



МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
«РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ –
МСХА имени К.А. ТИМИРЯЗЕВА»

Левшин А.Г., Левшин А.А., Бутузов А.Е., Майстренко Н.А.

«Планирование и организация эксперимента»
Учебное пособие

Москва
Издательство РГАУ-МСХА
2016

УДК 621.77:631.3

ББК

К 55

Левшин А.Г. Планирование и организация эксперимента: Учебное пособие/
А.Г. Левшин, А.А. Левшин, А.Е. Бутузов, Н.А. Майстренко- М.: Изд-во РГАУ-МСХА,
2016.- 65 с.

В учебном пособии представлена методика статистического анализа опытных данных, описания факторного пространства уравнениями регрессии и оценки их достоверности при выполнении индивидуальных заданий в ходе лабораторно-практических занятий по дисциплине «Планирование и организация эксперимента».

Учебное пособие предназначено для самостоятельной работы студентов, обучающихся по программам магистратуры по направлениям: «Агроинженерия», «Стандартизация и метрология», «Управление качеством» и «Эксплуатация транспортно-технологических машин и комплексов». Отдельные задания могут быть рекомендованы бакалаврам. Будут полезны для повышения квалификации специалистов, занимающихся вопросами испытания сельскохозяйственной техники и аспирантам.

Рецензенты: доктор техн. наук, проф. **Лобачевский Я.П.**;
доктор техн. наук, проф. **Девянин С.Н.**

ISBN

© Левшин А.Г., Левшин А.А.,
Бутузов А.Е., Майстренко Н.А.
© РГАУ-МСХА им. К.А. Тимирязева 2015;
© Издательство РГАУ-МСХА.

Содержание

Темы практических занятий

	Введение	4
1	Статистическая обработка одномерной выборки случайной величины.	6
2	Статистическая обработка выборки большого объема	14
3	Выбор закона распределения случайной величины	18
4	Оценка статистической взаимосвязи двух случайных величин.	22
5	Аппроксимация опытных данных математическими зависимостями по методу наименьших квадратов.	26
6	Планирование полного факторного эксперимента (ПФЭ) 2^n	32
7	Планирование дробного факторного эксперимента (ДФЭ) 2^{n-k}	36
8	Поиск области экстремума функции отклика методом крутого восхождения (спуска)	40
9	Описание области экстремума функции отклика уравнениями второго порядка.	43
	Приложение 1. Таблицы критических значений статистических распределений	49

Введение

Учебное пособие «Планирование и организация эксперимента» предназначено для самостоятельной работы студентов по изучению теоретических основ экспериментального изучения объекта исследования, формированию навыков и умений по планированию, проведению эксперимента и анализу полученных результатов.

Цель дисциплины - освоение студентами (магистрами, аспирантами) методики планирования и организации эксперимента для получения достоверной информации об исследуемом объекте при минимальных затратах труда и в кратчайшие сроки.

Задачи дисциплины – научить оценивать достоверность полученной экспериментальной информации; определять характеристики случайных величин и минимально необходимый объем выборок; выбирать и ранжировать наиболее существенные технологические факторы, влияющие на исследуемый процесс; описывать функцию отклика для выбранного факторного пространства и определять оптимальную область для исследуемого процесса; оценивать парную и множественную статистическую взаимосвязь между случайными величинами и аппроксимировать опытные данные математическими зависимостями; оценивать точность и адекватность математических моделей.

Требования к уровню освоения содержания дисциплины. В результате изучения дисциплины студент (магистр, аспирант) должен:

Иметь представление: о методах экспериментальных исследований процессов работы машин.

Знать: основы теории математической статистики и планирования эксперимента применительно к задачам формирования достоверных опытных данных об исследуемых объектах.

Уметь: оценивать достоверность результатов измерений и контроля; аппроксимировать опытные данные математическими зависимостями и оценивать их адекватность; планировать порядок проведения эксперимента при испытании с.-х. техники; обрабатывать опытные данные на ЭВМ; строить математические модели исследуемых процессов и их физическую интерпретацию.

Владеть навыками: планирования многофакторного эксперимента и обработки опытных данных; использования прикладных программ для обработки результатов экспериментов ЭВМ.

Сформированные знания, умения и навыки составляют методическую основу экспериментальных исследований сельскохозяйственных объектов и (или) процессов в процессе испытаний. **Испытания** - экспериментальное определение количественных и (или) качественных характеристик свойств объекта испытаний как результата воздействия на него, при его функционировании, при моделировании объекта и (или) воздействий. Основой испытаний является опыт.

Опыт — это метод исследования, который воспроизводится в описанных условиях неограниченное количество раз, и даёт идентичный результат (воспроизводимый опыт). Совокупность опытов для достижения поставленных задач в исследовании представляет собой эксперимент. **Эксперимент** (от лат. *Experimentum* — проба, опыт) — метод исследования некоторого явления или процесса в управляемых условиях. Отличается от наблюдения активным взаимодействием с изучаемым объектом. Обычно эксперимент проводится для проверки гипотезы, установления причинных связей между действующими факторами и откликом.

Лабораторно-практический курс при изучении дисциплины «Планирование и организация эксперимента» состоит из 9 расчетных заданий и выполняется по индивидуальным вариантам. В индивидуальных заданиях охвачены основные разделы курса планирования эксперимента и обработки статистической информации. Задания составлены таким образом, чтобы стимулировать инициативу и самостоятельную формулировку и доказательность выводов и рекомендаций студента. При этом в каждом задании содержатся элементы исследовательской работы.

Задания формируются в зависимости от образовательной программы и уровня подготовки обучающегося. Результаты выполнения индивидуального задания оформляются в «Рабочей тетради».

Выполненные индивидуальные задания необходимо защитить для формирования навыков научной дискуссии.

1. Статистическая обработка одномерной выборки случайной величины

1.1. Цель и задачи

Цель – изучить методику статистической обработки одномерной выборки случайной величины.

Задачи – освоить основные понятия (случайная величина, выборка, характеристики случайной величины), методику формулировки и проверки статистических гипотез, изучить требования к выборке и методику проверки их соблюдения, научиться определять достоверные статистические характеристики случайной величины.

1.2. Основные понятия и определения

Случайная величина — это переменная, которая принимает в результате опыта одно значение из множества исходов, причём появление того или иного значения этой величины до её измерения нельзя точно предсказать.

Все возможные значения случайной величины называют *генеральной совокупностью*. Если мы проведем n повторных измерений случайной величины X , то есть получим n конкретных различных численных значений X_1, X_2, \dots, X_n , то этот результат эксперимента можно считать *выборкой* объема n из гипотетической генеральной совокупности результатов единичных измерений. Оценивая характеристики выборки, мы можем иметь представление о свойствах генеральной совокупности.

Репрезентативная выборка – точно (достоверно) отражает свойства генеральной совокупности. Чтобы выборка правильно отражала основные свойства, присущие генеральной совокупности, она должна быть *случайной*, т.е. все объекты генеральной совокупности должны иметь равные шансы попасть в выборку. Для этого выборки формируются с помощью специальных методик. Репрезентативная выборка должна быть *достаточной* по объему для обеспечения необходимой точности определяемых показателей (характеристик).

Достоверная выборка не должна содержать *грубые ошибки* (промахи, не характерные значения реализации случайной величины).

Доверительная вероятность – вероятность того, что значение рассчитываемых оценочных характеристик для генеральной совокупности попадет в доверительный интервал. Чем больше доверительная вероятность, тем больше должен быть доверительный интервал.

Для оценки случайной величины используют **характеристики положения и рассеивания**.

Характеристики положения: математическое ожидание, мода и медиана.

Математическое ожидание - число, вокруг которого сосредоточены значения случайной величины представляет абсциссу центра тяжести плоской фигуры, ограниченной кривой распределения и осью абсцисс. Математическое ожидание случайной величины x обозначается $M(x)$. Математическое ожидание не-

прерывной случайной величины X равно

$$m_1(X) = M(x) = \int_{-\infty}^{+\infty} x f(x) dx$$

а математическое ожидание дискретной случайной величины равно

$$m_1(X) = M(x) = \sum_{i=1}^n x_i p_i$$

Модой дискретной случайной величины, обозначаемой Mo , называется ее наиболее вероятное значение, а модой непрерывной случайной величины – значение, при котором плотность вероятности максимальна.

Медианой непрерывной случайной величины X называется такое ее значение Me , для которого одинаково вероятно, окажется ли случайная величина меньше или больше Me , т.е. $P(X < Me) = P(X > Me)$.

Характеристики рассеивания: дисперсия, среднее квадратическое отклонение и коэффициент вариации.

Дисперсия случайной величины характеризует степень разброса случайной величины около ее математического ожидания и представляет собой **математическое ожидание квадрата ее отклонения:** $D(X) = M(X - M(X))^2$. Дисперсия случайной величины как характеристика разброса имеет одну неудобную особенность: ее размерность (из определения) равна квадрату размерности случайной величины X .

Средним квадратическим отклонением случайной величины X называется арифметический корень из дисперсии, т.е. $\sigma(X) = \sqrt{D(X)}$.

Коэффициент вариации $V[X]$ - отношение стандартного отклонения $\sigma[X]$ к математическому ожиданию $M[X]$, выраженное в процентах или в долях (в расчетах).

Для оценки приведенных выше истинных характеристик случайной величины используют некоторые оценочные функции этих величин

$f(X_1, X_2, \dots, X_n)$, которые называются **статистиками** (или **характеристиками**). Значения статистик зависят от объема выборки и свойств случайной величины.

Математическое ожидание $M[X]$ оценивается выборочным средним

$$\bar{X} = \frac{1}{n}(X_1 + X_2 + \dots + X_n) = \frac{1}{n} \sum_{i=1}^n X_i \quad (1.1)$$

Дисперсия $D[X]$ оценивается выборочной дисперсией

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \quad (1.2)$$

Оценочный коэффициент вариации вычисляется по формуле

$$V = \frac{S}{\bar{x}} \times 100\% , \quad (1.3)$$

где S – оценочное значение среднего квадратического отклонения $S_x = \sqrt{S_x^2}$.

1.3. Задание

Для сформированной на компьютере выборки объемом ($N=25$) выполнить проверки соблюдения требований к выборке:

- проверить наличие грубых ошибок в выборке;
- проверить соблюдение требований о случайном характере выборки;
- проверить соблюдения требований о достаточности выборки;
- определить оценочные статистические характеристики случайной величины.

1.4. Рекомендации по выполнению задания

1. С помощью специальной программы в среде MathCAD (рис.1.1) смоделируем $N=25$ значений случайной величины. Варианты индивидуального задания задаются преподавателем или задаются параметры рассеивания условной случайной величины (табл. 1.1.)

Моделирование условной выборки

$$N := 25 \quad a := 20 \quad b := 1.7 \quad X := (\text{norm}(N, a, b))$$

Для вывода данных набрать команду X=

Рис. 1.1 Программа моделирования условной выборки

2. Найдем статистические характеристики для полученной выборки.

Выборочное среднее:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{25} \cdot 429,825 = 17,193 \quad (1.4)$$

Выборочная дисперсия:

$$s_x^2 = \frac{1}{n-1} \sum_{i=1}^n \Delta i^2 = \frac{1}{25-1} \cdot 1074,624 = 44,776; \quad (1.5)$$

Среднее квадратическое отклонение:

$$s_x = \sqrt{s_x^2} = \sqrt{44,776} = 6,692. \quad (1.6)$$

Таблица 1.1. Данные для моделирования выборки на компьютере

Вариант	Параметры		Вариант	Параметры		Вариант	Параметры	
	a	b		a	b		a	b
1	15	1	6	15	2	11	15	1.75
2	18	1.5	7	18	2.5	12	18	2
3	21	2	8	21	3	13	21	2.25
4	24	2.5	9	24	3.5	14	24	2.75
5	27	3	10	27	3	15	27	3.2

3. Исключение грубых ошибок.

