



МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«ДОНСКОЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

Кафедра «Геодезия»

МЕТОДЫ ПЛАНИРОВАНИЯ ЭКСПЕРИМЕНТА В ПРИКЛАДНОЙ ГЕОДЕЗИИ

МЕТОДИЧЕСКИЕ УКАЗАНИЯ
к практическим занятиям

Ростов-на-Дону
2023

УДК 519.01242: 528.02

Составители: О.А. Губеладзе, А.Р. Губеладзе

Методы планирования эксперимента в прикладной геодезии:
методические указания к практическим занятиям / сост.

О.А. Губеладзе, А.Р. Губеладзе. – Ростов-на-Дону: Донской гос. техн.
ун-т, 2023. - 12 с.

Предназначены для обучающихся в магистратуре по направлению
подготовки 21.04.03 "Геодезия и дистанционное зондирование".

УДК 519.01242: 528.02

Печатается по решению редакционно-издательского совета
Донского государственного технического университета

Ответственный за выпуск зав. кафедрой «Геодезия»
канд. техн. наук, доцент М.А. Николенко

В печать 28.04.2023 г.
Формат 60×84/16. Объем 0,8 усл. п. л.
Тираж 50 экз. Заказ № 709

Издательский центр ДГТУ
Адрес университета и полиграфического предприятия:
344003, г. Ростов-на-Дону, пл. Гагарина, 1

© Донской государственный
технический университет, 2023

ПРАКТИЧЕСКОЕ ЗАНЯТИЕ № 1

Статистический анализ экспериментальных данных

Для графического изображения дискретных статистических рядов используется полигон. **Полигоном частот** (полигоном относительных частот) называется ломаная линия с вершинами в точках $(x_i^*; n_i)$. Для графического изображения интервальных статистических рядов используется гистограмма. **Гистограмма частот** (**гистограмма относительных частот**) — ступенчатая фигура, состоящая из прямоугольников, построенных на интервалах группировки и имеющих высоты, равные n_i / h (соответственно n_i / nh), где h — длина интервалов группировки. При больших значениях n гистограмма относительных частот является хорошим приближением для графика плотности распределения наблюдаемой случайной величины. Поэтому по виду гистограммы можно выдвинуть предположение (гипотезу) о распределении изучаемой величины. **Выборка** представляет собой ряд наблюдений за одной и той же случайной величиной. Для содержательного статистического анализа экспериментальных данных необходимо знать распределение этой величины. **Точечной оценкой** параметра Θ называется любая статистика $\hat{\Theta}_n$, предназначенная для оценивания этого параметра и определяемая одним числом. Точечная оценка практически никогда не совпадает с истинным значением параметра, она может только оценивать его с большей или меньшей точностью. Оценкой для математического ожидания является **выборочное среднее**, которое вычисляется по формуле

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i.$$

Оценка для **дисперсии** рассчитывается по формуле

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2.$$

Эту оценку называют **несмещенной оценкой дисперсии**. Число $f = n - 1$ называется **числом степеней свободы** этой оценки дисперсии и равно количеству независимых наблюдений (объему выборки), по которым

вычисляется данная оценка дисперсии, минус число параметров, которые оцениваются по этой выборке, кроме дисперсии.

Задача 1.1. При проверке прочности бетонных образцов получены следующие данные по маркам:

M50	M75	M50	M75	M50	M50	M75	M100	M150	M75
2	1	2	3	3	2	1	2	2	2

Составить статистический ряд, построить полигон частот, найти эмпирическую функцию распределения и построить ее график.

Задача 1.2. Из большой партии по схеме случайной повторной выборки было проверено 150 проб грунта с целью определения процента влажности. Получены следующие результаты:

$[x_{i-1}; x_i)$	[11; 13)	[13; 15)	[15; 17)	[17; 19)	[19; 21)
n_i	8	42	51	37	12

Построить гистограмму относительных частот, найти эмпирическую функцию распределения и построить ее график. Можно ли по виду гистограммы предложить, что наблюдаемая случайная величина имеет нормальное распределение?

