доверительной вероятностью 0,9025–0,95.

В отличие от *критерия серий*, рассматриваемый далее *критерий «восходящих» и «нисходящих» серий* выявляет смещение среднего значения не только монотонного характера (тренд или сдвиг), но и более общего, например, периодического характера. В нем также исследуется последовательность знаков, но закон ее построения, следующий: на месте значения h_i исходного ряда, ставится «+», если

$$h_{i+1} - h_i > 0, \tag{3.8}$$

и, соответственно, знак «–» при выполнении неравенства

$$h_{i+1} - h_i < 0. \tag{3.9}$$

Если несколько последовательных измерений равны, то используется только одно из них. Гипотеза об отсутствии систематического влияния принимается в случае выполнения неравенств:

$$v(N) > \frac{1}{2} \cdot (2 \cdot N - 1) - 1{,}96 \cdot \sqrt{\frac{16 \cdot N - 29}{90}}, \tag{3.10}$$

$$\tau(N) > \tau_0(N), \tag{3.11}$$

где $\tau_0(N) = 5$ при $N \leq 26$;

$\tau_0(N) = 6$ при $26 < N \leq 153$;

$\tau_0(N) = 7$ при $153 < N \leq 1170$.

В случае невыполнения одного из неравенств (3.10) и (3.11) гипотеза об отсутствии систематического влияния отвергается с уровнем значимости (вероятностью ошибки первого рода) от 0,05 до 0,0975.

3. Выявление эффектов автокорреляции

Под автокорреляцией ряда принято понимать тесноту связи между элементами одного ряда. Чтобы упорядочить эти связи, используют понятие лага – величины сдвига между исследуемыми элементами. Наиболее часто встречается автокорреляция лага (сдвига) 1, т. е. между рядом стоящими элементами в ряду: 1 и 2, 2 и 3, 3 и 4 и т. д. Самым известным и используемым тестом на исследование такого рода зависимости является критерий Дарбина-Уотсона, когда по статистике DW, вычисленной по величинам остатков V после аппроксимации ряда линейной функцией, делается вывод о виде и значимости автокорреляции. Эта статистика тесно связана с выборочным коэффициентом корреляции между рядом стоящими остатками V_{i-1} и V_i.

$$r = 1 - \frac{\mathrm{DW}}{2}, \tag{3.12}$$

$$\mathrm{DW} = \frac{\sum_{i=2}^{N}\left(V_i - V_{i-1}\right)^2}{\sum_{i=1}^{N} V_i^2}. \tag{3.13}$$

Тогда из (3.12) и (3.13) имеем:

– если DW ≈ 2, то $r \approx 0$ (отсутствие автокорреляции);

– если DW ≈ 0 , то $r \approx 1$ (положительная автокорреляции);

– если DW ≈ 4 , то $r \approx -1$ (отрицательная автокорреляции);

Общая схема критерия Дарбина-Уотсона, следующая:

1. Строят эмпирическое уравнение регрессии, например h_i от i, и находят остатки:

$$V_i = h_i - \hat{h}, \tag{3.14}$$

например, как в тесте на выявление мешающих параметров непараметрическими методами (см. ЛР2).

2. Рассчитывают по формуле (3.13) статистику DW и при приближенном оценивании по изложенному выше правилу смотрят, к какому числу из 0, 2 или 4 находится ближе вычисленное значение статистики.

Исходя из расчетов делают приближенный вывод о возможности того или иного исхода. Можно считать, что если 1,5 < DW < 2,5, то автокорреляция отсутствует, при –0,5 < DW < 0,5 имеем положительную автокорреляцию, т. е. остатки все время возрастают, а для 3,5 < DW < 4,5 – отрицательную автокорреляцию (остатки все время убывают). Результаты тем надежней, чем ближе статистика к ключевым точкам.

Расчетно-графическая работа

ИССЛЕДОВАНИЕ РЯДА ПОГРЕШНОСТЕЙ НА СООТВЕТСТВИЕ НОРМАЛЬНОМУ ЗАКОНУ РАСПРЕДЕЛЕНИЯ

Задание: освоить различные методики исследования погрешностей результатов измерений на соответствие нормальному закону распределения с предварительным анализом на систематические и грубые ошибки; получить основные вероятностно-статистические характеристики многократно измеренной величины.

Порядок выполнения работы:

1. Для ряда данных, соответствующего варианту, выданному преподавателем, выполнить оценку центра распределения и исключить из ряда промахи.

<u>Предварительные вычисления.</u>

2. Исследовать полученный ряд на наличие систематической составляющей и при ее наличии исключить из элементов, получив новые, преобразованные величины.

3. Исследовать выборку на наличие в ней грубых погрешностей и, если они обнаружатся, исключить их из ряда.

4. Исследовать по соответствующим формулам F-критерия Фишера ряд на степень неоднородности.

В пунктах 2–4 принять доверительную вероятность $\beta = 0{,}95$.

<u>Приближенные исследования.</u>

5. Вычислить оценки основных параметров нормального распределения.

6. Выполнить приближенное исследование ряда погрешностей на соответствие нормальному закону, вычислив среднюю абсолютную и вероятную ошибки и их отношение к средней квадратической ошибке ряда и дополнительные характеристики статистической совокупности в виде асимметрии и эксцесса, установить их значимость.

7. Сделать выводы о соответствии закону Гаусса по приближенным критериям.

<u>Графические исследования.</u>

8. Рассчитать размер интервалов, необходимых для построения гистограммы. Интервалы должны накрывать все исследуемые значения.

9. Сосчитать количество элементов, попадающих в каждый интервал. Откладывая на горизонтальной оси величины интервалов, а на вертикальной – высоты прямоугольников, построить гистограмму.

10. Построить на гистограмме кривую распределения нормального закона по вычисленным в п. 5 основным характеристикам.

11. Визуально оценить степень приближения эмпирического распределения, представленного в виде гистограммы, к теоретическому распределению, представленному в виде непрерывной кривой.

Вероятностные исследования.

12. Оценить, верна ли с вероятностью β принятая гипотеза о законе распределения погрешностей по χ^2 – критерию Пирсона.

13. Произвести вычисление вероятности точного выполнения гипотезы.

14. Сделать общий вывод о соответствии исследуемого ряда погрешностей нормальному закону распределения. Заключительный анализ результатов исследований выполняется в виде пояснительной записки, включающей текстовый материал, со ссылками на формулы и источники литературы.