Смоделированную на компьютере выборку представим в виде вариационного ряда, т.е. ряда упорядоченного по мере возрастания (колонка 3). Проанализируем наибольшее значение $X_{\max} = 29,114$ и наименьшее значение - $X_{\min} = 0,117$. Если эти значения выделяются из приведенной совокупности, то их можно считать грубыми ошибками и исключить из выборки.

Проверку этого утверждения относительно $X_{\max} = 29,114$ выполним в следующем порядке:

3.1. Сформулируем основную и альтернативную статистические гипотезы:

H_0 : X_{\max} является грубой ошибкой;

H_1 : X_{\max} не является грубой ошибки.

3.2. Выбираем статистический критерий, который представляет собой отношение удаления подозреваемой величины $X_{\max} = 29,114$ от среднего значения $\bar{x} = 17,193$ к среднему квадратическому отклонению и находим расчетное значение

$$V_p = \frac{X_{\max} - \bar{X}}{S_x} = \frac{29,114 - 17,19306}{6,692} = 1,782 \quad (1.7)$$

3.3. Определяем критическое значение статистического критерия таблица 1 (приложения1)

$$V_{кр} \equiv V_{\alpha=0.05; n=25} = 2,717.$$

3.4. Сравниваем расчетное значение с критическим: так как $V_p < V_{кр}$, то X_{\max} не является грубой ошибкой, справедлива гипотеза H_1 .

Таблица 1.2. Расчетная таблица.

№ п/п	Выборочные значения		$d_i = x_{i+1} - x_i$	d_i^2	$\Delta_i = x_i - \bar{x}$	Δ_i^2
	в порядке появления	в порядке возрастания				
1	10,723	0,117	-5,66	32,01	-6,47	41,86
2	16,381	10,353	-4,06	16,50	-0,81	0,66
3	20,4435	10,723	4,56	20,80	3,25	10,57
4	15,883	10,807	-1,80	3,25	-1,31	1,72
5	17,685	11,264	-6,09	37,08	0,49	0,24
6	23,774	11,908	2,24	5,04	6,58	43,31
7	21,53	12,276	-1,31	1,70	4,34	18,81
8	22,835	12,956	10,93	119,40	5,64	31,83
9	11,908	13,511	0,64	0,41	-5,29	27,93
10	11,264	13,854	-1,01	1,02	-5,93	35,15
11	12,276	15,883	-4,87	23,75	-4,92	24,18
12	17,149	16,381	6,34	40,22	-0,04	0,00
13	10,807	17,149	10,69	114,28	-6,39	40,78
14	0,117	17,185	-28,82	830,36	-17,08	291,59
15	28,933	17,685	3,40	11,53	11,74	137,83
16	25,538	20,443	4,54	20,58	8,35	69,64
17	21,001	21,001	8,05	64,72	3,81	14,50
18	12,956	21,53	-0,56	0,31	-4,24	17,95
19	13,511	21,944	-0,34	0,12	-3,68	13,56
20	13,854	22,662	-8,09	65,45	-3,34	11,15
21	21,944	22,835	4,76	22,65	4,75	22,57
22	17,185	23,774	-11,93	142,30	-0,01	0,00
23	29,114	25,538	18,76	351,98	11,92	142,11
24	10,353	28,933	-12,31	151,51	-6,84	46,79
25	22,662	29,662				
	$\sum x_i = 429,825$			$\sum d_i^2 = 2076,96$		$\sum \Delta_i^2 = 1044,72$

Проверяем, является ли X_{\min} грубой ошибкой.

H_0 : X_{\min} является грубой ошибкой

H_1 : X_{\min} не является грубой ошибки

$$V_p = \frac{\bar{X} - X_{\min}}{S_x} = \frac{17,19306 - 0,117}{6,692} = 2,5517$$

Так как $V_p < V_{кр}$, то X_{\min} так же не является грубой ошибкой.

4. Проверка случайности выборки.

Проверку гипотезы о случайности выборки проведем на основе метода разностей. Сформируем новую случайную величину $d_i = x_{i+1} - x_i$ как разницу между предшествующим и последующим значением (в порядке получения).

4.1. Сформулируем гипотезу:

H_0 : выборка случайна;

H_1 : выборка неслучайна.

4.2. В качестве критерия возьмем отношение двух характеристик рассеивания

$$\tau_p = \frac{C_x^2}{S_x^2}, \quad (1.8)$$

где C_x^2 – характеристика рассеивания случайной величины, подсчитанное по методу разностей.

$$C_x^2 = \frac{1}{2(n-1)} \sum_{i=1}^{24} d_i^2 = 45,151 \quad (1.9)$$

Расчетное значение критерия получим

$$\tau_p = \frac{C_x^2}{S_x^2} = 1,0084 \quad (1.10)$$

4.3. Критическое значение выбранного критерия для объема выборки $n \leq 20$ $\tau_{кр} \equiv \tau_{\alpha;n}$ находим по таблице 2 (приложения1).

Для выборки объемом больше $n > 20$ τ распределено по нормальному закону распределения с параметрами

$$m_x = 1 ; S_x = \sqrt{\frac{1}{n+1} \cdot \left(1 - \frac{1}{n-1}\right)} \approx \frac{1}{\sqrt{n+1}} \quad (1.11)$$

В этом случае $\tau_{кр}$ определяется из условия

$$\frac{\sqrt{n+1}}{\sqrt{2\pi}} \int_{-\infty}^{\tau_{кр}} \exp\left[-\frac{(\tau-1)^2 \cdot (n+1)}{2}\right] d\tau = \alpha. \quad (1.12)$$

Для $n=25$, $\alpha=0,05$ находим параметр нормирования ЗНР $t = \frac{x-m_x}{S_x}$ по таблице 10 (приложения1) соответствующий уровню доверительной вероятности 0,95, получим $t = 1,65$ ($0,5+0,45053$), а для 0,05 $t = -1,65$ ($0,5 - 0,45053$).

Искомое значение будет равно $\tau_{кр} = x = t \cdot S_x + m_x$. Значение $m_x = 1, a$ S_x подсчитываем по формуле (1.11) $S_x = \frac{1}{\sqrt{25+1}} = 0,196$. Тогда

$$\tau_{кр} \equiv \tau_{\alpha=0,05;n=25} = -1,65 \cdot 0,196 + 1 = 0,6766 \quad (1.13)$$

4.4. Вывод: так как $\tau_p < \tau_{кр}$, то выборка случайна.

5. Проверка достаточности выборки.

5.1. Рассчитаем оценочное значение среднего (1.4) для первых 3, 5, 10 и 25 значений и построим график зависимости среднего от объема выборки.

Объем выборки	3	5	10	25
$m_x = \bar{x}$				



Рис. 1.2 Зависимость среднего \bar{X} от объема выборки n

Из графика видно, что среднее зависит от анализируемой величины и от объема выборки. Для достаточной выборки должно выполняться условие

$$P(\bar{X} - I_{доо} \leq m_x \leq \bar{X} + I_{доо}) = \alpha, \quad (1.14)$$

где $I_{доо}$ – доверительный интервал для нахождения среднего значения; m_x – математическое ожидание среднего значения; α – доверительная вероятность (в расчетах принимаем $\alpha=0,95$).

Для соблюдения условия (1.14) должно выполняться условие

$$\Delta = \frac{I_{доо}}{\bar{X}} \leq \frac{t_{\gamma;n} \cdot S_x}{\bar{X} \cdot \sqrt{n}}, \quad (1.15)$$

где Δ – относительная погрешность определения среднего; $t_{\gamma;n}$ - коэффициент Стьюдента (таблица , приложения 1), $t_{\gamma=0,95;n=25} = 2,06$.

Значение относительной погрешности задается методикой испытаний или измерений (в расчетах принимаем $\Delta=10\%$ или 0,1).

5.2. Минимально необходимый объем подсчитываем из условия равенства (1.15)

$$n_p = \left(\frac{t_{\gamma;n} \cdot S_X}{\Delta \cdot \bar{X}} \right)^2 = \left(\frac{2,06 \cdot 6,692}{0,1 \cdot 17,193} \right)^2 = 64,29 \quad . \quad (1.16)$$

Полученное расчетное значение округляем до большего целого значения $n_p = 65$.

Сравниваем полученное значение минимально необходимого числа опытов с объемом выборки и делаем вывод о достаточности имеющейся выборки. В нашем примере $n_p = 65 > N = 25$, следовательно, выборка не достаточная по объему. В этом случае, необходимо провести дополнительную серию опытов.

5.3. Определение минимально необходимого числа измерений при разработке методики исследования.

В практической работе исследователя чаще всего встречается задача обоснования необходимого числа измерений при разработке рабочей методики испытаний. На этом этапе выборочных значений не имеем. Значения величин по формуле (1.16) найти не можем, так как каждое из них зависит от искомого объема выборки.

При выполнении задания необходимо самостоятельно разобраться и объяснить, как поступить в этом случае?

Контрольные вопросы:

1. Дать определение случайной величины и привести примеры из своей практики.
2. Как будут изменяться статистические характеристики случайной величины при увеличении и выборки?
3. Объяснить порядок проверки статистических гипотез.
4. Объяснить понятие статистического критерия.
5. Как сравнить два исследуемых идентичных процесса с разными средними и дисперсиями одного признака?

2. Статистическая обработка выборки большого объема

2.1. Основные понятия и определения.

По мере увеличения объема выборки увеличивается риск ввода ошибочных данных. В практической статистике выборки более 30-40 значений относят к *выборкам большого объема*. Для определения статистических характеристик выборку в этом случае преобразуем в форму представления случайной величины в виде ряда распределения.

Ряд распределения – это таблица, ставящая в соответствие значения случайной величины и вероятности их появления.

x_i	x_1	x_2	x_n
p_i	p_1	p_2	p_n

В статистике *ряд распределения* — представляет собой упорядоченное распределение единиц изучаемой совокупности по группам, разделяемым по определенному варьирующему признаку. В зависимости от признака, положенного в основу образования ряда распределения различают **атрибутивные** и **вариационные** ряды распределения.

Атрибутивными — называют ряды распределения, построенные по качественными признакам.

Сгруппированные данные в порядке возрастания или убывания значений количественного признака называются *вариационным рядом*. Количественный признак может иметь дискретные (целочисленные) значения или непрерывные значения. Непрерывные значения признака могут быть разбиты на интервалы и оценены частотой попадания выборочных значений в тот или иной интервал.

Ряд распределения характеризуется двумя элементами: вариантой (X) и частотой (f). *Варианта* – это отдельное значение признака отдельной единицы или группы совокупности. Число, показывающее, сколько раз встречается то или иное значение признака, называется *частотой*. Сумма всех частот должна быть равна численности единиц всей совокупности. Если частота выражена относительным числом, то она называется *частостью (опытной вероятностью)*.

2.2. Задание

Для анализа методики статистической обработки выборки большого объема необходимо определить по данным из 1-ого задания:

- преобразовать выборку объемом $n=25$ (взять из 1-ой работы) в ряд распределения;
- определить статистические характеристики;

- сравнить значения статистических характеристик, полученных в 1-ом и 2-ом задании.