Задача 1.3. При проверке прочности бетонных кубиков-образцов были получены следующие результаты, МПа: 200; 220; 250; 210; 230. Найти несмещенные оценки для математического ожидания и дисперсии, указать число степеней свободы оценки дисперсии.

Задача 1.4. По выборке

$[x_{i-1}; x_i)$	[5; 10)	[10; 15)	[15; 20)	[20; 25)	[25; 30)	[30; 35)	[35; 40)
n_i	4	6	16	36	24	10	4

найти несмещенные оценки для математического ожидания и дисперсии, указать число степеней свободы оценки дисперсии.

Задача 1.5. Для определения прочности бетона было испытано три бетонных кубика. Результаты испытаний - 19,8; 20,1; 20,4 МПа. Сколько надо

провести таких испытаний, чтобы с надежностью 0,95 ошибка при определении средней прочности была в пределах 0,2 МПа, если считается, что ошибки прибора нормальны?

Задача 1.6. Даны результаты 5 независимых равноточных измерений толщины металлической пластинки: 2,015; 2,020; 2,025; 2,020; 2,015 мм. Нужно: 1) оценить с помощью доверительного интервала истинную толщину пластинки при доверительной вероятности $\gamma = 0,95$;

2) найти минимальное число измерений, которое надо выполнить, чтобы с надежностью 0,95 можно было утверждать, что предельная погрешность точечной оценки истинной толщины металлической пластинки не превышает 0,003.

ПРАКТИЧЕСКОЕ ЗАНЯТИЕ № 2

Определение коэффициентов эмпирического уравнения регрессии в случае линейной однофакторной зависимости. Криволинейная регрессия

Пусть имеется выборка объема N наблюдений над двумя величинами x и y и принята гипотеза о линейной регрессионной зависимости между ними, т. е. оценкой для модельного уравнения регрессии y на x является эмпирическое линейное уравнение регрессии $\hat{y} = b_0 + b_1 x$. Далее для нахождения коэффициентов b_0 и b_1 используем **метод наименьших квадратов (МНК)**. Идея МНК заключается в следующем: значения коэффициентов b_1 и b_0 выбирают так, чтобы сумма квадратов отклонений наблюдаемых значений y_i от предсказываемых по уравнению регрессии $\hat{y} = b_0 + b_1 x_i$ была наименьшей.

Задача 2.1. Для разработки методики прогнозирования морозостойкости керамического кирпича на предприятии проводился сбор статистической информации. Измеряемые свойства и их величины приведены в таблице:

Плотность x_1 , %	Водопоглощение под вакуумом x_2 , %	Прочность на сжатие x_3 , %	Морозостойкость y , цикл
19,5	18,4	15,4	30
18,5	17,9	15,4	30
37,9	18,9	13,7	76
25,8	19	15,1	41
31,7	18,7	15,4	60
33,1	18,2	14,2	54
34	18,9	14,1	65
27,2	19,2	15,1	40

По имеющимся данным для переменных x и y требуется:

- 1) построить корреляционное поле;
- 2) найти выборочный коэффициент корреляции и проверить его значимость при $\alpha = 0,05$;
- 3) определить коэффициенты эмпирического линейного уравнения регрессии, если это целесообразно;
- 4) построить прямую на корреляционном поле.

Вспомогательные величины: $\sum x_1 i = 227,7$; $\sum x_1^2 i = 6820,89$; $\sum y i = 396$; $\sum y^2 = 21598$; $\sum x_1 i y i = 12065,6$.

На практике зависимость между двумя величинами далеко не всегда можно выразить линейной функцией. По расположению точек на корреляционном поле исследователем определяется вид зависимости с точностью до нескольких параметров. Предварительно с помощью замены переменных функцию стараются свести к линейной (если это возможно), а далее для определения искомых параметров может быть использован метод наименьших квадратов. Следует иметь в виду, что МНК может быть использован для функции любого типа, но в данном случае исследователь сталкивается с решением достаточно сложных систем нелинейных уравнений.