Теоретические основы выполнения исследований

Обработка результатов измерений имеет место всегда, когда одна из определяемых величин получена несколько раз с отличными друг от друга значениями. При этом корректная оценка результатов возможна только при известных правилах, определяющих поведение погрешностей измерений Δ. К таким правилам относят законы поведения погрешностей в дифференциальной $f(\Delta)$ и интегральной $F(\Delta)$ формах, их основные численные характеристики и представления законов в графическом виде. Интегральная форма называется *функцией распределения погрешностей* Δ_i, дифференциальная форма – *функцией плотности распределения погрешностей* Δ_i. К основным характеристикам законов относят наиболее вероятное значение определяемой величины, называемое *математическим ожиданием* и обозначаемое МО(Δ) или $M(\Delta)$, μ; меру рассеивания измерений вокруг математического ожидания, называемую *дисперсией* и обозначаемую $\sigma^2(\Delta)$ или $D(\Delta)$ (чаще используют просто величину $\sigma(\Delta)$, называемую *стандартом,* так как он не имеет квадратичной размерности, как у дисперсии). К дополнительным характеристикам законов относят меру скошенности относительно вертикальной оси симметрии, называемую *асимметрией* и обозначаемую A или S_k, и меру крутости, называемую *эксцессом* и обозначаемую E.

Множество теоретических и практических исследований показывают, что результаты геодезических измерений подчиняются *нормальному закону распределения (закону Гаусса)* и имеют вид

$$F(X) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot \int_{-\infty}^{x} e^{\frac{-(t-\mu)^2}{2\cdot\sigma^2}} \mathrm{d}t, \qquad (4.1)$$

с основными характеристиками, представленными выше. Следует твердо усвоить, что математическое ожидание и стандарт являются теоретическими, заведомо неизвестными значениями. Получить эти значения можно только при числе измерений $n = \infty$. Поэтому в качестве приближений к этим параметрам используются их оценки.

Например, оценка математического ожидания μ в виде *среднего весового* (общей *арифметической середины*) $X_{\text{ср}}$. Или оценка стандарта в виде *средней квадратической погрешности m*. Эти оценки вычисляются достаточно просто и имеют минимальное отклонение от своих теоретических значений, если поведение погрешностей подчинено нормальному закону Гаусса. Необходимо помнить, что теоретическое значение математического ожидания из ряда случайных погрешностей для закона Гаусса МО = 0. В упрощенном виде под законами распределения случайной величины будем понимать совокупность частот попадания ее реализации в заданные интервалы.

Таким образом, для правильного выбора алгоритма обработки результатов измерений необходимо установить, что с достаточной вероятностью β ошибки подчинены тому или иному закону распределения. Для геодезических измерений часто требуется исследование на соответствие погрешностей нормальному распределению Гаусса. Тогда в качестве наиболее надежных оценок результатов измерений можно использовать среднее взвешенное (или среднее арифметическое) и среднюю квадратическую погрешность.

Из теории погрешностей известно, что источниками элементарных погрешностей измерений является множество факторов разного характера, учесть которые практически невозможно. Поэтому суммарные ошибки, вызванные этими факторами, носят случайный характер и подчиняются законам теории вероятностей. Практическую реализацию формул теории вероятностей при конечном числе наблюдений рассматривает математическая статистика.

Одна из основных теорем теории вероятностей, называемая *центральной предельной теоремой Ляпунова,* гласит, что если элементарные составляющие общей ошибки измерения примерно одного порядка малы по величине, независимы между собой и не имеют систематической составляющей, то общая ошибка распределена по закону, очень близкому к нормальному. Вместе с тем, в процессе измерений часто имеет место наличие *грубых погрешностей* (погрешностей, которые больше заданного допуска) или промахов (т. е. очень сильно отличающихся от других). Кроме этого, в результатах наблюдений могут содержаться и какие-либо постоянные составляющие, называемые *систематическими ошибками.* Не редки случаи включения в обработку измерений с разной точностью (неоднородных измерений).

Методики производства геодезических измерений рассчитываются так, чтобы по возможности полнее исключить систематические влияния на результаты, которые к тому же должны быть свободны от грубых погрешностей и однородны. Наличие как грубых систематических погрешностей, так и неоднородности, нарушает условия теоремы Ляпунова и делает проблематичным использование для обработки измерений алгоритмов, полученных на основе закона распределения Гаусса.

Вычислительные схемы, основанные на нормальном законе (например, разные формы *метода наименьших квадратов* (*МНК*)), имеют самый простой вид и дают вполне надежные результаты при условии выполнения приведенных выше условий. Из всего сказанного совершенно очевидно, что для использования оценок, соответствующих закону Гаусса, необходимо выполнить исследование на степень соответствия ряда погрешностей нормальному закону. Но для достоверности результатов исследований необходимо провести предварительные вычисления и выявить наличие грубых погрешностей, значимость систематической составляющей и неоднородности. После этого закон распределения устанавливается на основе имеющейся статистической совокупности погрешностей. При этом результаты будут тем более надежны, чем больше используется в исследованиях величин погрешностей.

Основными задачами, решаемыми вероятностно-статистическими методами в контексте исследования закона распределения случайных величин для общего случая, являются:

1. Определение вида закона распределения ряда погрешностей – задача *сглаживания или выравнивания статистического ряда.*

2. Определение степени согласования опытных данных с гипотезой о том, что случайная величина подчинена полученному в п. 1 закону распределения с вероятностью β. Для этого используются различные «критерии согласия» – задача *проверки правдоподобия гипотез*.

На основе перечисленных пунктов производится определение «наилучших» оценок неизвестных параметров распределения в виде оценок математического ожидания МО(Δ) как центра распределения и дисперсии $D(\Delta)$ как меры рассеивания – задача *наилучшего оценивания.*

Эти же задачи характерны в целом и для всей математической статистики. Обычно используют три группы алгоритмов исследования на закон распределения случайных величин:

1. Приближенные – вычисляется ряд характеристик распределения и сравнивается с теоретическими значениями.

2. Графические – по результатам измерений строится какое-либо графическое представление закона в виде многоугольника распределения, полигона распределения, гистограммы и др.

3. Вероятностные – выдвигается гипотеза о соответствии ряда погрешностей закону распределения с вероятностью β и проверяется ее правильность, или вычисляется величина-критерий, по которой выносят суждение о вероятности соответствия результатов измерений выдвинутой гипотезе.

Очевидно, что перед началом работы по определению соответствия ряда нормальному закону, следует убедиться, что он достаточно случаен, т. е. достаточно неплохо выполняются условия теоремы Ляпунова. Следует иметь в виду, что невыполнение этих условий полностью обесценит результаты исследований. Для этого производятся предварительные вычисления.

Оценки центра распределения

Координата центра распределения определяет положение случайной величины на числовой оси. Дать однозначное определение этого понятия невозможно. Центр распределения может быть найден несколькими способами:

– как медиана распределения;

– как мода распределения;

– как математическое ожидание.

По возможности наиболее точная оценка центра распределения по выборке случайных величин исключительно важна, так как центр распределения используется в формулах для вычисления дисперсии, среднеквадратичного отклонения, коэффициента асимметрии и эксцесса распределения. Некорректное определение центра влечет за собой ошибки в определении всех этих величин.