2.3.Выполнение задания

1. Задаемся числом интервалов m и разбиваем весь диапазон изменения случайной величины от \min до \max на равные участки. Процедура не формализованная. Существует достаточно много эмпирических соотношений. В частности в ряде учебников приводится формула Стерджесса $m = 1 + 3.22 \cdot \lg N$; есть практические рекомендации: при $n < 40-50$ число интервалов выбирают в пределах $5 \dots 8$, а для большего числа – $8 \dots 12$. В нашем примере для выбора числа интервалов воспользуемся соотношением $m = \sqrt{n}$. Принимаем $m = \sqrt{25}, m = 5$.

2. Подсчитаем ширину интервала

$$d = \frac{R}{m} = \frac{X_{\max} - X_{\min}}{m} = \frac{29,114 - 0,117}{5} = 5,7994 \quad (2.1)$$

3. Определим левые и правые границы для каждого интервала и заполним расчетную таблицу (столбцы 2 и 3). Левая граница первого интервала равна $X_{\min} = 0,117$, правая граница будет равна $X_{\min} + d$. Для второго интервала левая граница будет равна правой границе первого интервала, правая граница равна $X_{\min} + 2d$. Правая граница последнего интервала будет равна X_{\max} .

Таблица 2.1. Преобразование выборки в ряд распределения

№	Граница		\bar{x}	n_i	p_i
	левая	правая			
1	0,117	5,9164	3,0167	1	0,04
2	5,9164	11,7158	8,8161	4	0,16
3	11,7158	17,5152	14,6155	9	0,36
4	15,5152	23,3146	20,4149	7	0,28
5	23,3146	29,114	26,2143	4	0,16

4. Определим середины интервалов x_i и запишем их в 4 столбец таблицы. Примем *допущение 1* – случайная величина может принимать любое значение в пределах интервала, но в расчетах их будем относить к середине интервала.

5. Подсчитаем число значений выборки, попадающие в каждый интервал n_i . Значения запишем в 5 столбец таблицы и проверим условие: $\sum_{i=1}^{m=5} n_i = N = 25$.

6. Определим опытные вероятности $p_i = \frac{n_i}{N}$ и записываем их значения в 6 столбец таблицы. Проверим условие нормирования, так как события представляют полную группу, то сумма опытных вероятностей должна равняться 1, т.е. $\sum_1^5 p_i = 1$.

Принимаем **допущение 2** – вероятность попадания случайной величины в интервал различна для левой и правой границы, а в расчетах примем постоянной или распределенной по закону равномерной плотности. С учетом принятых допущений 1 и 2 получили *ряд распределения*.

7. Определим среднее значение

$$\bar{X}_2 = \sum x_i \cdot p_i = 16,70328 \quad (2.2)$$

8. Определим дисперсию

$$S_{x2}^2 = \sum_1^5 (x_i - \bar{x})^2 \cdot p_i = 37,586 \quad (2.3)$$

9. Сравним значения средних значений, полученных в первом и втором задании

$$\Delta_{\bar{x}_1} = \frac{\bar{x}_1 - \bar{x}_2}{\bar{x}_1} \cdot 100 = \frac{17,193 - 16,703}{17,193} = 2,848\% \quad (2.4)$$

10. Сравним значения дисперсий, полученных в первом и втором задании

$$\Delta_{S_x^2} = \frac{S_1^2 - S_2^2}{S_1^2} \cdot 100 = \frac{44,776 - 37,586}{44,776} \cdot 100 = 16,05\% \quad (2.5)$$

Анализ полученных данных показывает удовлетворительную разницу между средними значениями 2,85%, а для дисперсии разницу существенная – 16,05%. В этом случае необходимо проводить дополнительные опыты. Если провести дополнительные опыты невозможно, используют поправку Шепарда $\Delta_{S_x^2} = \frac{d^2}{12}$. Для приближения оценочной дисперсии S_x^2 к истинному значению при ограниченном числе опытных данных к дисперсии прибавляют поправку Шепарда.

$$s_{x2}^2 = s_{x2}^2 + \frac{d^2}{12} \quad (2.6)$$

Для уменьшения относительной погрешности можно попробовать увеличить число интервалов и (или) провести дополнительные опыты.

11. Опытные данные в графическом виде представляют в виде гистограммы и опытной функцией плотности распределения (*многоугольник распределения, полигон распределения*) (рис. 2.1).

Гистограмма - столбчатая диаграмма, высота столбиков которой соответствует частоте или относительной частоте (частости, опытной вероятности) попадания данных в каждый из интервалов.

Для построения гистограммы на горизонтальной оси в выбранном масштабе отметим границы соответствующих интервалов. На вертикальной оси выбирают масштаб в соответствии с максимальным значением опытной вероятности.

Строят столбчатую диаграмму, затем середины столбчатой диаграммы соединяют отрезками ломаной прямой и получают полигон распределения (рис. 2.1). Полигон представляет собой опытную функцию плотности распределения.

12. Опытная функция распределения (*кумулята*) показана на рис. 2.2. График накопленных относительных частот представляет собой опытную функцию распределения. Для его построения по горизонтальной оси, как и для гистограммы, отмечают границы интервалов, а по вертикальной оси (ординате) в интервале (0,1) откладывают накопленные частоты из условия: $F^*(x) = P(X < x) = \alpha$, где x – заданное значение. Так задаваясь $x=5,9$, вероятность будет равна $p_1=0,04$; при

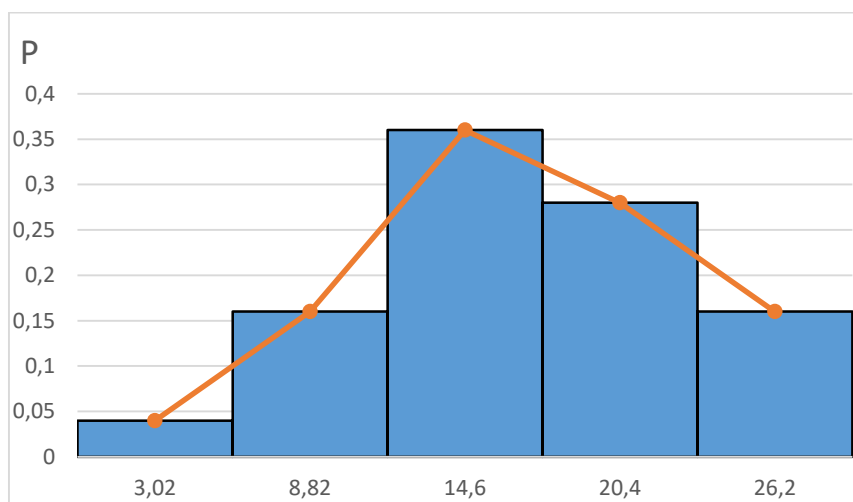


Рис. 2.1. Гистограмма распределения и опытная функция плотности распределения

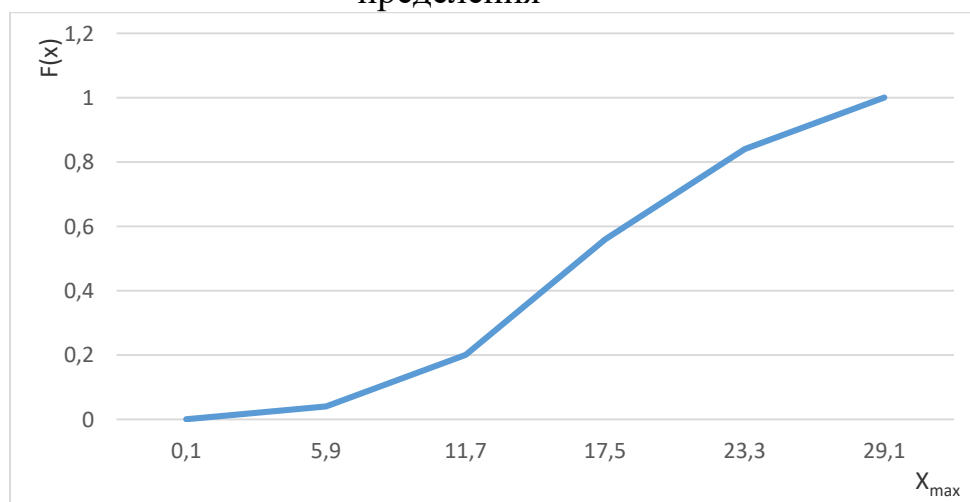


Рис. 2.2. Опытная функция распределения (кумулята)

$x=11,7$, ордината будет равна p_1+p_2 , так для всех других значений границ интервалов.

Контрольные вопросы:

1. В какой форме представлена случайная величина?
2. Какие причины могут вызвать большую относительную погрешность определения оценочной дисперсии?
3. Как будут изменяться частоты (опытные вероятности) при увеличении числа измерений?
4. Как будут изменяться частоты (опытные вероятности) при увеличении числа интервалов?
5. Как будет изменяться полигон распределения при увеличении числа измерений?

3. Выбор закона распределения случайной величины

3.1. Цель и задачи

Цель – изучить методику выбора закона распределения для описания одномерной выборки случайной величины.

Задачи – освоить основные понятия (закон распределения, функция распределения, функция плотности распределения, вероятность согласия), методику выбора закона распределения, изучить свойства функции распределения, научиться определять критические значения критериев согласия Пирсона и Колмогорова и вероятность согласия опытных данных определенному закону распределения.

3.2. Основные понятия и определения

Закон распределения – математическая зависимость, связывающая значения случайной величины и вероятности ее определения $P = f(x_i)$. По характеру этой зависимости классифицируют теоретические законы распределения (закон нормального распределения; логарифмически нормальный, Вейбулла и другие).

Закон распределения используется для сглаживания статистических данных. Задача сглаживания заключается в том, чтобы подобрать теоретическую плавную кривую распределения, с той или иной точки зрения наилучшим образом описывающую данное статистическое распределение. Как правило, принципиальный вид теоретической кривой выбирается заранее из соображений, связанных с существом задачи, а в некоторых случаях просто с внешним видом статистического распределения.

Функцией распределения – называется функция $F(X)$, описывающая вероятности появления значений случайной величины $P(X \leq x)$ меньше наперед заданного числа равной α .

$$F(X) = P(X \leq x) = \alpha. \quad (3.1)$$

Функция распределения $F_X(x)$ любой случайной величины обладает следующими свойствами:

1. F_X — функция неубывающая;

2. $\lim_{x \rightarrow -\infty} F_X(x) = 0, \lim_{x \rightarrow \infty} F_X(x) = 1;$
3. F_X непрерывна справа.

Функция плотности распределения случайной величины $f(x)$ представляет собой первую производную функции распределения $F(X)$

$$f(x) = \frac{dF(X)}{dx} . \quad (3.2)$$

Близость опытных данных к функции распределения оценивается вероятностью согласия.

3.3. Задание:

Для имеющихся данных (задание 2, таблица 2.1)

- проверить гипотезу о возможности использования закона нормального распределения (ЗНР) для сглаживания опытных данных по критерию Пирсона (χ^2);
- проверить гипотезу о возможности использования закона нормального распределения для сглаживания опытных данных по критерию Колмогорова (λ);
- определить по одному из критериев вероятность согласия опытных данных и теоретической функции распределения с опытными параметрами.

3.4. Порядок выполнения

1. Для ряда распределения формулируем гипотезу:

H_0 : случайная величина X подчинена закону нормального распределения)

H_1 : случайная величина X не подчинена закону нормального распределения.