Задача 2.2. Установить зависимость твердения бетона от времени:

t , сут.	1	3	6	10	16	21	28
V , %	12	32	51	72	93	97	100

Во многих случаях необходимо исследовать зависимость величины y от нескольких переменных X_1, X_2, \dots, X_n . Эти переменные называются **факторами**, а зависимая переменная y - **откликом (параметром оптимизации)**. Зависимость отклика от изучаемых факторов $y = f(X_1, X_2, \dots, X_n)$ называется **функцией отклика**. Поскольку вид функции отклика, как правило, неизвестен, ее представляют в виде полинома (многочлена) и, находя по экспериментальным данным оценки коэффициентов полинома, получают эмпирическое уравнение регрессии

Задача 2.3. По имеющимся результатам эксперимента

Номер опыта	X_{1i}	X_{2i}	y_i
1	1	0	6
2	0	2	1
3	1	-1	8
4	0	1	2

получить (если это возможно):

- 1) линейное уравнение регрессии;
- 2) уравнение вида $\hat{y} = b_0 + b_{12}X_1X_2 + b_{22}X_2^2$;
- 3) квадратичное уравнение регрессии;
- 4) уравнение $\hat{y} = b_0 + b_1X_1 + b_2X_2 + b_{11}X_1^2$.

ПРАКТИЧЕСКОЕ ЗАНЯТИЕ № 3

Матрица эксперимента

Эксперимент включает в себя постановку задачи, которая будет решаться. На этом этапе выбирается независимая переменная (или несколько переменных). Решается вопрос, насколько точно можно измерить исследуемые величины.

Устанавливается, какие независимые переменные (факторы) могут влиять на зависимую величину или отклик и каким распределением они могут быть описаны. На этапе планирования определяется размер выборки, порядок рандомизации, строится математическая модель для описания эксперимента. Окончательным этапом является анализ. К нему относятся процессы сбора и упорядочения данных вычисления некоторых статистик, необходимых для принятия решений, а также выбор соответствующих правил принятия решений для проверки математических гипотез относительно математической модели.

Задача 41. Исходные параметры технологического процесса составляют:

толщина объекта 55 мм;

время воздействия нагрузки 30 с.

Составить матрицу эксперимента и получить линейную модель процесса.

Решение:

$$x_{1c} = 55; \quad x_{2c} = 30.$$

Возьмем верхние и нижние значения обоих факторов так, чтобы они располагались симметрично относительно текущего значения:

$$x_{1в} = 60; \quad x_{2в} = 35;$$

$$x_{1н} = 50. \quad x_{2н} = 25.$$

Расположение экспериментальных точек в двумерном факторном пространстве (рис.3.1).

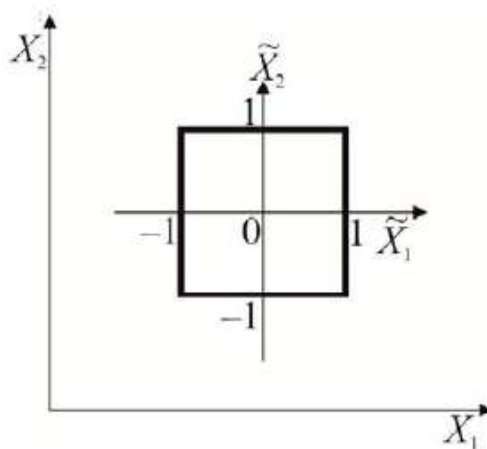


Рис.3.1

Составим таблицу, в которой значения обоих факторов находятся во всех возможных сочетаниях, и проведем измерения в этих точках (значения отклика приводим условно):

x_1	x_2	y
50	25	140
50	35	210
60	25	170
60	35	220

На основании полученных результатов составляем систему четырех уравнений с двумя переменными. Ниже показана эта система, а также ее сокращенная запись в виде матрицы.

