Министерство образования Республики Беларусь

БЕЛОРУССКИЙ НАЦИОНАЛЬНЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

Факультет транспортных коммуникаций

Кафедра «Геодезия и аэрокосмические геотехнологии»

**ОТЧЁТ**

по тематической контрольной работе №1 «Исследование ряда погрешностей на соответствие нормальному закону распределения»

по дисциплине «Теория математической обработки геодезических измерений»

Вариант 11

|  |  |
| --- | --- |
| Выполнил: Студент гр. 31401123 | Д.Г. Подсуконный |
| Проверил:  Старший преподаватель кафедры ГиАКГТ | А.Ю. Будо |

Минск

2025

**ВВЕДЕНИЕ**

Цель: исследовать ряд погрешностей результатов измерений спутников на соответствие нормальному закону распределения с предварительным анализом на систематические и грубые ошибки.

Исходными данными для работы являются данные измерения спутников.

**Ход работы**

**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 Хцентр мы будем вычислять его пятью различными способами:

Хмедиана, Хцентр\_сгибов,, 50% , Хцентр\_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. Хмедиана - обычная медиана. (0,00041)
2. - среднее арифметическое. (0,000656)
3. Хцентр\_размаха- среднее между максимальным и минимальным значением в выборке. (0,031316)

4) Хцентр\_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.

М=ЦЕЛОЕ(22418/4)= 5604. (1)

Тогда центр сгибов определяется по формуле:

Хцентр\_сгибов= = 0,062356. (2)

5) Ẋ50% - среднее арифметическое по 50%-му интерквантильному промежутку вычисляется по формуле:

Ẋ50%== 0,00092; (3)

Хцентр=0,00092.

**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 введём понятие коэффициента цензурирования - безразмерной величины G, при которой промахами считаются все значения из выборки, лежащие за пределами интервала

Хцентр ≤ x ≤ Хцентр, (4)

-2,446477 ≤ x ≤ 2,448317,

где

= 0,118218. (5)

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 (затем более плосковершинного) распределения, для островершинного распределения 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+0.8lg(N/10) = 20,702418, (6)

где

Е= = 52,053193, (7)

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.

**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 Xист и произведено *N* ее измерений хi. В этом случае пользуются зависимостью

, (8)

где

, (9)

средняя квадратическая погрешность среднего арифметического

, (10)

m – средняя квадратическая погрешность (СКП) исследуемой величины,

N – число элементов в ряду.

zq – 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 или вычисляется, например, в Excel. Если неравенство выполняется, то с вероятностью *β* = 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

= 1,072493, (11)

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

= 0,118218, (12)

и

=0,126788, (13)

где уклонение *i*-той величины от среднего

, (14)

последовательные разности

. (15)

В (13) 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 δq. Тогда, при δ>δq принимается гипотеза об отсутствии систематической ошибки с вероятностью Р=1 - q. В противном случае (δ<δq) 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 Хисп с уменьшенной по сравнению с исходным рядом систематической составляющей.

**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 её теоретической кривой закона Гаусса с параметрами 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 Qможно получить, например, используя следующие формулы

(16)

если известно число интервалов k, и

, (17)

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 соответствующих им единицах произведенных измерений (секунды, метры и т.д.).

В геодезии чаще всего в такого рода исследованиях ряд делят на 12 интервалов, каждый из которых должен иметь размер 0.5 m.

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

. (18)

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

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

. (19)

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

Таблица №1 - Вычисление значений для гистограммы эмпирического распределения

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Интервал в долях m | Интервал в ед. измерения | | Абсолютная частота Nj | Относительная частота Qj=Nj/N | Сторона прямоуг.  hj=Qj/(0,5\*m) |
| [-3,0m; -2,5m] | -0,354653 | -0,295544 | 179 | 0,008021 | 0,135695 |
| [-2,5m; -2,0m] | -0,295544 | -0,236435 | 451 | 0,020209 | 0,341892 |
| [-2,0m; -1,5m] | -0,236435 | -0,177326 | 754 | 0,033786 | 0,571588 |
| [-1,5m; -1,0m] | -0,177326 | -0,118218 | 1921 | 0,086078 | 1,456261 |
| [-1,0m; -0,5m] | -0,118218 | -0,059109 | 3190 | 0,142940 | 2,418258 |
| [-0,5m; 0m] | -0,059109 | 0 | 4672 | 0,209347 | 3,541724 |
| [0m; 0,5m] | 0 | 0,059109 | 4909 | 0,219967 | 3,721388 |
| [0,5m; 1,0m] | 0,059109 | 0,118218 | 2990 | 0,133979 | 2,266643 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Продолжение таблицы №1 | | | | | |
| [1,0m; 1,5m] | 0,118218 | 0,177326 | 1799 | 0,080611 | 1,363776 |
| [1,5m; 2,0m] | 0,177326 | 0,236435 | 861 | 0,038580 | 0,652702 |
| [2,0m; 2,5m] | 0,236435 | 0,295544 | 468 | 0,020971 | 0,354779 |
| [2,5m; 3,0m] | 0,295544 | 0,354653 | 224 | 0,010037 | 0,169809 |
|  | Σ= | | 22418 | 1 |  |