2. Проверяем гипотезу по критерию Пирсона (χ^2). Из таблицы 2.1 перенести значения середин интервалов X_i и опытные вероятности p_i в таблицу 3.1. Накладываем теоретическую функцию плотности распределения на полигон распределения. Для этого вычислим параметр нормирования t_i (3.3) для нормированного закона нормального распределения и запишем в столбец 4 таблицы 3.1. Параметры закона распределения $a = \bar{x}$, $b = S_x$ (среднее \bar{x} и среднее квадратическое S_x берутся из задания 1)

$$t_i = \frac{x-a}{b} = \frac{x-\bar{x}}{S_x}; \quad (3.3)$$

Определим значения теоретической функции плотности распределения для закона нормального распределения для середин интервалов $f(x_i)$ по значению t_i (таблица 9, приложение 1).

Подсчитаем расчетное значение критерия Пирсона

$$\chi_p^2 = \sum \frac{(p^* - f(x_i))^2}{f(x_i)} = 0,02 .$$

Расчетное значение критерия Пирсона является мерой расхождения теоретической функции плотности распределения для нормального закона распределения с опытными параметрами и опытной функцией распределения (полигоном распределения).

Таблица 3.1. Расчетная таблица для критерия Пирсона

№	X_i	P_i	t_i	$f(x_i)$	$\frac{(p^* - f(x_i))^2}{f(x_i)}$
1	3,0167	0,04	-2,118	0,04217	0,0001
2	8,8161	0,16	-1,252	0,18266	0,0028
3	14,6155	0,36	-0,385	0,37115	0,0003
4	20,4149	0,28	0,481	0,35553	0,0160
5	26,2143	0,16	1,35	0,16039	0,0004

Найдем критическое значение критерия Пирсона по таблице 4 (приложение 1) для уровня доверительной вероятности 0,95 (ошибка 0,05) и $k=m-1=5-1=4$.

Сравниваем расчетное значение критерия Пирсона с критическим значением и делаем вывод. Если $\chi_p^2 \leq \chi_{кр}^2$, то принимаем гипотезу о том, что случайная величина подчинена нормальному закону распределения, в противном случае отвергаем эту гипотезу.

3. Проверка гипотезы по критерию Колмогорова А.Н. (λ_p). Накладываем теоретическую функцию распределения с опытными параметрами на опытную функцию (рис. 2.2). Для этого подсчитываем значения опытной функции распределения, соответствующее серединам интервалов $F^*(x_i)$

$$F^*(x_1) = \frac{1}{2}p_1 ;$$

$$F^*(x_2) = p_1 + \frac{1}{2}p_2 ;$$

$$F^*(x_3) = p_1 + p_2 + \frac{1}{2}p_3 ;$$

$$F^*(x_4) = p_1 + p_2 + p_3 + \frac{1}{2}p_4 ;$$

$$F^*(x_5) = p_1 + p_2 + p_3 + p_4 + \frac{1}{2}p_5 .$$

Полученные значения запишем в столбец 5 таблицы 3.2.

Определим значения теоретической функции распределения нормального закона распределения $F(x_i)$ и запишем в столбец 6 таблицы 3.2. Для этого по значениям t_i находим значения $\Phi(t_i)$ по таблице 10 (приложение 1). Необходимо учитывать особенность приведенной таблицы: при $t < 0$; $F(x) = 0,5 - \Phi(t)$, при $t > 0$; $F(x) = 0,5 + \Phi(t)$.

Таблица 3.2 Расчетная таблица для критерия Колмогорова

№	X_i	P_i	t_i	$F^*(x_i)$	$F^t(x_i)$	$D_i = F^* - F^t $
---	-------	-------	-------	------------	------------	---------------------

1	3,0167	0,04	-2,118	0,02	0,0174	0,0026
2	8,8161	0,16	-1,252	0,12	0,1057	0,0143
3	14,6155	0,36	-0,385	0,38	0,3483	0,0317
4	20,4149	0,28	0,481	0,7	0,6844	0,0156
5	26,2143	0,16	1,35	0,92	0,9115	0,0085

Находим разницу между опытными и теоретическими значениями функции распределения по абсолютной величине

$$D_i = |F^*(x_i) - F(x_i)| \quad (3.4)$$

Выбираем максимальное значение D_i и определяем расчетное значение критерия Колмогорова А.Н.

$$\lambda_p = D_{i \max} \cdot \sqrt{m} = 0,0317 \cdot \sqrt{5} = 0,07. \quad (3.5)$$

По таблице 5 (приложение 1) находим $\lambda_{кр} = 0,52$ или вероятность согласия $P(\chi_p^2)=1$. Так как $\lambda_p < \lambda_{кр}$, следовательно, расхождение между теоретическими и опытными значениями не превышает допустимого значения значит принимаем гипотезу о том, что случайная величина подчинена нормальному закону распределения.

Контрольные вопросы:

1. Что оценивает расчетное значение критерия Пирсона?
2. Что оценивает расчетное значение критерия Колмогорова?
3. С какой вероятностью делается вывод о справедливости проверяемой гипотезы?
4. Как определить вероятность согласия опытных данных и теоретического закона распределения?
5. Как выбрать лучший закон распределения для имеющихся опытных данных?

4. Оценка статистической взаимосвязи двух случайных величин

5.1. Цель и задачи

Цель – изучить методику статистической оценки системы случайных величин.

Задачи – освоить основные понятия (система случайных величин, эллипс рассеивания, ковариация и корреляционный момент, коэффициент парной и множественной корреляции), методику оценки статистической взаимосвязи двух случайных величин, изучить виды взаимосвязи случайных величин, научиться определять коэффициент парной корреляции и оценивать его значимость.

5.2. Основные понятия и определения

В экспериментальных исследованиях, связанных с сельскохозяйственными объектами (растениями, животными, технологическими процессами) применительно к одному и тому же объекту рассматривают две и более связей.

Связь – совокупность зависимостей свойств одного элемента от свойств других элементов системы. Установить связь между двумя элементами – это значит выявить наличие зависимостей их свойств. Связи, взаимодействующие на основе законов природы, называют **функциональными** $y = f(x)$. Для них характерно для любого x однозначное соотношение другой переменной y .

На практике чаще всего встречаются такие соотношения между переменными, когда каждому значению признака x соответствует ни одно, а множество возможных значений признака y . Такие связи в отличие от функциональных связей, называются **стохастическими (вероятностными) или корреляционными**. При изучении таких связей возникают два основных вопроса – о тесноте связи и форме связи.

Для измерения тесноты связи используют специальный статистический метод, называемый **корреляцией**.

Для анализа линейной корреляции между x и y необходимо иметь n – независимых парных наблюдений, исходом каждого из которых является пара чисел (x_1, y_1) , (x_2, y_2) , ..., (x_n, y_n) . Эти наблюдения могут быть получены сбором статистических данных (*пассивный эксперимент*) или при измерении значений y для заданных в определенном порядке x (*активный эксперимент*). Порядок задания значений x называют **планом эксперимента**.

Для каждой переменной определяют статистические характеристики (см. задание 1), а для оценки взаимосвязи определяют **корреляционный момент** $K_{x,y}$, который является статистической оценкой **ковариации** $Cov(x,y)$

$$K_{x,y} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y}) \quad . \quad (4.1)$$

Для функциональной зависимости $K_{x,y} = S_x^2 = S_y^2$, а при отсутствии зависимости $K_{x,y} = 0$ (свойство среднего).

В качестве числового показателя простой линейной корреляции, указывающего на тесноту (силы) и направление связи x с y , используют коэффициент парной корреляции $r_{x,y}$.

Это безразмерная величина, изменяющаяся от -1 до +1, при этом $r_{x,y}=1$ для функциональной зависимости и $r_{x,y}=0$ при отсутствии зависимости. Значение коэффициента парной корреляции рассчитывают по формуле

$$r_{x,y} = \frac{K_{x,y}}{S_x \cdot S_y} . \quad (4.2)$$

Для оценки наличия взаимосвязи оценивают значимость коэффициента парной корреляции.

При наличии связи находят уравнение большой оси эллипса рассеивания

$$\hat{y} = \bar{y} + r_{x,y} \cdot \frac{S_y}{S_x} (x - \bar{x}) . \quad (4.3)$$

Данное уравнение легко приводится к каноническому виду прямой $y = a + bx$.

Если анализируется наличие взаимосвязи более двух случайных характеристик (множественная корреляция), то подсчитывают симметричную корреляционную матрицу с диагональными элементами равными 1.

5.3. Задание

1. Для заданного типа сельскохозяйственных машин разных марок и производителей по справочным данным выбрать два показателя (4...6 значений) и оценить их взаимосвязь.

1. Оценить значимость взаимосвязи между выбранными параметрами однотипных машин.

2. Для области рассеивания опытных данных получить уравнение большой оси эллипса рассеивания и привести его к каноническому виду.

4. Выбрать масштаб осей и нанести опытные данные на координатное поле и в пределах поля рассеивания точек нарисовать полученное уравнение.

5.4. Порядок выполнения

Исходными данными для задания представляют пары чисел, относящиеся к одному объекту или к однотипным объектам, если задачами исследования необходимо их оценить и найти зависимость между оценочными характеристиками. Такие данные можно получить измерением одной из величин для заданных значений (активный эксперимент) или выбрать из каталогов или справочников (пассивный эксперимент).

1. Проанализируем значения ширины $V_k(x_i)$ захвата (м) и мощности $N_T(y_i)$ (кВт), необходимой для выполнения технологического процесса (таблица 4.1).

Таблица 4.1. Исходные данные для анализа

№	X_i	Y_i	$(X_i - \bar{X})$	$(X_i - \bar{X})^2$	$(Y_i - \bar{Y})$	$(Y_i - \bar{Y})^2$
1	3	66	-1,286	1,653	-28,857	832,73

2	3.5	77	- 0,786	0,618	- 17,857	318,87
3	4	88	- 0,286	0,082	- 6,857	47,018
4	4.5	103	0,214	0,046	8,143	66,308
5	4	88	- 0,286	0,082	- 6,857	47,018
6	5	110	0,714	0,510	15,143	229,310
7	6	132	1,714	2,938	37,143	1379,60

2. Находим среднее значение x , y .

$$\bar{x} = \frac{1}{n} \sum x_i = \frac{30}{7} = 4,286 \quad (4.4)$$

$$\bar{y} = \frac{1}{n} \sum y_i = \frac{664}{7} = 94,857 \quad (4.5)$$

3. Определяем дисперсии X и Y

$$S_x^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2 = \frac{5,929}{6} = 0,988 \quad (4.6)$$

$$S_y^2 = \frac{1}{n-1} \sum (y_i - \bar{y})^2 = \frac{2920,857}{6} = 486,81 \quad (4.7)$$

4. Определяем средние квадратические отклонения X и Y

$$S_x = \sqrt{0,988} = 0,994 \quad (4.8)$$

$$S_y = \sqrt{486,81} = 22,064 \quad (4.9)$$

5. Корреляционный момент

$$K_{xy} = \frac{1}{n-1} \sum (x_i - \bar{x})(y_i - \bar{y}) = \frac{131,286}{7-1} = 21,881 \quad (4.10)$$

6. Коэффициент парной корреляции

$$R_{xy} = \frac{K_{xy}}{S_x * S_y} = \frac{21,881}{0,994 * 22,064} = 0,998 \quad (4.11)$$

7. Оценим значимость коэффициента парной корреляции. Для этого расчетное значение R_{xy} сравним с критическим значением $R_{кр} = 0,755$ (для $k=n-2$ и доверительной ошибки равной 0,05) (таблица 4.2). Так как $R_{xy} > R_{кр}$, следовательно между шириной захвата $B_k(x_i)$ и мощностью $N_T(y_i)$ есть статистическая взаимосвязь или эти величины взаимокоррелированы.