$$\begin{cases} a_0 + 50a_1 + 25a_2 = 140, \\ a_0 + 50a_1 + 35a_2 = 210, \\ a_0 + 60a_1 + 25a_2 = 170, \\ a_0 + 60a_1 + 35a_2 = 220. \end{cases} \quad \begin{pmatrix} x_0 & x_1 & x_2 & y \\ 1 & 50 & 25 & 140 \\ 1 & 50 & 35 & 210 \\ 1 & 60 & 25 & 170 \\ 1 & 60 & 35 & 220 \end{pmatrix}.$$

Матрицу данного вида назовем матрицей эксперимента.

В матрице эксперимента второй и третий столбцы представляют собой значения факторов, четвертый столбец — значения отклика системы, а первый содержит единицы, соответствующие единичным коэффициентам свободного члена модели a_0 . Будем считать этот столбец некоторым виртуальным фактором x_0 , который всегда принимает единичные значения. Решим систему, переходя к нормированным координатам. Чтобы облегчить решение системы, проведем нормировку факторов. Верхним значениям факторов присвоим нормированное значение $+1$, нижним значениям — нормированное значение -1 , среднему значению — нормированное 0 . В общем виде нормировка фактора выражается формулой

$$\tilde{x} = \frac{2(x_i - x_{i\text{с}})}{x_{i\text{в}} - x_{i\text{н}}} = \frac{2x_i - x_{i\text{в}} - x_{i\text{н}}}{x_{i\text{в}} - x_{i\text{н}}} = \frac{x_i - x_{i\text{с}}}{\Delta},$$

где $\Delta = (x_{iB} - x_{iH})/2$.

С учетом нормировки факторов система уравнений и матрица эксперимента примут следующий вид:

$$\begin{cases} \tilde{a}_0 - \tilde{a}_1 - \tilde{a}_2 = 140, \\ \tilde{a}_0 - \tilde{a}_1 + \tilde{a}_2 = 210, \\ \tilde{a}_0 + \tilde{a}_1 - \tilde{a}_2 = 170, \\ \tilde{a}_0 + \tilde{a}_1 + \tilde{a}_2 = 220. \end{cases} \quad \begin{pmatrix} \tilde{x}_0 & \tilde{x}_1 & \tilde{x}_2 & y \\ +1 & -1 & -1 & 140 \\ +1 & -1 & +1 & 210 \\ +1 & +1 & -1 & 170 \\ +1 & +1 & +1 & 220 \end{pmatrix}.$$

Поскольку сумма членов во втором и третьем столбце матрицы равны нулю, свободный член модели можно найти, сложив все четыре уравнения:

$$4\tilde{a}_0 = 140 + 210 + 170 + 220 = 740;$$

$$\tilde{a}_0 = 185.$$

Чтобы найти какой-либо другой коэффициент модели, нужно изменить знаки в уравнениях таким образом, чтобы в соответствующем столбце оказались одни единицы, после чего сложить все четыре уравнения.

В общем случае решение системы будет выглядеть как

$$\tilde{a}_k = \frac{1}{2^n} \sum_{i=1}^{2^n} y_i \tilde{x}_{k,i}.$$

Тогда остальные параметры будут равны

$$4\tilde{a}_1 = -140 - 210 + 170 + 220 = 40;$$

$$\tilde{a}_1 = 10.$$

$$4\tilde{a}_2 = -140 + 210 - 170 + 220 = 120;$$

$$\tilde{a}_2 = 30.$$

Таким образом, линейная модель технологического процесса в окрестностях точки (55, 30) имеет вид

$$y = 185 + 10\tilde{x}_1 + 30\tilde{x}_2.$$

Возвращаемся к ненормированным факторам. Переход от нормированных к ненормированным факторам осуществляется обратным преобразованием:

$$x_i = \tilde{x}_i \Delta + x_{i\text{с}} = \tilde{x}_i \Delta + \frac{x_{i\text{в}} + x_{i\text{н}}}{2}.$$