Оценку центра распределения по выборке можно проводить различными способами. Не зная априорно закона распределения случайной величины, невозможно заранее указать наиболее приемлемый способ. К тому же, некоторые из этих оценок чувствительны к наличию аномальных значений в выборке (промахов).

Для нахождения центра распределения можно использовать альтернативные L, R-оценки, адаптивную оценку Хогга и др.

Но в данной работе для корректной оценки центра распределения $X_{\text{центр}}$ мы будем вычислять его пятью различными способами:

$X_{\text{медиана}}$, $X_{\text{центр\_сгибов}}$, $\bar{X}$, $\bar{X}_{50\,\%}$, $X_{\text{центр\_размаха}}$.

После этого пять полученных оценок упорядочим по возрастанию и выберем из них в качестве центра распределения серединное, то есть третье по счету, значение.

1) $X_{\text{медиана}}$ – обычная медиана;

2) $\bar{X}$ – среднее арифметическое;

3) $X_{\text{центр\_размаха}}$ – среднее между максимальным и минимальным значением в выборке;

4) $X_{\textit{центр\_сгибов}}$ – центр 50 %-го интерквантильного промежутка (нечувствительна к промахам в выборке). Перед вычислением упорядочивают выборку по возрастанию. Вычисляют четвертую часть от объема выборки, то есть

$$M=ЦЕЛОЕ(N/4). \qquad (4.2)$$

Тогда центр сгибов определяется по формуле:

$$X_{\text{центр\_сгибов}} = \frac{X_{M+1} + X_{N-M}}{2}; \qquad (4.3)$$

5) $\bar{X}_{50\,\%}$ – среднее арифметическое по 50 %-му интерквантильному промежутку вычисляется по формуле

$$\bar{X}_{50\,\%} = \frac{1}{N-2M} \cdot \sum_{k=M+1}^{N-M} x_k, \qquad (4.4)$$

где M вычисляется по формуле (4.2).

Исключение промахов из выборки

Промахами в выборке случайных величин будем называть аномально отклоняющиеся от центра распределения значения по сравнению с основной массой данных. Исключать промахи из выборки необходимо, т. к. они могут существенно исказить оценку параметров распределения. Для исключения промахов введем понятие коэффициента цензурирования – безразмерной величины G, при которой промахами считаются все значения из выборки, лежащие за пределами интервала

$$X_{\text{центр}} - G \cdot \bar{\sigma} \le x \le X_{\text{центр}} + G \cdot \bar{\sigma}, \qquad (4.5)$$

где

$$\bar{\sigma} = \sqrt{\frac{1}{N-1} \cdot \sum_{i=1}^{N} \left(x_i - X_{\text{центр}}\right)^2}. \qquad (4.6)$$

Интуитивно понятно, что коэффициент цензурирования должен зависеть от объема выборки и рассчитанного по выборке значения эксцесса. Действительно, такое отклонение от центра, которое является промахом для средневершинного (а тем более плосковершинного) распределения, для островершинного распределения с его длинными «тяжелыми» спадами может безусловно принадлежать выборке.

Эмпирическая формула для коэффициента цензурирования как функции от объема выборки N и эксцесса E, пригодная к применению для широкого класса распределений, следующая:

$$G = 1{,}55 + 0{,}8 \cdot \lg\left(N/10\right) \cdot \sqrt{E-1}, \tag{4.7}$$

где

$$E = \frac{1}{\overline{\sigma}^4} \cdot \frac{N^2 - 2N + 3}{\left(N-1\right)\left(N-2\right)\left(N-3\right)} \sum\nolimits_{i=1}^{N} \left(x_i - X_{\text{центр}}\right)^4 - \\ - \frac{3\left(2N-3\right)\left(N-1\right)}{N\left(N-2\right)\left(N-3\right)}. \tag{4.8}$$

После удаления промахов нужно пересчитать параметры распределения. При этом в качестве центра распределения уже можно использовать среднее арифметическое как состоятельную и несмещенную оценку математического ожидания. Повторно выполнять поиск промахов с использованием неравенства (4.5) не требуется.

Предварительные вычисления для исследования

В предварительных вычислениях ряд исследуется на наличие значимых систематических и грубых погрешностей, а также меры однородности результатов по точности на основе различных критериев.

Определение значимости систематического влияния. Следует иметь в виду, что систематические влияния в рядах присутствуют всегда, но они могут быть значимы и не значимы. При определении наличия значимых систематических погрешностей в ряду имеют место два случая:

1) известно истинное значение определяемой величины $X_{\text{ист}}$ и произведено N ее измерений x_i. В этом случае пользуются зависимостью

$$\frac{\left|\Delta_i\right|}{M_{\overline{X}}} \le z_q, \tag{4.9}$$

где

$$\Delta_i = x_i - X_{\text{ист}}, \qquad (4.10)$$

средняя квадратическая погрешность среднего арифметического

$$M_{\bar{X}} = \frac{m}{\sqrt{N}}, \qquad (4.11)$$

где m – средняя квадратическая погрешность (СКП) исследуемой величины,

N – число элементов в ряду.

z_q – квантиль t-распределения Стьюдента определяется по *уровню значимости* q (или вероятности β) и числу *избыточных измерений* (*числу степеней свободы*) $k = N - 1$ и выбирается из статистических таблиц (прил. А) или вычисляется, например, в Excel. Если неравенство (4.9) выполняется, то с вероятностью $\beta = 1 - q$ считаем, что значимые систематические погрешности в ряду измерений отсутствуют;

2) истинное значение величины неизвестно. Тогда наличие в результатах наблюдений постоянной составляющей может быть выяснено по наиболее распространенному в геодезии *критерию Аббе*. Для этого выдвигается гипотеза, что с вероятностью β в предложенном ряду отсутствует значимое систематическое влияние. По исследуемым величинам вычисляется практическая величина

$$\delta = \frac{\tilde{m}^2}{m^2}, \qquad (4.12)$$

которая является отношением двух оценок дисперсий, средние квадратические ошибки которых получены как

$$m = \sqrt{\frac{\sum_{i=1}^{N} v_i^2}{N-1}} \qquad (4.13)$$

и

$$\tilde{m} = \sqrt{\frac{\sum_{i=1}^{N} d_i^2}{2 \cdot (N-1)}}, \qquad (4.14)$$

где уклонение i-й величины от среднего

$$v_i = x_i - \overline{X} \qquad (4.15)$$

последовательные разности

$$d_i = x_{i+1} - x_i. \qquad (4.16)$$

В (4.14) суммирование производится по (N – 1) элементам. Для сравнения, по заданной вероятности β (или уровню значимости q), числу степеней свободы r и с использованием статистических таблиц *критерия Аббе* получают контрольную величину δ_q. Тогда, при $\delta > \delta_q$ принимается гипотеза об отсутствии систематической ошибки с вероятностью $\beta = 1 - q$. В противном случае ($\delta \le \delta_q$) следует принять гипотезу о постоянной составляющей в статистической совокупности и для корректной оценки исследуемых параметров ее необходимо исключить из ряда измерений. Для этого получают усредненную величину систематического влияния, равную среднему арифметическому из всех элементов. Данную величину исключают из измерений, получая новый ряд $X_{\text{исп}}$ с уменьшенной по сравнению с исходным рядом систематической составляющей.

$$x_{\text{ист},i} = x_i - \overline{X}. \qquad (4.17)$$

При $N > 60$ величина δ_q может быть вычислена по формуле

$$\delta_q = 1 + \frac{n_q}{\sqrt{N + 0{,}5 \cdot \left(1 + n_q^2\right)}}, \qquad (4.18)$$

где n_q – квантиль нормированного нормального распределения. Для его нахождения в Microsoft Excel 2007 используется формула

= НОРМСТОБР(q) (4.19)

Например, при $q = 0{,}05$ будет получено значение $n_q = -1{,}64485$.