8. Найдем уравнение большой оси эллипса рассеивания по формуле (4.3)

$$\hat{y} = \bar{y} + r_{x,y} \cdot \frac{S_y}{S_x} (x - \bar{x}) = 94,857 + 0,998 \cdot \frac{22,064}{0,994} (x - 4,286) = -0,048 + 22,14x.$$

Таблица 4.2. Критические значения коэффициента парной корреляции

K=n-2	Доверительная ошибка α			K=n-2	Доверительная ошибка α		
	0.10	0.05	0.01		0.10	0.05	0.01
1	0.988	0.997	0.999	6	0.621	0.707	0.834
2	0.900	0.950	0.990	7	0.582	0.666	0.798
3	0.805	0.878	0.959	8	0.549	0.632	0.765
4	0.729	0.811	0.971	9	0.521	0.602	0.735
5	0.669	0.755	0.875	10	0.497	0.576	0.708

9. Построим график. Для этого выбираем масштаб координатных осей, затем в выбранном масштабе наносим координаты опытных точек.

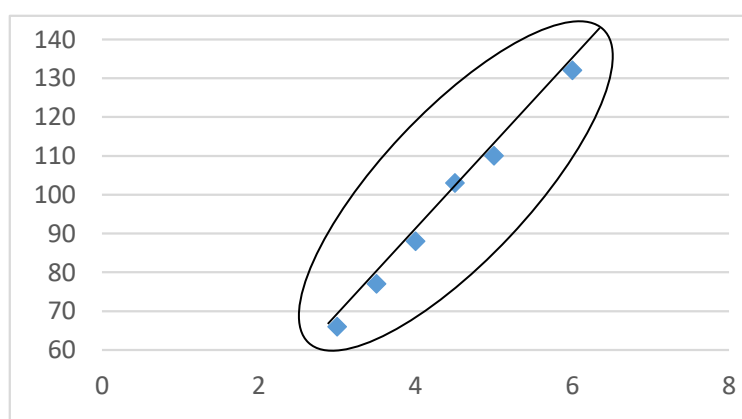


Рис. 4.1. Эллипс рассеивания опытных данных

Для функциональной зависимости эллипс рассеивания сходится к линии, а при отсутствии корреляционной связи в окружность.

Контрольные вопросы:

1. Какие могут быть виды вероятностной взаимосвязи между двумя случайными характеристиками?
2. Может ли оценить наличие значимой взаимосвязи коэффициент парной корреляции для нелинейных зависимостей?
3. Как будет изменяться коэффициент парной корреляции при увеличении объема выборки?
4. С какой достоверностью принимается или отвергается гипотеза о наличии вероятностной (стохастической) взаимосвязи между двумя случайными величинами?

5. Аппроксимация опытных данных математическими зависимостями по методу наименьших квадратов

5.1. Цель и задачи

Цель – изучить методику аппроксимации опытных данных математическими зависимостями.

Задачи – освоить основные понятия (аппроксимация, линеаризация, дисперсия адекватности, регрессия) и методику аппроксимации опытных данных математическими зависимостями и оценки их достоверности, изучить метод наименьших квадратов, научиться определять эмпирические коэффициенты и оценивать точность математических уравнений на основе дисперсионного анализа.

5.2. Основные понятия и определения

Корреляционный анализ позволяет оценить наличие линейной стохастической (вероятностной) взаимосвязи и получить одно из возможных линейных уравнений. Для анализа и описания нелинейных зависимостей используют методы регрессионного анализа.

Регрессионный анализ — это статистический метод исследования зависимости случайной величины y (*показатель, отклик*) от переменных (*аргументов, действующих факторов*) x_j ($j = 1, 2, \dots, k$) случайных величин, оцениваемых средними значениями, независимо от истинного закона распределения x_j . Методику применения регрессионного анализа рассмотрим для одномерной зависимости, т.е. для $j=1$. В этом случае уравнение регрессии будет иметь вид функции $y = f(x)$.

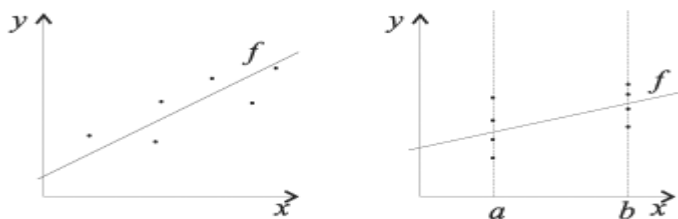


Рис.5.1. Вид функции отклика $y = f(x)$.

Для анализа о характере зависимости между x и y , как и в предыдущем задании, необходимо иметь n – независимых парных наблюдений, исходом каждого из которых является пара чисел (x_1, y_1) , (x_2, y_2) , ..., (x_n, y_n) .

Задача регрессионного анализа заключается в поиске такой функции f , которая описывает эту зависимость. Регрессия может быть представлена в виде суммы неслучайной и случайной составляющих

$$y = f(x) + v, \quad (5.1)$$

где f — функция регрессионной зависимости, а v — аддитивная случайная величина с нулевым математическим ожиданием. Обычно предполагается, что

величина v имеет гауссовское распределение (подчиненное нормальному закону) с нулевым средним $m_x=0$ и дисперсией $\sigma^2(v)$.

Вид функции $y = f(x)$ может быть линейным или нелинейным.

Линейная регрессия – это такая зависимость, когда при любом значении аргумента x одинаковые приращения его вызывают одинаковые изменения функции y

$$\hat{y} = a + b \cdot x + v. \quad (5.2)$$

Если при одинаковых приращениях аргумента функция имеет неодинаковые изменения, регрессия называется криволинейной, в простейшем случае – парабола

$$\hat{y} = a + b \cdot x + c \cdot x^2 + v. \quad (5.3)$$

Для нахождения лучших эмпирических коэффициентов a и b функция $F(x)$ **эффективности аппроксимации опытных данных** зависимостью (5.2) в виде суммы квадратов отклонений опытных значений y_i и \hat{y}_i – подсчитанных по уравнению **должна быть наименьшей**

$$F(x) = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - a - b \cdot x_i)^2 \Rightarrow \min. \quad (5.4)$$

Для нахождения минимума функции $F(x)$ двух переменных a и b возьмем две частные производные и приравняем их к нулю, получим систему уравнений

$$\left. \begin{aligned} \frac{dF(x)}{da} &= 2 \sum_{i=1}^n (y_i - a - b \cdot x_i)^1 \cdot (-1) = 0 \\ \frac{dF(x)}{db} &= 2 \sum_{i=1}^n (y_i - a - b \cdot x_i)^1 \cdot (-x_i) = 0 \end{aligned} \right\} \quad (5.5)$$

Преобразуем полученную систему уравнений

$$\left. \begin{aligned} a \cdot n + b \sum x_i &= \sum y_i \\ a \cdot \sum x_i + b \sum x_i^2 &= \sum y_i \cdot x_i \end{aligned} \right\} \quad (5.6)$$

Решая систему (5.6) по методу Крамера, получим

$$a = \frac{\sum y_i \cdot \sum x_i^2 - \sum x_i \cdot \sum y_i x_i}{n \cdot \sum x_i^2 - (\sum x_i)^2}; \quad (5.7)$$

$$b = \frac{n \cdot \sum y_i x_i - \sum x_i \cdot \sum y_i}{n \cdot \sum x_i^2 - (\sum x_i)^2}. \quad (5.8)$$

Функция (5.3) выражает суть **метода наименьших квадратов** (МНК) и является базовым методом регрессионного анализа для оценки неизвестных параметров регрессионных моделей по выборочным данным.

Важным моментом при аппроксимации является обоснование **достаточной точности** описания опытных данных той или иной зависимостью. Для этого рассмотрим **дисперсию адекватности**, оценивающую рассеивание опытных данных относительно уравнения (5.2)

$$S_{ad1}^2 = \frac{1}{f_{ad}} \cdot \sum_{i=1}^n (y_i - \hat{y}_i)^2, \quad (5.9)$$

где $f_{ad} = N - d - 1$ - число степеней свободы, а d количество независимых связей (для прямой $d=1$).

Повысим степень уравнения и аппроксимируем наши данные уравнением второго порядка (5.3) и определим S_{ad2}^2 . Мы получим две дисперсии, оценивающие рассеивание опытных данных для прямой S_{ad1}^2 и параболы S_{ad2}^2 , которые можем сравнить по критерию Фишера (F критерий), представляющий собой функцию плотности распределения отношения двух дисперсий $F = \frac{S_1^2}{S_2^2}$, где $S_1^2 > S_2^2$ со степенями свободы f_1 и f_2 .

5.3. Задание

1. Для имеющихся данных однотипных сельскохозяйственных машин разных марок и производителей (исходные данные задания 4) найти коэффициенты уравнения (5.2) и дисперсию адекватности (5.9).

2. Для тех же данных с помощью компьютерной программы («Регрессия») найти коэффициенты уравнения параболы (5.3) и дисперсию адекватности S_{ad2}^2 по формуле (5.9).

3. Сравнить полученные дисперсии и сделать вывод о виде зависимости, рекомендуемой для аппроксимации опытных данных.

5.4. Порядок выполнения задания

1. Для исходных данных подсчитаем значения коэффициентов системы уравнений (5.6) для определения коэффициентов a и b уравнения прямой по методу наименьших квадратов (таблица 5.1)

Таблица 5.1. Исходные данные для анализа

№	X_i	Y_i	$X_i \cdot Y_i$	X_i^2
1	3	66	198	9
2	3.5	77	269,5	12,25
3	4	88	352	16
4	4.5	103	463,5	20,25
5	4	88	352	16
6	5	110	550	25
7	6	132	792	36
	$\sum x_i = 30$	$\sum y_i = 664$	$\sum x_i y_i = 2977$	$\sum x_i^2 = 134,5$

2. Составим систему уравнений (5.6)

$$\begin{aligned} a \cdot 7 + 30b &= 664 \\ a \cdot 30 + 134,5b &= 2977 \end{aligned}$$

3. Определим значения коэффициентов по формулам (5.7 и 5.8)

$$a = \frac{\sum y_i \cdot \sum x_i^2 - \sum x_i \cdot \sum y_i x_i}{n \cdot \sum x_i^2 - (\sum x_i)^2} = \frac{664 \cdot 134,5 - 30 \cdot 2977}{7 \cdot 134,5 - (30)^2} = 0,048$$

$$b = \frac{n \cdot \sum y_i x_i - \sum x_i \cdot \sum y_i}{n \cdot \sum x_i^2 - (\sum x_i)^2} = \frac{7 \cdot 2977 - 30 \cdot 664}{7 \cdot 134,5 - (30)^2} = 22,14$$

Получим уравнение прямой $\hat{y} = 0,048 + 22,14 \cdot x$.

4. Подсчитаем дисперсию адекватности для прямой по формуле (5.9)

$$S_{ad1}^2 = \frac{1}{f_{ad}} \cdot \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \frac{1}{5} \cdot 13,968 = 2,794$$

Число степеней свободы равно $f_{ad} = N - d - 1 = 7 - 1 - 1 = 5$.