Чтобы найти параметры модели для ненормированных координат, подставим выражения для нормированных координат в уравнение модели:

$$\begin{aligned} y &= \tilde{a}_0 + \tilde{a}_1 \tilde{x}_1 + \tilde{a}_2 \tilde{x}_2 = \tilde{a}_0 + \tilde{a}_1 \frac{2(x_1 - x_{1\text{с}})}{x_{1\text{в}} - x_{1\text{н}}} + \tilde{a}_1 \frac{2(x_2 - x_{2\text{с}})}{x_{2\text{в}} - x_{2\text{н}}} = \\ &= \tilde{a}_0 - \tilde{a}_1 \frac{2x_{1\text{с}}}{x_{1\text{в}} - x_{1\text{н}}} - \tilde{a}_2 \frac{2x_{2\text{с}}}{x_{2\text{в}} - x_{2\text{н}}} + \frac{2\tilde{a}_1}{x_{1\text{в}} - x_{1\text{н}}} x_1 + \frac{2\tilde{a}_2}{x_{2\text{в}} - x_{2\text{н}}} = \\ &= \tilde{a}_0 - \tilde{a}_1 \frac{x_{1\text{в}} + x_{1\text{н}}}{x_{1\text{в}} - x_{1\text{н}}} - \tilde{a}_2 \frac{x_{2\text{н}} + x_{2\text{н}}}{x_{2\text{н}} - x_{2\text{н}}} + \frac{2\tilde{a}_1}{x_{1\text{в}} - x_{1\text{н}}} x_1 + \frac{2\tilde{a}_2}{x_{2\text{в}} - x_{2\text{н}}} x_2. \end{aligned}$$

Сравнивая последнее выражение с выражением для линейной модели в ненормированных координатах, получим выражение для параметров модели:

$$\begin{aligned} a_0 &= \tilde{a}_0 - \tilde{a}_1 \frac{x_{1\text{в}} + x_{1\text{н}}}{x_{1\text{в}} - x_{1\text{н}}} - \tilde{a}_2 \frac{x_{2\text{в}} + x_{2\text{н}}}{x_{2\text{в}} - x_{2\text{н}}}; \\ a_1 &= \frac{2\tilde{a}_1}{x_{2\text{в}} - x_{1\text{н}}}; \\ a_2 &= \frac{2\tilde{a}_2}{x_{2\text{в}} - x_{2\text{н}}}. \end{aligned}$$

Для приведенного выше примера:

$$\begin{aligned} a_0 &= 185 - 10 \frac{60 + 50}{60 - 50} - 30 \frac{35 + 25}{35 - 25} = -105; \\ a_1 &= \frac{2 \cdot 10}{60 - 50} = 2; \\ a_2 &= \frac{2 \cdot 30}{35 - 25} = 6. \end{aligned}$$

Окончательно получаем модель в естественных координатах:

$$y = -105 + 2x_1 + 6x_2.$$

ЛИТЕРАТУРА

1. ГОСТ 31937-2011 Межгосударственный стандарт. Здания и сооружения. Правила обследования и мониторинга технического состояния (принят Межгосударственной научно-технической комиссией по стандартизации, техническому нормированию и сертификации в строительстве. Протокол № 39 от 08 декабря 2011 г.)
2. СТО СРО-П 60542948 00005-2012 Обследование строительных конструкций.
3. Калинин, В.М. Обследование и испытание конструкций зданий и сооружений [Текст] / – М.: ИНФРА-М, 2005. – 336 с.
4. Красовский Г. И. Планирование эксперимента. — Минск : БГУ, 1982. — 302 с.
5. Планирование эксперимента [Электронный ресурс]: учебно-практическое пособие / Т. В. Ерещенко, Н. А. Михайлова ; М-во образования и науки Рос.Федерации, Волгогр. гос. архит.-строит. ун-т. — Волгоград : ВолгГАСУ, 2014.