Анализ формул (4.12)–(4.14) говорит о том, что постоянная часть будет значима в измерениях, когда величины уклонений от среднего минимальны, а последовательные разности максимальны.

Исследование на наличие грубых погрешностей. При отсутствии систематической составляющей или после их исключения по (4.17) проводят исследование на наличие в ряду грубых погрешностей. В зависимости от требований задачи существует масса критериев, решающих поставленную задачу: критерий Граббса, Диксона, Шарлье, Шовенэ и др. В работе для выявления грубых погрешностей предлагается использовать критерий Граббса. Критерий дает вероятность выполнения выдвинутой гипотезы о том, что максимальное или минимальное значение из ряда не являются грубыми погрешностями. Для этого по экстремальным значениям выборки $X_{\max}$ и $X_{\min}$, среднему арифметическому и средней квадратической погрешности, вычисляют значения

$$\left.\begin{aligned} z_{\max} &= \frac{X_{\max} - \bar{X}}{m} \\ z_{\min} &= \frac{\bar{X} - X_{\min}}{m} \end{aligned}\right\}. \tag{4.20}$$

Если $z_{выч} \le z_q$ для максимального и минимального значения, то следует принять гипотезу об отсутствии в ряду грубых погрешностей, так как экстремальные значения не являются грубыми. Значения теоретической величины критерия z_q получают по заданному аргументу q и числу элементов в выборке N по специальным статистическим таблицам критерия Граббса для z_q, [8, табл. IX]. Если же $z_{выч} > z_q$, тогда крайнее значение вариационного ряда из дальнейшей обработки следует исключить, а к новым крайним значениям еще раз применить критерий *Граббса*. В рамках данной работы при необходимости удаляется одно или оба крайних значения; повторное применение критерия Граббса студентом не выполняется.

Для $N > 25$ значения z_q можно с приемлемой точностью рассчитать по уравнениям, указанным в таблице.

Таблица 4.1

Вычисление критических значений z_q

q	z_q **при** $N > 25$
0,10	$0{,}3053 \cdot \mathrm{Ln}(N) + 1{,}6513$
0,05	$0{,}2849 \cdot \mathrm{Ln}(N) + 1{,}9517$
0,01	$0{,}2648 \cdot \mathrm{Ln}(N) + 2{,}4839$

Выявление неоднородности в исследуемом ряду. Для выявления степени *неоднородности* (неравноточности) результатов измерений используется самый простой *F-критерий Фишера*. Для этого разбивают исследуемую выборку на две примерно одинаковые подвыборки и для каждой вычисляются средние квадратические погрешности m_1 (для первой половины выборки) и m_2 (для второй половины выборки). В данной работе в случае нечетного количества элементов в исходной выборке размер первой подвыборки делаем меньше размера второй. Тогда, если вычисленное значение статистики Фишера $F_{\text{выч}} < F_{\text{крит}}$ (критического значения), то принимается гипотеза с вероятностью β об отсутствии неравноточности (разнородности) в ряду исследуемых величин. Здесь величина $F_{\text{выч}} = m^2{}_1 \,/\, m^2{}_2$, причем в числителе должна быть большая дисперсия. Значение $F_{\text{крит}}$ выбирают из таблиц распределения Фишера по $r_1 = N_1 - 1 - 1$ (для числителя) и $r_2 = N_2 - 1 - 1$ (для знаменателя) степеням свободы и вероятности $(1 + \beta) / 2$ (прил. В). Также расчет можно произвести в Excel.

Таким образом, ряд исследуемых случайных величин будет подготовлен в соответствии с *центральной предельной теоремой Ляпунова* для корректного исследования на соответствие нормальному закону распределения.

Вычисление основных характеристик ряда

Как было показано выше, к основным характеристикам закона распределения относят *математическое ожидание* $M(X)$ (так называемый *начальный момент первого порядка*) и стандарт σ (корень из *центрального момента второго порядка).* Эти величины теоретические и могут быть точно определены только при числе элементов в ряду $N = \infty$. В практической деятельности используют один из видов их оценок, вычисленных при конечном N. Для нормального закона это, соответственно, среднее арифметическое $X_{ср}$ и *средняя квадратическая погрешность* m.

Оценки математического ожидания, дисперсии и стандарта получим по следующим формулам:

$$M(X) = \frac{1}{N} \cdot \sum_{i=1}^{N} x_i, \tag{4.21}$$

$$D(X) = \frac{1}{N-1} \cdot \sum_{i=1}^{N} \left(x_i - \overline{X}\right)^2, \tag{4.22}$$

$$\sigma = \sqrt{D(X)} \approx m. \tag{4.23}$$

Если в качестве исследуемых величин берутся истинные ошибки (например, невязки), то сумму в формулах (4.21)–(4.22) для усреднения делят на количество элементов N (форма Гаусса); если в качестве погрешностей используют уклонения от оценки (например, от среднего арифметического), то в качестве делителя берут число степеней свободы $N - 1$ (формула Бесселя).

Приближенные критерии исследования

Приближенные критерии исследования ряда погрешностей на соответствие нормальному закону распределения используют сравнение некоторых известных теоретических характеристик нормального закона и их вычисленных по результатам измерений аналогов. Степень отличия теоретических величин от вычисленных оценок говорит о степени приближения исследуемого ряда к нормальному закону.