Таблица 5.2. Определение дисперсии адекватности для прямой

№	X_i	Y_i	\hat{y}_i	$(y_i - \hat{y}_i)$	$(y_i - \hat{y}_i)^2$
1	3	66	66,47	0,47	0,221
2	3.5	77	77,54	0,54	0,292
3	4	88	88,61	0,61	0,372
4	4.5	103	99,63	3,37	11,357
5	4	88	88,61	0,61	0,372
6	5	110	110,75	0,75	0,562
7	6	132	132,89	0,89	0,792
					$\sum (y_i - \hat{y}_i)^2 = 13,968$

5. Определяем коэффициенты уравнения параболы для данных таблицы 5.1. по программе «Полиномиальная регрессия» (рис. 5.2).

Полиномиальная регрессия

Ввод данных :

$X := (32 \ 36 \ 37 \ 42 \ 43 \ 51)^T$ $Y := (1.5 \ 2.7 \ 2.7 \ 3 \ 3.5 \ 4.2)^T$ rows(X) = 6

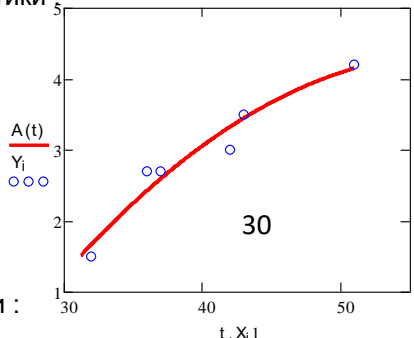
Расчет коэффициентов регрессии: $S := \frac{\sum (Y_i - f1(x_i))^2}{n - 3}$

$S := \text{regress}(X, Y, 2)$ t := 31.250, 31.26.. 51 n := rows(X)

$A(t) := \text{interp}(S, X, Y, t)$ i := 0.. n - 1

Статистические характеристики :

var(X) = 37.139
var(Y) = 0.682
corr(X, Y) = 0.956
x_i := X_i



Коэффициенты

$$S = \begin{pmatrix} 3 \\ 3 \\ 2 \\ -8.823 \\ 0.452 \\ -3.869 \times 10^{-3} \end{pmatrix}$$

Расчет дисперсии адекватности :

$$f(x) := S_3 + S_4 \cdot x \quad \text{sad} = 0.075$$

$$f1(x) := S_3 + S_4 \cdot x + S_5 \cdot x^2$$

Рис. 5.2. Программа «Полиномиальная регрессия»

Запускаем программу «Полиномиальная регрессия» в математическом пакете MathCAD. Для этого курсором нажимаем кнопку «файл» и в меню выбираем нужную программу.

Вводим данные:

- нажимаем курсором на вектор X , выделяем значения (кроме знака «Т») вектора и удаляем;
- нажимаем курсором на иконку «матрица» и задаем число столбцов в транспонированной матрице равной числу пар данных и число строк равное 1;
- в образовавшийся макет вводим данные (в качестве разделителя дробной части ставит точку);
- вводим аналогично значения вектора Y ;
- выписываем значения коэффициентов уравнения параболы, дисперсию адекватности

$$\hat{y} = a + b \cdot x + c \cdot x^2 = -12,327 + 27,831 x - 0,629 x^2;$$

$$S_{a02}^2 = 2,718 .$$

6. Подсчитываем расчетное значение критерия Фишера для

$$S_{\hat{a}d1}^2 > S_{\hat{a}d2}^2 \quad F_p = \frac{S_{\hat{a}d1}^2}{S_{\hat{a}d2}^2} = \frac{2,813}{2,718} = 1,035$$

7. Находим критическое значение критерия Фишера для $\alpha=0,05$, $k_1=5$ и $k_2=4$ (таблица 7 приложения 1)

$$F_{кр} \equiv F_{\alpha=0,05; k_1=5; k_2=4} = 5,192 .$$

8. Сравниваем расчетное и критическое значения и делаем вывод о равенстве дисперсий. Если $F_{кр} < F_p$, то дисперсии однородны, не отличаются друг от друга, поэтому можно и прямой и параболой описать опытные данные. В противном случае дисперсии различаются существенно и выбираем уравнение, имеющее меньшую дисперсию адекватности.

5.5. Функции, приводимые к линейному виду (справочно)

Метод наименьших квадратов не может быть применим для нахождения эмпирических коэффициентов нелинейной зависимости (кроме степенных). Для таких функций подбирают порядок преобразования координат так, чтобы геометрический образ принял линейную форму (проводят *линеаризацию зависимости*).

Для вычисления параметров функции $Y=ax^b$ необходимо предварительно ее прологарифмировать $\ln Y = \ln ax^b = \ln a + b \ln x$. После чего замена $Z = \ln Y$, $X = \ln x$, $A = \ln a$ приводит заданную функцию к линейному виду $Z = bX + A$, где коэффициенты A и b вычисляются по формулам (5.7 и 5.8).

Аналогично можно подобрать параметры функции вида $Y=ae^{bx}$. Прологарифмируем заданную функцию $\ln y=\ln a+bx$, $\ln y=\ln a+bx$. Проведём замену $Y=\ln y$, $A=\ln a$ и получим линейную зависимость $Y=bx+A$.

Для подбора параметров функции (гипербола) $y = \frac{1}{ax+b}$ сделаем замену $Z=1/Y$. В результате получим линейную зависимость $Z=ax+b$.

Аналогичными приемами (логарифмированием, заменами и т. п.) можно многие подбираемые зависимости преобразовать к такому виду, что получаемая при решении задачи оптимизации (5.4) была системой линейных алгебраических уравнений.

Контрольные вопросы:

1. В чем разница корреляционного и регрессионного уравнения?
2. В чем суть метода наименьших квадратов?
3. Если подобрать значения фактора так, что сумма будет равна нулю $\sum x_i = 0$ как изменятся формулы для определения коэффициентов?
4. Что оценивает дисперсия адекватности?
5. Как поступить, если расчетное значения критерия Фишера будет меньше 1?

6. Планирование полного факторного эксперимента (ПФЭ) 2ⁿ.

6.1. Цель и задачи

Цель – изучить методику планирования многофакторного эксперимента для получения линейной формы.

Задачи – освоить основные понятия планирования многофакторного эксперимента и методику дисперсионного анализа полученных результатов, метод наименьших квадратов для ортогональных матриц, научиться планировать и проводить многофакторный эксперимент и оценивать адекватность полиномиального уравнения и значимость его коэффициентов.

6.2. Основные понятия и определения

Многофакторные процессы описываются зависимостью случайной величины y (*показатель, отклик*) от переменных (*аргументов, действующих факторов*) x_j ($j = 1, 2, \dots, k$) случайных величин, оцениваемых средними значениями, независимо от истинного закона распределения x_j .

$$\hat{y} = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_j x_j + v, \quad (6.1)$$

где $\alpha_0, \alpha_1, \dots, \alpha_j$ - коэффициенты уравнения регрессии.

Факторы в физической системе имеют размерность и изменяются в диапазоне (x_{nj}, x_{oj}). Описание искомой зависимости для области определения факторов осуществляют в нормированной системе координат

$$x_j^o = \frac{x_j - x_{oj}}{J_j}, \quad (6.2)$$

где x_j - текущее значение, x_{oj} - основной уровень и J_j - интервал варьирования фактора в физической системе.

Проведение эксперимента предполагает одновременное изменение всех действующих факторов в соответствии с блоком планирования План-матрицы эксперимента (Таблица 6.2).

Для повышения точности оценочных значений коэффициентов уравнения регрессии (6.1) проводят параллельные (*воспроизводимые*) опыты (*повторности*).

Оценки коэффициентов регрессионного уравнения определяют по средним значениям показателя y , полученным в ходе эксперимента **методом наименьших квадратов** (МНК).

6.3. Задание

С помощью имитационной модели для условного 3 факторного пространства получить линейное уравнение регрессии $\hat{y} = f(x_1; x_2; x_3)$, для этого:

- задаться диапазоном варьирования каждого фактора;
- построить план-матрицу полного факторного эксперимента (ПФЭ 2³;

- провести имитационное моделирование в соответствии с блоком планирования план-матрицы и выбранными уровнями варьирования;
- провести дисперсионный анализ полученных результатов;
- получить уравнение регрессии и оценить его адекватность;
- оценить значимость коэффициентов уравнения.

6.4. Порядок выполнения задания

1. Выбираем уровни варьирования действующих факторов. Верхний и нижний уровень варьирования каждого действующего фактора задаются произвольно из допустимого диапазона (0,100) для условной физической системы и записывают в таблицу 6.1

Таблица 6.1 Уровни варьирования действующих факторов

Уровни	X_1	X_2	X_3
1.Нижний	7	13	24
2.Верхний	33	37	48
3.Основной	20	25	36
4. Интервал варьирования	13	12	12

2. Строим план-матрицу полного факторного эксперимента 2^3 . Область нормированного факторного пространства представляет собой куб и для проведения полной серии опытов потребуется 8 опытов. Координаты вершин куба указываются в блоке планирования. Для повышения точности оценочных коэффициентов уравнения регрессии предлагается проведение 3 повторностей (параллельных опытов). Добавляем столбцы для записи среднего значения отклика \bar{y}_j , дисперсии для каждого опыта S_j^2 и расчетные значения \hat{y}_j , подсчитанные по уравнению регрессии.

Таблица 6.2 План-матрица ПФЭ 2^3

№ опыта	X_0	X_1	X_2	X_3	Повторности			\bar{y}_j	S_j^2	\hat{y}_j
					1	2	3			
1	+	+	+	+	8.38	8.24	8.21	8.28	0.0051	8,292
2	+	-	+	+	8.18	8.50	8.26	8.32	0.019	8,272
3	+	+	-	+	8.39	8.74	8.15	8.43	0.059	8,478
4	+	-	-	+	8.27	8.60	8.54	8.47	0.021	8,458
5	+	+	+	-	7.51	7.73	8.44	7.62	0.158	8,032
6	+	-	+	-	8.10	7.93	8.32	8.12	0.025	8,012
7	+	+	-	-	8.54	8.40	8.32	8.42	0.009	8,218
8	+	-	-	-	7.99	8.09	8.02	8.03	0.0020	8,198

3. Проведение имитационного моделирования. Проведение имитационного эксперимента осуществляется в следующем порядке.

Для каждого опыта в соответствии с блоком планирования (*столбцы переменных выделены жирной чертой*) в программе «Планирование эксперимента» (MathCAD, рис. 6.1) устанавливают значения переменных в физической системе (табл. 6.1) на верхнем (+1) или на нижнем (-1) уровне.

Задают моделирование трех значений ($n:=3$) центрированных случайных поправок.

Результаты моделирования заносят в таблицу 6.2 :

- значения функции отклика для каждой повторности – 1,2 и 3;
- среднее $[\text{mean}(y)]$;
- дисперсию $[\text{var}(y)]$.

В указанном порядке проводят 8 опытов.

Планирование эксперимента

ввод значений переменных из области определения факторов в исходной размерности

$$x_1 := 2 \quad x_2 := 12 \quad x_3 := 14$$

ввести каждый раз число повторных опытов $n := 3$

Значения функции отклика в повторностях $c := \text{norm}(n, 0, 0.3)$ $y := y_0 + c$

$$y = \begin{pmatrix} 8.029 \\ 7.957 \\ 8.019 \end{pmatrix}$$

среднее $\text{mean}(y) = 8.001$

дисперсия

$$\text{var}(y) = 1.015 \times 10^{-3}$$

Рис.

6.1.

Программа имитационного моделирования «Планирование эксперимента»

4. Дисперсионный анализ результатов эксперимента.