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

. (20)

Учитывая, что *t=(X-µ)/m* – это нормированные границы интервалов, введём обозначение

, (21)

тогда формулу (20) можно представить как

, (22)

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 (22) 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 (22) 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 (Рисунок 1).

Рисунок 1 – Гистограмма и огива

**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, β + 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 Пирсона следует вычислить величину

. (23)

Здесь – теоретическая вероятность попадания случайной величины в соответствующий интервал, Данная величина вычисляется по значениям интеграла вероятностиФ(), которые выбираются из статистических таблиц по величинам границ интервалов

, (24)

и не деля на два, если используются таблицы нормированной функции Лапласа. 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 Пирсона или из какой-либо программы (например, Excel) получают эталонное значение . Есливычисленное значение меньше эталонного, то с вероятностью β можно принять выдвинутую гипотезу о соответствии ряда нормальному закону распределения.

При втором подходе по вычисленной величине 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:

β> 0.5 – отличное согласование с выдвинутой гипотезой о нормальности;

0.3 <β< 0.5 – согласие считают хорошим;

0.1 <β< 0.3 – согласие удовлетворительное;

β< 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 были сведены таблицу (Таблица 2).

Таблица №2 – Вычисления критерия x2 Пирсона

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Интервал в долях m от M(X) | Границы вероятностной  ф-ии  ½\*Ф(ti) | Вероятность попадания в интервал pi | Практическая абсолютная частота Ni | Теоретическая абсолютная частота  N\*pi | Ni-N\*pi | *X2* |
| [-3,0m; -2,5m] | -0,499 | 0,005 | 179 | 109,353 | 69,647 | 0,637 |
| [-2,5m; -2,0m] | -0,494 | 0,017 | 451 | 368,231 | 82,770 | 0,225 |
| [-2,0m; -1,5m] | -0,477 | 0,044 | 754 | 984,180 | -230,180 | -0,234 |
| [-1,5m; -1,0m] | -0,433 | 0,092 | 1921 | 2050,932 | -129,932 | -0,063 |
| [-1,0m; -0,5m] | -0,341 | 0,150 | 3190 | 3345,318 | -155,318 | -0,046 |
| [-0,5m; 0m] | -0,191 | 0,192 | 4672 | 4273,706 | 398,295 | 0,093 |
| [0m; 0,5m] | 0,000 | 0,192 | 4909 | 4273,706 | 635,295 | 0,149 |
| [0,5m; 1,0m] | 0,191 | 0,150 | 2990 | 3345,318 | -355,318 | -0,106 |
| Продолжение таблицы №2 | | | | | | |
| [1,0m; 1,5m] | 0,341 | 0,092 | 1799 | 2050,932 | -251,932 | -0,123 |
| [1,5m; 2,0m] | 0,433 | 0,044 | 861 | 984,180 | -123,180 | -0,125 |
| [2,0m; 2,5m] | 0,477 | 0,017 | 468 | 368,231 | 99,770 | 0,271 |
| [2,5m; 3,0m] | 0,494 | 0,005 | 224 | 109,353 | 114,647 | 1,048 |
|  |  | 1 |  |  |  | 1,725 |

Х2выч=1,725 < Х2эт= 2,700; т.к вычисленное значение Х2выч меньше Х2эт то можно принять гипотезу о соответствии ряда нормальному закону распределения.

**ВЫВОД**

В процессе выполнения тематической контрольной работы я исследовал ряд погрешностей результатов измерений спутников на соответствие нормальному закону распределения с предварительным анализом на систематические и грубые ошибки.

**СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ**

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

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

3. Чеботарёв А.С. Способ наименьших квадратов с основами теории вероятностей. Издательствогеодезическойлитературы, 1958. – 610 с.

4. Leick A. Adjustment Computations. – Department of Spatial Information Science and Engineering. University of Maine, 1980. – 245 p.

5. Leick A., Humphrey D. Adjustments with examples. – University of Maine, 1986

6. Дегтярёв А.М. Вероятностно-статистические методы в геодезии. Конспект лекций. – Новополоцк: ПГУ, 2005. – 208 с.

7. 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.