Кроме наиболее распространенной средней квадратической погрешности m используют *средние абсолютные* v и *вероятные* (*срединные*) ошибки r, которые являются оценками теоретических *абсолютных центральных моментов первого порядка* и 0,5 (или 2) – *квантили закона распределения Гаусса* соответственно. Между тремя ошибками m, v и r для нормального закона распределения величин имеются теоретические строгие соотношения

$$\begin{cases} k_1 = \dfrac{m}{v} = 1{,}25; \\ k_2 = \dfrac{m}{r} = 1{,}48; \\ k_3 = \dfrac{v}{r} = 1{,}18. \end{cases} \tag{4.24}$$

По величине отклонения вычисленных значений от их теоретических аналогов можно судить о степени приближения ряда к закону Гаусса, а также возможно использовать (4.24) для вычисления любых других двух погрешностей, если известна одна из трех. В качестве меры значимости отличия вычисленной величины от ее теоретического значения можно использовать, например, «критерий ничтожных влияний», гласящий, что величина считается неизменной, если ее вариация составляет не более 11 % от самой величины. Например, для первой формулы из (4.24) имеем

$$\left| \frac{m}{v} - 1{,}25 \right| \le 0{,}11 \cdot k_1. \tag{4.25}$$

Тогда отклонения коэффициентов считаются допустимыми, если их абсолютная разность не более 0,138, 0,163 и 0,130 соответственно.

Для вычисления *средней абсолютной ошибки* пользуются формулами

$$v = \frac{1}{N} \sum_{i=1}^{N} |x_i| \tag{4.26}$$

и

$$v = \frac{1}{N-1} \sum_{i=1}^{N} |x_i - \bar{X}|. \tag{4.27}$$

При этом формула (4.26) используется, если ошибки x_i – истинные и оценка математического ожидания равна нулю.

Чтобы определить *вероятную ошибку*, величины располагают в так называемый абсолютный вариационный ряд (по возрастанию их абсолютных значений). В данной работе каждый элемент ряда рассчитывается по формуле

$$z_i = \frac{x_i - M(X)}{\sigma}. \quad (4.28)$$

Если в ряду нечетное число элементов, то искомая ошибка равна значению величины, находящейся точно посередине (величина с номером $(N + 1) / 2$ в построенном вариационном ряду). Если число исследуемых величин четное, то значение ошибки находится как среднее арифметическое из двух чисел, стоящих в середине ряда (среднее между элементами с номерами $N / 2$ и $(N + 1) / 2$). Другими словами, находится медиана ряда

$$r = \mathrm{med}(z). \quad (4.29)$$

К дополнительным характеристикам распределения случайных величин, составляющих статистическую совокупность, называемым характеристиками формы, относят *эксцесс* E (меру «крутости») и *асимметрию* A (меру «скошенности»), которые получаются с использованием моментов более высокого порядка. При этом следует помнить, что для нормального закона распределения теоретические значения *асимметрии* и *эксцесса* равны нулю. Тогда значения эксцесса и асимметрии можно считать несущественными при условиях

$$|A| \le t \cdot m_A \quad (4.30)$$

и

$$|E| \le t \cdot m_E, \quad (4.31)$$

где

$$A = \frac{1}{N \cdot m^3} \cdot \sum_{i=1}^{N} \left(x_i - \overline{X}\right)^3, \quad (4.32)$$

$$E = \frac{1}{N \cdot m^4} \cdot \sum_{i=1}^{N} \left(x_i - \overline{X}\right)^4 - 3. \quad (4.33)$$

Значение вероятностного коэффициента t в (4.30), (4.31) выбирают в зависимости от вероятности, изменяя ее как $(1 + \beta) / 2$, так как интервал двухсторонний, и используя таблицы нормального распределения. Для наиболее часто используемой вероятности 0,95 коэффициент для двухстороннего интервала будет равен 1,96.

Эмпирические значения средних квадратических погрешностей дополнительных характеристик распределения могут быть вычислены по сокращенным формулам:

$$m_A = \sqrt{\frac{6}{N}} \tag{4.34}$$

и

$$m_E = \sqrt{\frac{24}{N}}. \tag{4.35}$$

Следует иметь в виду, что если при вычислении асимметрии A ее значение окажется больше нуля, то кривая эмпирического распределения будет скошена влево, а когда $A < 0$ – то вправо. При величине эксцесса $E > 0$ эмпирическое распределение «высоковершинное» (его вершина выше вершины теоретической кривой нормального распределения). В противном случае $(E < 0)$ распределение «низковершинное».

Вычисления оценок параметров нормального распределения исследуемого ряда величин (среднего арифметического и средней квадратической погрешности), а также ее дополнительных характеристик (коэффициентов k, асимметрии и эксцесса) позволяют сделать только предварительное заключение о соответствии эмпирического распределения теоретическому нормальному закону Гаусса и только по близости соответствующих теоретических характеристик распределения их вычисленным аналогам.

Графический критерий исследования

Для дальнейших исследований погрешностей на соответствие их нормальному закону распределения строят для ряда одно из его графических представлений, например, в виде гистограммы или многоугольника распределения, с нанесенной поверх нее теоретической кривой закона Гаусса с параметрами $\bar{X}$ и m, называемой *огивой*.

В данной работе предлагается использовать *гистограмму*. Построение гистограммы начинают с разбиения ряда погрешностей на интервалы. Число интервалов k зависит от точности измерений, количества элементов в выборке и является в некотором смысле произвольным. Основное требование к количеству и величине интервалов заключается в том, чтобы полученный на их основе график был наглядным и правдоподобным. Длину интервала Q можно получить, например, используя следующие формулы:

$$Q = \frac{X_{\max} - X_{\min}}{k}, \tag{4.36}$$

если известно число интервалов k, и

$$Q = \frac{X_{\max} - X_{\min}}{1 + N}, \tag{4.37}$$

если используется количество элементов N в ряду (формула *Г. Стерджесса*). Полученное значение длины интервала Q округляют в сторону уменьшения до удобного для дальнейших вычислений числа. При этом значения крайних элементов ряда (максимальное и минимальное) должны обязательно лежать в пределах вычисленного интервала. Сами длины интервалов принято выражать в долях средней квадратической погрешности m.

В геодезии чаще всего в такого рода исследованиях ряд делят на 12 интервалов, каждый из которых должен иметь размер 0,5m.

Далее необходимо подсчитать число N_j элементов ряда, принадлежащих j-му интервалу, и вычислить практические оценки неизвестных вероятностей (частоты Q_j) по формуле

$$Q = \frac{N_j}{N}. \tag{4.38}$$

При этом необходимо проследить, чтобы сумма частот N_j по всем интервалам равнялась числу всех элементов в выборке, а сумма частот Q_j была равна единице в пределах ошибки округления.

Для построения гистограммы исследуемых случайных величин значение j-й частоты делят на длину принятого интервала, обычно равного половине средней квадратической ошибки. Таким образом, получают вертикальные составляющие гистограммы, называемые высотами прямоугольников:

$$h_j = \frac{Q_j}{0,5 \cdot m}. \tag{4.39}$$

Результаты вычислений целесообразно представлять в таблице с соответствующими контролями по числу элементов и частоте в интервалах (табл. 4.2).