4.1. Для оценки наличия грубых ошибок в полученных данных проверяют гипотезу об однородности ряда дисперсий, для этого выбирают максимальное значение дисперсий из проведенных опытов и подсчитывают расчетное значение критерия Кохрена

$$G_p = \frac{s_{jmax}^2}{\sum s_j^2} = \frac{0,158}{0,2981} = 0,53. \quad (6.3)$$

4.2. Критическое значения критерия находят по таблице 8 (приложение 1) для относительной ошибки $\alpha=0,05$, для 8 опытов и 3-х повторностей $n=3$

$$G_{кр} \equiv G_{\alpha=0.05; N=8; n-1=2} = 0,715. \quad (6.4)$$

4.3. Расчетное значение сравниваем с критическим и делаем вывод: так как $G_p < G_{кр}$ следовательно ряд дисперсий однороден и серии опытов (столбцы 6-8) не содержат грубых ошибок.

4.4. Определяем дисперсию воспроизводимости, оценивающую точность определения показателя y

$$S_e^2 = \frac{\sum S_j^2}{N} = \frac{0,2981}{8} = 0,037. \quad (6.5)$$

Число степеней свободы для дисперсии воспроизводимости S_e^2 равно $f_e = N(n - 1) = 8 \cdot 2 = 16$.

4.5. Определяем коэффициенты a_i уравнения регрессии (6.6) по методу наименьших квадратов

$$\hat{y} = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_3. \quad (6.6)$$

Для ортогональной матрицы ($\sum x_i = 0$) *метод наименьших квадратов* примет вид

$$a_i = \frac{1}{N} \sum_{j=1}^N x_{ij} \cdot \bar{y}_j. \quad (6.7)$$

$$a_0 = \frac{1}{8}(65,96) = 8,245$$

$$a_1 = \frac{1}{8}(0,08) = 0,01$$

$$a_2 = \frac{1}{8}(-0,74) = -0,093$$

$$a_3 = \frac{1}{8}(1,04) = 0,13$$

Получим уравнение

$$\hat{y} = 8,245 + 0,01x_1 - 0,093x_2 + 0,13x_3.$$

4.6. Определяем дисперсию адекватности. Для этого подсчитаем значение отклика для заданных переменных:

$$\hat{y}_1 = 8,245 + 0,01 - 0,093 + 0,13 = 8,292$$

$$\hat{y}_2 = 8,245 - 0,01 - 0,093 + 0,13 = 8,272$$

$$\hat{y}_3 = 8,245 + 0,01 + 0,093 + 0,13 = 8,478$$

$$\hat{y}_4 = 8,245 - 0,01 + 0,093 + 0,13 = 8,458$$

$$\hat{y}_5 = 8,245 + 0,01 - 0,093 - 0,13 = 8,032$$

$$\hat{y}_6 = 8,245 - 0,01 - 0,093 - 0,13 = 8,012$$

$$\hat{y}_7 = 8,245 + 0,01 + 0,093 - 0,13 = 8,218$$

$$\hat{y}_8 = 8,245 - 0,01 + 0,093 - 0,13 = 8,198$$

$$S_{ad}^2 = \frac{1}{f_{ad}} \cdot \sum_{j=1}^N (y_j - \hat{y}_j)^2 = \frac{1}{4} \cdot 0,255 = 0,064 \quad (6.8)$$

Число степеней свободы равно $f_{ad} = N - d - 1 = N - 3 - 1 = 4$.

4.7. Оцениваем адекватность полученного уравнения, для этого сравниваем дисперсию воспроизводимости S_e^2 и дисперсию адекватности S_{ad}^2 и подсчитываем расчетное значение критерия Фишера

$$F_p = \frac{S_{ad}^2}{S_e^2} = \frac{0,064}{0,037} = 1,725. \quad (6.9)$$

Критическое значение критерия находим по таблице (приложения 1) для $\alpha=0.05$; $f_{ad} = 4$ и $f_e = 16$

$$F_{кр} \equiv F_{\alpha=0.05; f_{ad}=4; f_e=16} = 3,007. \quad (6.10)$$

Сравниваем расчетное значение с критическим и делаем вывод: так, как расчетное значение меньше критического, то делаем вывод о том, что сравни-

ваемые дисперсии отличаются не существенно, следовательно модель адекватна. Это означает то, что провести дополнительный опыт или рассчитать значение по уравнению регрессии равнозначно.

4.8. Оцениваем значимость коэффициентов уравнения регрессии. Для этого определяем среднее квадратическое отклонение для коэффициентов

$$S_{a_i} = \sqrt{\frac{S_6^2}{N \cdot n}} = \sqrt{\frac{0,037}{8 \cdot 3}} = 0,039 \quad (6.11)$$

Далее определяем расчетные значения критерия Стьюдента для каждого коэффициента при неизвестных

$$t_{p1} = \frac{a_1}{S_{a_i}} = \frac{0,01}{0,039} = 0,256;$$

$$t_{p2} = \frac{a_2}{S_{a_i}} = \frac{0,093}{0,039} = 2,38;$$

$$t_{p3} = \frac{a_3}{S_{a_i}} = \frac{0,13}{0,039} = 3,33.$$

По таблице 3 (приложения 1) определяем критическое значение критерия Стьюдента для $\alpha=0.05$, $k = N \cdot n = 8 \cdot 3 = 24$

$$t_{кр} \equiv t_{\alpha=0.05; k=24} = 2.064.$$

Сравниваем расчетные значения с критическим и делаем вывод: коэффициент $a_1=0.01$ оказался незначим, так как t_{p1} меньше критического и его можно исключить из уравнения регрессии, коэффициенты a_2 и a_3 – значимы.

Контрольные вопросы:

1. Что собой представляет нормированное факторное пространство с единичным ортом?
2. Как подсчитать значения отклика у для заданной точки факторного пространства: x_1 , x_2 и x_3 ?
3. Можно ли по значениям коэффициентов оценить силу влияния каждого фактора?

7. Планирование дробного факторного эксперимента (ДФЭ) 2^{n-k} .

7.1. Цель и задачи

Цель – изучить методику планирования дробного многофакторного эксперимента для получения линейной формы уравнения.

Задачи – освоить основные понятия планирования дробного многофакторного эксперимента, научиться планировать, проводить и обрабатывать результаты многофакторного эксперимента и оценивать адекватность уравнения.

7.2. Основные понятия и определения.

Многофакторные процессы описываются зависимостью случайной величины y (показатель, отклик) от переменных (аргументов, действующих факторов) x_j ($j = 1, 2, \dots, k$) случайных величин, оцениваемых средними значениями, независимо от истинного закона распределения x_j .

$$\hat{y} = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_j x_j + v, \quad (7.1)$$

где $\alpha_0, \alpha_1, \dots, \alpha_j$ - коэффициенты уравнения регрессии.

Проведение эксперимента предполагает одновременное изменение всех действующих факторов в соответствии с блоком планирования План-матрицы эксперимента (Таблица 7.2).

Для повышения точности оценочных значений коэффициентов уравнения регрессии (7.1) проводят параллельные (воспроизводимые) опыты (повторности).

7.3. Задание.

С помощью имитационной модели для условного 3-х факторного пространства и диапазонов варьирования факторов, принятых при моделировании полного факторного эксперимента (задание 6):

- построить план-матрицу дробного факторного эксперимента ДФЭ 2^{3-1} ;
- провести имитационное моделирование в соответствии с блоком планирования План-матрицы и заданными уровнями варьирования;
- провести дисперсионный анализ полученных результатов;
- получить уравнение регрессии и оценить его адекватность.

7.4. Порядок выполнения задания

1. Задаемся уровнями варьирования действующих факторов (таблица 7.1). В данном задании уровни варьирования факторов возьмем из задания 6.

Таблица 7.1. Уровни варьирования действующих факторов

Уровни	X_1	X_2	X_3
1.Нижний, x_n	7	13	24
2.Верхний, x_v	33	37	48
3.Основной, x_o	20	25	36
4.Интервал варьирования, J_i	13	12	12

2. Составим План-матрицу (таблица 7.2) дробного факторного эксперимента 2^{3-1} . Для этого берем за основу План-матрицу ПФЭ 2^2 и приравниваем $x_3 = x_1 \cdot x_2$. Это означает, что a_3 будет учитывать одновременное влияние фактора x_3 и двух смешанных эффектов. Соответственно и коэффициенты a_1 и a_2 будут учитывать влияние на функцию отклика факторов и смешанные парные эффекты

Таблица 7.2 План-матрица ДФЭ 2^{3-1}

№ опыта	X_0	X_1	X_2	$X_1 X_2 = X_3$	\bar{y}_j	S_j^2	\hat{y}_j
1	+	+	+	+	8,172	0,045	8,172
2	+	-	+	-	8,55	0,085	8,55
3	+	+	-	-	8,577	0,005	8,577
4	+	-	-	+	8,322	0,068	8,322

3. Проведем имитационное моделирование.

Проведение эксперимента осуществляется в том же порядке, что и в предыдущем задании б.

Результаты моделирования заносят в таблицу 7.2 :

- среднее [$\text{mean}(y)$];
- дисперсию [$\text{var}(y)$].

4. Дисперсионный анализ результатов эксперимента

4.1. Проверка однородности ряда дисперсий

$$G_p = \frac{S_{jmax}^2}{\sum S_j^2} = \frac{0,085}{0,203} = 0,419. \quad (7.2)$$

Критические значения критерия Кохрена находим по таблице 8 (приложения 1). Для $N=4$, $k=n-1=3-1=2$ равно

$$G_{kp} \equiv G_{\alpha=0.05; N=4; n-1=2} = 0,906 \quad . \quad (7.3)$$

Расчетное значение критерия Кохрена сравниваем с критическим и делаем вывод: так, как $G_p < G_{kp}$, то принимаем гипотезу о том, что ряд дисперсий однороден, и следовательно в серии опытов грубых ошибок нет.

4.2. Определяем дисперсию воспроизводимости

$$S_e^2 = \frac{\sum S_j^2}{N} = \frac{0,203}{4} = 0,051. \quad (7.4)$$

Число степеней свободы для дисперсии воспроизводимости S_e^2 равно $f_e = N(n - 1) = 4 \cdot 2 = 8$.

4.3. Определяем коэффициенты a_i уравнения регрессии (6.6) по методу наименьших квадратов

$$a_i = \frac{1}{N} \sum_{j=1}^N x_{ij} \cdot \bar{y}_j \quad (7.4)$$

$$a_1 = \frac{1}{4} \sum_{j=1}^4 x_{1j} \cdot \bar{y}_j = 0,25(-0,123) = -0,031;$$

$$a_2 = \frac{1}{4} \sum_{j=1}^4 x_{2j} \cdot \bar{y}_j = 0,25(-0,177) = -0,044.$$

$$a_3 = \frac{1}{4} \sum_{j=1}^4 x_{3j} \cdot \bar{y}_j = 0,25(-0,633) = -0,158.$$

Получим уравнение $\hat{y} = 8,405 - 0,031x_1 - 0,044x_2 - 0,158x_3$.

4.4. Определяем дисперсию адекватности

$$S_{ad}^2 = \frac{1}{f_{ad}} \cdot \sum_{j=1}^N (y_j - \hat{y}_j)^2 = \frac{1}{3} \cdot 0 = 0 \quad (7.5)$$

Число степеней свободы равно $f_{ad} = N - 0 - 1 = 3$. Если дисперсия адекватности близка к 0, точки принадлежат поверхности и п 4.5 не проводят, считаем модель адекватной.