Таблица 4.2

Значения для гистограммы эмпирического распределения

Интервал в долях m	Интервал в ед. измерения		N_j	Q_j	h_j
[–3,0m; –2,5m]	–2,7499	–2,2916	0	0,00	0,00
(–2,5m; –2,0m]	–2,2916	–1,8332	2	0,04	0,09
(–2,0m; –1,5m]	–1,8332	–1,3749	0	0,00	0,00
(–1,5m; –1,0m]	–1,3749	–0,9166	3	0,06	0,13
(–1,0m; –0,5m]	–0,9166	–0,4583	8	0,16	0,35
(–0,5m; 0m]	–0,4583	0,0000	7	0,14	0,31
(0m; 0,5m]	0,0000	0,4583	10	0,20	0,44
(0,5m; 1,0m]	0,4583	0,9166	10	0,20	0,44
(1,0m; 1,5m]	0,9166	1,3749	7	0,14	0,31
(1,5m; 2,0m]	1,3749	1,8332	2	0,04	0,09
(2,0m; 2,5m]	1,8332	2,2916	1	0,02	0,04
(2,5m; 3,0m]	2,2916	2,7499	0	0,00	0,00
		Σ =	50	1,00	

Примечание:
N_j – абсолютная частота
$Q_j = N_j / N$ – относительная частота
$h_j = Q_j / (0,5m)$ – сторона прямоугольника.

При таком представлении гистограммы площадь построенного прямоугольника будет равна величине соответствующей частоты, а общая площадь – примерно единице. Построение эмпирического распределения производят по значениям длин интервалов (ось абсцисс) и высот прямоугольников (ось ординат). На выбор масштабов накладывается лишь условие наглядности. Вид гистограммы дает возможность предположить о мере соответствия исследуемых величин нормальному закону распределения Гаусса.

На этом же графике необходимо построить теоретическую кривую, соответствующую нормальному закону, которая наилучшим образом сглаживает (выравнивает) данное эмпирическое распределение. Кривая строится на основе формулы плотности вероятности для закона Гаусса

$$\varphi\left(X\right)=\frac{1}{m\cdot\sqrt{2\pi}}\cdot e^{\frac{-\left(X-\mu\right)^2}{2\cdot m^2}}. \tag{4.40}$$

Учитывая, что $t = (X - \mu) / m$ – это нормированные границы интервалов, введем обозначение

$$K=\frac{1}{m\cdot\sqrt{2\pi}}, \tag{4.41}$$

тогда формулу (4.40) можно представить как

$$\varphi\left(X\right)=K\cdot e^{\frac{-t^2}{2}}. \tag{4.42}$$

Обычно величина t изменяется от –3 до 3 через 0,5, так что вычисления не представляют трудности. Необходимо учитывать, что функция симметричная, т. е. $f(-x) = -f(x)$. Значения функции (4.42) с $m = 1$ приведены в книгах по статистике или обработке измерений в виде таблиц и также могут быть использованы при вычислениях.

По значениям t и функции (4.42) на рисунок гистограммы наносится ряд точек, которые соединяются плавной сплошной линией (огивой). Эта линия и будет соответствовать теоретической кривой нормального закона распределения со стандартными характеристиками, которая визуально сравнивается с практическим представлением распределения в виде гистограммы (рис. 4.1).

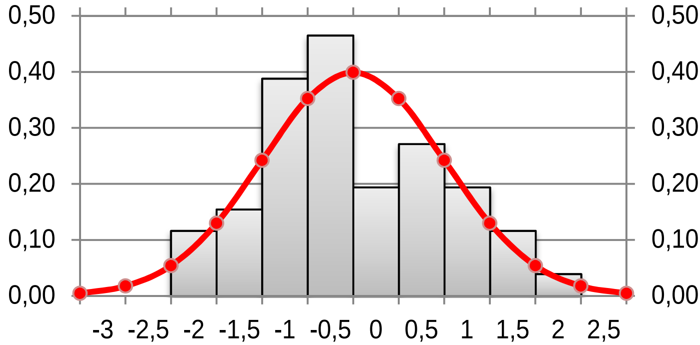

Рис. 4.1. Гистограмма и огива на землю

Точные критерии исследования

Строгую оценку степени соответствия эмпирического распределения теоретическому выполним на основе статистического аппарата *проверки гипотез*. Для этого задается уровень значимости q (вероятность невыполнения гипотезы) или β (вероятность выполнения гипотезы) и выдвигается гипотеза, что ряд исследуемых величин подчинен нормальному распределению Гаусса с вероятностью β. При этом, естественно, $\beta + q = 1$.

К точным критериям исследования соответствия эмпирического распределения теоретическому относят критерии согласия Колмогорова и Пирсона. Критерий согласия Колмогорова предназначен для проверки гипотезы о принадлежности выборки некоторому закону распределения, то есть проверки того, что эмпирическое распределение соответствует предполагаемой модели. Данный критерий рекомендуется применять при $N > 200$.

Для оценки степени приближения к нормальному закону по критерию χ^2 *Пирсона* следует вычислить величину

$$\chi^2 = \sum_{i=1}^{N} \frac{\left(N_i - N \cdot p_i\right)^2}{N \cdot p_i}. \qquad (4.43)$$

Здесь p_i – теоретическая вероятность попадания случайной величины в соответствующий интервал. Данная величина вычисляется по значениям интеграла вероятности $\Phi(t_i)$, которые выбираются из статистических таблиц по величинам границ интервалов t_i:

$$p_i = \frac{1}{2} \cdot \left(\Phi(t_{i+1}) - \Phi(t_i) \right) \tag{4.44}$$

и не деля на два, если используются таблицы нормированной функции Лапласа $\Phi(t_i)$ [3, с. 330]. Для вычислений можно использовать также любой статистический программный пакет для расчета вероятности попадания величин в соответствующий интервал.

Степень согласованности эмпирического распределения с теоретическим можно провести двумя способами. Во-первых, задать вероятность β, по которой по числу степеней свободы r из статистических таблиц χ^2 Пирсона (прил. Б) или из какой-либо программы (например, Excel) получают эталонное значение $\chi^2_{эт}$. Если вычисленное значение меньше эталонного, то с вероятностью β можно принять выдвинутую гипотезу о соответствии ряда нормальному закону распределения.

При втором подходе по вычисленной величине χ^2 и числу степеней свободы из статистических таблиц выбирают вероятность β того, что выдвинутая гипотеза принимается. При этом для целей геодезии можно использовать следующие характеристики для вероятностей:

$\beta > 0{,}5$ – отличное согласование с выдвинутой гипотезой о нормальности;

$0{,}3 < \beta < 0{,}5$ – согласие считают хорошим;

$0{,}1 < \beta < 0{,}3$ – согласие удовлетворительное;

$\beta < 0{,}1$ – согласие считается неудовлетворительным.

Число степеней свободы r определяется как $r = k - 1 - s$, где k – число интервалов, s – число определяемых для характеристики закона распределения параметров (чаще всего определяют математическое ожидание и дисперсию, тогда $s = 2$).

Все вычисления целесообразно свести в табл. 4.3.

Таблица 4.3

Критерий Пирсона

Доли *m*	$\frac{1}{2}\cdot\Phi(t_i)$	p_i	N_i	$N \cdot p_i$	Δ	χ^2
–3*m*	–0,499	0,0049	1	0,24	0,76	2,358
–2,5*m*	–0,494	0,0165	0	0,83	–0,83	0,827
–2*m*	–0,477	0,0441	1	2,20	–1,20	0,657
–1,5*m*	–0,433	0,0919	5	4,59	0,41	0,036
–1*m*	–0,341	0,1499	8	7,49	0,51	0,034
–0,5*m*	–0,191	0,1915	11	9,57	1,43	0,213
0*m*	0,000	0,1915	9	9,57	–0,57	0,034
0,5*m*	0,191	0,1499	8	7,49	0,51	0,034
1*m*	0,341	0,0919	4	4,59	–0,59	0,076
1,5*m*	0,433	0,0441	2	2,20	–0,20	0,019
2*m*	0,477	0,0165	0	0,83	–0,83	0,827
2,5*m*	0,494	0,0049	1	0,24	0,76	2,358
3*m*	0,499					
	Σ =	0,9973	50	49,87		7,474

Примечание:
Доли *m* – интервал в долях *m* от *M*(*X*);
$\frac{1}{2}\cdot\left(\Phi\left(t_i\right)\right)$ – границы вероятностной функции;
p_i – вероятность попадания в интервал;
N_i – практическая абсолютная частота;
$N \cdot p_i$ – теоретическая абсолютная частота;
$\Delta = N_i - N \cdot p_i$.

Список вопросов для защиты РГР

1. Для чего исследуют результаты опыта на соответствие какому-либо закону распределения?
2. Для чего исследуют результаты опыта на соответствие нормальному закону распределения?
3. Характеристики и вид нормального закона распределения.
4. Выявление мешающих параметров, его необходимость и суть.
5. Выявление грубых величин, его необходимость и суть.
6. Выявление систематического влияния, его необходимость и суть.
7. Когда ряд величин является случайным?
8. Определение основных оценок случайного ряда и их необходимость для достижения задачи исследования.
9. Приближенные критерии исследования ряда случайных величин на соответствие нормальному закону распределения.
10. Характеристики формы, их вычисление и суть.
11. Графические методы исследования на соответствие ряда нормальному закону распределения.
12. Определение частот попадания в интервал.
13. Построение теоретической гауссианы.
14. Выбор интервалов при построении гистограммы.
15. Вычисление теоретической вероятности попадания случайной величины в заданный интервал.
16. Точные критерии исследования ряда случайных величин на соответствие нормальному закону распределения, суть и последовательность вычислений.
17. Критерий хи-квадрат Пирсона.
18. Функция Лапласа и случаи ее использования при оценивании характеристик случайных величин.
19. Критерий Колмогорова, его недостатки.
20. Оценки центра распределения, их вычисление и суть.

ЛИТЕРАТУРА

1. Правила оформления [Электронный ресурс]. – Режим доступа: https://drive.google.com/file/d/1Z7EfWQckHqlSvEzHYDtEWR1mDAboH4pc/view. – Дата доступа: 10.10.2020.

2. Большаков, В. Д. Теория ошибок наблюдений : учебник для вузов / В. Д. Большаков. – 2-е изд., перераб. и доп. – М. : Недра, 1983. – 223 с.

3. Большаков, В. Д. Практикум по теории математической обработки геодезических измерений : учебное пособие для вузов / В. Д. Большаков, Ю. И. Маркузе. – М. : Недра, 1984. – 352 с.

4. Чеботарев, А. С. Способ наименьших квадратов с основами теории вероятностей / А. С. Чеботарев. – М. : Издательство геодезической литературы, 1958. – 610 с.

5. Leick, A. Adjustment Computations / A. Leick. – University of Maine, 1980. – 245 p.

6. Leick, A. Adjustments with examples / A. Leick, D. Humphrey. – University of Maine, 1986. – 450 p.

7. Дегтярев, А. М. Вероятностно-статистические методы в геодезии : конспект лекций / А. М. Дегтярев. – Новополоцк : ПГУ, 2005. – 208 с.

8. Михелев, Д. Ш. Геодезические измерения при изучении деформаций крупных инженерных сооружений / Д. Ш. Михелев, И. В. Рунов, А. И. Голубцов. – М. : Недра, 1977. – 152 с.

ПРИЛОЖЕНИЕ А

Односторонние и двусторонние критические значения коэффициента Стьюдента (t-критерий)

Односторонний	P = 0,90	0,95	0,975	0,99	0,995	0,9975	0,999	0,9995
Двусторонний	0,80	0,90	0,95	0,98	0,99	0,995	0,998	0,999
r = 1	3,0770	6,3130	12,7060	31,820	63,656	127,656	318,306	636,619
2	1,8850	2,9200	4,3020	6,964	9,924	14,089	22,327	31,599
3	1,6377	2,35340	3,182	4,540	5,840	7,458	10,214	12,924
4	1,5332	2,13180	2,776	3,746	4,604	5,597	7,173	8,610
5	1,4759	2,01500	2,570	3,649	4,0321	4,773	5,893	6,863
6	1,4390	1,943	2,4460	3,1420	3,7070	4,316	5,2070	5,958
7	1,4149	1,8946	2,3646	2,998	3,4995	4,2293	4,785	5,4079
8	1,3968	1,8596	2,3060	2,8965	3,3554	3,832	4,5008	5,0413
9	1,3830	1,8331	2,2622	2,8214	3,2498	3,6897	4,2968	4,780
10	1,3720	1,8125	2,2281	2,7638	3,1693	3,5814	4,1437	4,5869
11	1,363	1,795	2,201	2,718	3,105	3,496	4,024	4,437
12	1,3562	1,7823	2,1788	2,6810	3,0845	3,4284	3,929	4,178
13	1,3502	1,7709	2,1604	2,6503	3,1123	3,3725	3,852	4,220
14	1,3450	1,7613	2,1448	2,6245	2,976	3,3257	3,787	4,140
15	1,3406	1,7530	2,1314	2,6025	2,9467	3,2860	3,732	4,072
16	1,3360	1,7450	2,1190	2,5830	2,9200	3,2520	3,6860	4,0150
17	1,3334	1,7396	2,1098	2,5668	2,8982	3,2224	3,6458	3,965
18	1,3304	1,7341	2,1009	2,5514	2,8784	3,1966	3,6105	3,9216
19	1,3277	1,7291	2,0930	2,5395	2,8609	3,1737	3,5794	3,8834
20	1,3253	1,7247	2,08600	2,5280	2,8453	3,1534	3,5518	3,8495
21	1,3230	1,7200	2,2,0790	2,5170	2,8310	3,1350	3,5270	3,8190
22	1,3212	1,7117	2,0739	2,5083	2,8188	3,1188	3,5050	3,7921
23	1,3195	1,7139	2,0687	2,4999	2,8073	3,1040	3,4850	3,7676
24	1,3178	1,7109	2,0639	2,4922	2,7969	3,0905	3,4668	3,7454
25	1,3163	1,7081	2,0595	2,4851	2,7874	3,0782	3,4502	3,7251
26	1,315	1,705	2,059	2,478	2,778	3,0660	3,4360	3,7060
27	1,3137	1,7033	2,0518	2,4727	2,7707	3,0565	3,4210	3,6896
28	1,3125	1,7011	2,0484	2,4671	2,7633	3,0469	3,4082	3,6739
29	1,3114	1,6991	2,0452	2,4620	2,7564	3,0360	3,3962	3,8494
30	1,3104	1,6973	2,0423	2,4573	2,7500	3,0298	3,3852	3,6460

ПРИЛОЖЕНИЕ Б

Квантили распределения χ^2 для различной доверительной вероятности P и числа степеней свободы r

r	***P***										
	0,975	**0,95**	**0,9**	**0,8**	**0,7**	**0,6**	**0,5**	**0,4**	**0,3**	**0,2**	**0,1**
1	0,0010	0,0039	0,0158	0,0642	0,1485	0,2750	0,4549	0,7083	1,0742	1,6424	2,7055
2	0,0506	0,1026	0,2107	0,4463	0,7133	1,0217	1,3863	1,8326	2,4079	3,2189	4,6052
3	0,2158	0,3518	0,5844	1,0052	1,4237	1,8692	2,3660	2,9462	3,6649	4,6416	6,2514
4	0,4844	0,7107	1,0636	1,6488	2,1947	2,7528	3,3567	4,0446	4,8784	5,9886	7,7794
5	0,8312	1,1455	1,6103	2,3425	2,9999	3,6555	4,3515	5,1319	6,0644	7,2893	9,2364
6	1,2373	1,6354	2,2041	3,0701	3,8276	4,5702	5,3481	6,2108	7,2311	8,5581	10,6446
7	1,6899	2,1673	2,8331	3,8223	4,6713	5,4932	6,3458	7,2832	8,3834	9,8032	12,0170
8	2,1797	2,7326	3,4895	4,5936	5,5274	6,4226	7,3441	8,3505	9,5245	11,0301	13,3616
9	2,7004	3,3251	4,1682	5,3801	6,3933	7,3570	8,3428	9,4136	10,6564	12,2421	14,6837
10	3,2470	3,9403	4,8652	6,1791	7,2672	8,2955	9,3418	10,4732	11,7807	13,4420	15,9872
11	3,8157	4,5748	5,5778	6,9887	8,1479	9,2373	10,3410	11,5298	12,8987	14,6314	17,2750
12	4,4038	5,2260	6,3038	7,8073	9,0343	10,1820	11,3403	12,5838	14,0111	15,8120	18,5493
13	5,0088	5,8919	7,0415	8,6339	9,9257	11,1291	12,3398	13,6356	15,1187	16,9848	19,8119
14	5,6287	6,5706	7,7895	9,4673	10,8215	12,0785	13,3393	14,6853	16,2221	18,1508	21,0641
15	6,2621	7,2609	8,5468	10,3070	11,7212	13,0297	14,3389	15,7332	17,3217	19,3107	22,3071
16	6,9077	7,9616	9,3122	11,1521	12,6243	13,9827	15,3385	16,7795	18,4179	20,4651	23,5418
17	7,5642	8,6718	10,0852	12,0023	13,5307	14,9373	16,3382	17,8244	19,5110	21,6146	24,7690
18	8,2307	9,3905	10,8649	12,8570	14,4399	15,8932	17,3379	18,8679	20,6014	22,7595	25,9894
19	8,9065	10,1170	11,6509	13,7158	15,3517	16,8504	18,3377	19,9102	21,6891	23,9004	27,2036
20	9,5908	10,8508	12,4426	14,5784	16,2659	17,8088	19,3374	20,9514	22,7745	25,0375	28,4120
21	10,2829	11,5913	13,2396	15,4446	17,1823	18,7683	20,3372	21,9915	23,8578	26,1711	29,6151
22	10,9823	12,3380	14,0415	16,3140	18,1007	19,7288	21,3370	23,0307	24,9390	27,3015	30,8133
23	11,6886	13,0905	14,8480	17,1865	19,0211	20,6902	22,3369	24,0689	26,0184	28,4288	32,0069
24	12,4012	13,8484	15,6587	18,0618	19,9432	21,6525	23,3367	25,1063	27,0960	29,5533	33,1962
25	13,1197	14,6114	16,4734	18,9398	20,8670	22,6156	24,3366	26,1430	28,1719	30,6752	34,3816
26	13,8439	15,3792	17,2919	19,8202	21,7924	23,5794	25,3365	27,1789	29,2463	31,7946	35,5632
27	14,5734	16,1514	18,1139	20,7030	22,7192	24,5440	26,3363	28,2141	30,3193	32,9117	36,7412
28	15,3079	16,9279	18,9392	21,5880	23,6475	25,5093	27,3362	29,2486	31,3909	34,0266	37,9159
29	16,0471	17,7084	19,7677	22,4751	24,5770	26,4751	28,3361	30,2825	32,4612	35,1394	39,0875
30	16,7908	18,4927	20,5992	23,3641	25,5078	27,4416	29,3360	31,3159	33,5302	36,2502	40,2560

ПРИЛОЖЕНИЕ В

Значения критерия Фишера (F-критерия)
для уровня значимости $\alpha = 0{,}05$

	r_1										
r_2	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**15**
1	161,45	199,50	215,71	224,58	230,16	233,99	236,77	238,88	240,54	241,88	245,95
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,43
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,70
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,86
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,62
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	3,94
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,51
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,22
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,01
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,85
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,72
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,62
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,53
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,46
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,40
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,35
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,31
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,27
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,23
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,20

Примечание:
r_1 – число степеней свободы большей дисперсии;
r_2 – число степеней свободы меньшей дисперсии.

Содержание

Учебное издание

БУДО Андрей Юрьевич

ТЕОРИЯ ПОГРЕШНОСТЕЙ ИЗМЕРЕНИЙ

Учебно-методическое пособие
для студентов специальности
1-56 02 01 «Геодезия»

Редактор *Н. А. Костешева*
Компьютерная верстка *Н. А. Школьниковой*

Подписано в печать 28.06.2022. Формат 60×84 $^{1}/_{16}$. Бумага офсетная. Ризография.
Усл. печ. л. 2,91. Уч.-изд. л. 2,27. Тираж 50. Заказ 290.

Издатель и полиграфическое исполнение: Белорусский национальный технический университет.
Свидетельство о государственной регистрации издателя, изготовителя, распространителя
печатных изданий № 1/173 от 12.02.2014. Пр. Независимости, 65. 220013, г. Минск.