4.5. Для оценки адекватности полученного уравнения подсчитываем расчетное значение критерия Фишера

$$F_p = \frac{S_{ad}^2}{S_g^2} = \dots = \dots \quad (7.6)$$

Критическое значение критерия находим по таблице (приложения 1) для $\alpha=0.05$; $f_{ad} = 3$ и $f_g = 8$

$$F_{кр} \equiv F_{\alpha=0.05; f_{ad}=3; f_g=8} = 4,134. \quad (7.9)$$

Сравниваем расчетное значение с критическим, если дисперсия адекватности имеет численной значение.

Контрольные вопросы:

1. По какому принципу формируется блок планирования?
2. Что оцениваю коэффициенты уравнения?
3. Можно ли по значению коэффициентов однозначно утверждать о силе влияния того или иного фактора ?
4. Какое преимущество и недостатки имеет метод дробного планирования?
5. Что означает ДФЭ 2^{n-k} ?

8. Поиск области экстремума функции методом крутого восхождения (спуска)

8.1. Цель и задачи

Цель – изучить методику планирования многофакторного эксперимента для поиска области расположения экстремума функции.

Задачи – освоить основные понятия планирования многофакторного эксперимента по методу крутого спуска (восхождения), научиться планировать, проводить и обрабатывать результаты многофакторного эксперимента при поиске опытным путем экстремума функции.

8.2. Основные понятия и определения.

Для нахождения оптимального решения гладкой дифференцируемой функции (рис.8.1), для которой нет теоретического описания, применяют экспериментальные методы многофакторного планирования. **Метод крутого восхождения** был предложен Боксом и Уилсоном в 1951 году, поэтому его называют также методом Бокса-Уилсона.

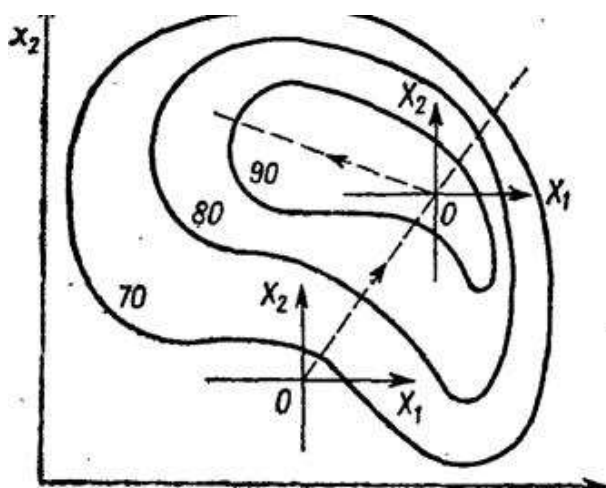


Рис. 8.1 Линии одинакового уровня для исследуемой функции и движение по градиенту

Одним из наиболее эффективных методов является **метод крутого восхождения**, который представляет собой сочетание градиентного метода оптимизации с дробным факторным экспериментом, применяемым для аппроксимации поверхности отклика полиномом первого порядка, обеспечивающим получение оценки градиента функции отклика в заданной точке с минимальной дисперсией.

Градиентом называют вектор, показывающий направление наискорейшего изменения некоторой величины, значение которой меняется от одной точки пространства к другой. Порядок оптимизации:

1. Выбирается (на основе априорной информации) начальная точка x^0 , из которой будет осуществляться поиск оптимальных (близких к нему) значений входных переменных.

2. Используя эту точку как центр плана, строится полный или дробный факторный план, по результатам эксперимента в котором строится линейная модель функции отклика:

$$y(x,b)=b_0+b_1x_1+\dots+b_kx_k \quad (8.1)$$

Вид плана выбирается таким образом, чтобы перенасыщенность его была минимальной (число точек плана по возможности не намного превышало число определяемых коэффициентов).

Данное уравнение является уравнением гиперплоскости в $(k+1)$ -мерном пространстве, касательной к поверхности отклика в точке x^0 . Коэффициенты b_i характеризуют её наклон относительно осей координат и являются составляющими вектора *градиента*, обеспечивающего движение в области изменения переменных x_i в направлении крутого (наискорейшего) восхождения по этой гиперплоскости в сторону возрастания целевой функции :

$$\overline{grad} y = \frac{dy}{dx_1} \overline{i_1} + \dots + \frac{dy}{dx_k} \overline{i_k} = \tilde{b}_1 \cdot \overline{i_1} + \dots + \tilde{b}_k \cdot \overline{i_k} \quad (8.2)$$

где $\overline{i_i}$ -единичный вектор в направлении переменных x_i (орты).

3.Осуществляется движение вдоль линии крутого восхождения, определяя в каждой точке значения функции отклика по результатам наблюдений. Движение в данном направлении осуществляется до тех пор, пока не будет достигнуто наибольшее значение функции отклика Y_{max} .

8.3. Задание.

Для уровней варьирования 3-х факторов, принятые в задании 6, и уравнения регрессии (6.6) :

- рассчитать допустимые смещения вдоль осей для движения вдоль вектора градиента и составить план эксперимента;
- с помощью имитационной модели для условного 3-х факторного пространства проводить опыты, начиная с центра факторного пространства;
- после каждого опыта провести анализ динамики изменения частного сечения неизвестной функции отклика и проводить опыты до достижения максимума.

8.4. Порядок выполнения задания

1.Задаем уровни варьирования факторов для перехода в кодированное пространство. При выполнении задания уровни варьирования возьмем из задания 6 (таблица 8.1).

2. Уравнение регрессии, полученное при проведении полного факторного эксперимента (перенести из задания 6)

$$\hat{y} = 8,245 - 0,093 x_2 + 0,13x_3. \quad (8.3)$$

Таблица 8.1. Уровни варьирования действующих факторов

Уровни	X ₁	X ₂	X ₃
1.Нижний	7	13	24
2.Верхний	33	37	48
3.Основной	20	25	36
4. Интервал варьирования	13	12	12

3. Построение расчетной таблицы и матрицы планирования

3.1. В первую строчку запишем коэффициенты уравнения с соблюдением знака (учитываем только значимые значения).

3.2. Значения коэффициентов умножаем на величину интервала варьирования (табл.81) и результат записываем во вторую строчку.

3.3. Находим базовый коэффициент, соответствующий минимуму абсолютных значений второй строчки.

Таблица 8.2 Расчетная таблица матрица планирования эксперимента

Показатель	X ₂	X ₃	Расчеты
1. Коэффициенты уравнения, a_i	-0,093	0,13	$Min\{ a_i \cdot J_i \} = 1,12$ $a_{\text{баз}} = 0,093$. Масштабный коэффициент: $k_m = 1/a_{\text{баз}} = 10,75$
2. Произведение $a_i \cdot J_i$	-1,12	1,56	
3. Расчетное смещение (шаг движения), $a_i \cdot J_i \cdot k_m$	-12,04	16,77	
4. Округленный шаг смещения Δ_i	-12	16	
-1	32	9	$y_{-1} = 8,259$
0	20	25	$y_0 = 8,296$
1	8	41	$Y_1 = 8,136$
2			

3.4. Подсчитываем масштабный коэффициент и рассчитываем расчетный шаг движения вдоль градиента (строчка 3). Расчетные значения округляют до удобного для реализации эксперимента значения Δ_i .

4. Проведение имитационного моделирования

Задать начальную точку, соответствующую центру эксперимента (точка x_{i0}). Проведение эксперимента осуществляется в том же порядке, что и задании б.

Результаты моделирования среднее $[\text{mean}(y)]$ заносят в таблицу 8.2. и анализируем значения функции отклика. Если происходит увеличение функции, продолжаем движение вдоль градиента пока значения не начнут уменьшаться. В окрестности максимальной точки далее следует искать искомый максимум. В приведенном примере максимум функции отклика $y_{\max} = 8,296$ достигается в точке 1 с координатами $X_2 = 20; X_3 = 25$.

Контрольные вопросы:

1. Как выбрать начальную точку планирования эксперимента?
2. Как определить направление движения в факторном пространстве?
3. Как определить шаг движения вдоль градиента?
4. При каком условии прекращается проведение эксперимента?

9. Описание области экстремума функции отклика уравнениями второго порядка

9.1. Цель и задачи

Цель – изучить методику планирования многофакторного эксперимента для описания области экстремума функции уравнениями второго порядка.

Задачи – освоить основные понятия планирования многофакторного эксперимента для планов второго порядка, научиться планировать, проводить и обрабатывать результаты многофакторного эксперимента при описании области экстремума функции уравнениями второго порядка.

9.2. Основные понятия и определения

Для достижения почти стационарной области (области наилучших значений выходной переменной) используется два подхода.

1. Методы, использующие аппроксимацию поверхности отклика в локальной области гиперплоскостью с целью определения наилучшего направления движения в сторону оптимума (например, метод крутого восхождения - при максимизации функции отклика).

2. Методы, определяющие направление движения на каждом шаге по результатам очередного наблюдения, сравнивая его с результатами предыдущих наблюдений (например, симплекс-метод).

Второй подход, по сравнению с первым, допускает применение на действующих системах (установках) за счёт незначительных отклонений от номинальных режимов работы и не приводит к недопустимым режимам (метод эволюционного планирования или промышленный эксперимент).

Для исследования почти стационарной области целесообразно аппроксимировать поверхность отклика квадратичным полиномом второго порядка. В обобщенном виде уравнение второго порядка представлено выражением

$$\hat{y} = a_0 + \sum a_i x_i + \sum a_{ij} x_i x_j + \sum a_{ij} x_i^2 \quad (9.1)$$

Для двухфакторного пространства ($x_2 = const$) развернутое выражение полинома второго порядка имеет вид

$$\hat{y} = a_0 + a_1 x_1 + a_2 x_2 + a_{12} x_1 x_2 + a_{11} x_1^2 + a_{22} x_2^2 \quad (9.2)$$

Для нахождения экстремума функции найдем частные производные и решим систему уравнений

$$\left. \begin{aligned} \frac{dy}{dx_1} = a_1 + a_{12} x_2 + 2a_{11} x_1 = 0 \\ \frac{dy}{dx_2} = a_2 + a_{12} x_1 + 2a_{22} x_2 = 0 \end{aligned} \right\} \quad (9.3)$$

Для оценивания коэффициентов регрессии в этой модели необходимо использовать 3^n - факторный эксперимент.

Для описания области экстремума применяются полиномиальные модели с учетом квадратов, а иногда и более высоких степеней факторов, которые позволяют установить точное положение экстремума методами аналитической геометрии.

Рассмотрим наиболее часто встречающийся случай построения поверхности второго порядка. Чтобы построить полином, содержащий квадраты факторов, требуется каждый фактор варьировать не менее чем на трех уровнях. Если при этом плану эксперимента необходимо придать некоторые свойства (ортогональности, рототабельности), опытные точки в факторном пространстве следует расположить специальным образом.

На практике часто используют план V_m . При выполнении задания ограничимся двух факторным пространством ($m=2$).

9.3. Задание

С помощью имитационной модели для условного 2-х факторного пространства:

- задаться уровнями варьирования факторов для области нахождения экстремума;
- построить план-матрицу эксперимента V_2 ;
- провести имитационное моделирование в соответствии с блоком планирования план-матрицы и заданными уровнями варьирования;
- рассчитать коэффициенты уравнения второго порядка и определить экстремум функции отклика.

9.4. Порядок выполнения задания

1. Задаем уровни варьирования факторов. Для этого в качестве координат центра эксперимента берем координаты точки локального экстремума, полученные в задании 8.

Максимум функции отклика вдоль градиента $y_{\max} = 8,296$ достигается в точке 1 с координатами $X_2 = 20$; $X_3 = 25$ (значение $X_1 = 20$). Затем задаем интервалами варьирования и рассчитываем верхний и нижний уровни варьирования факторов.

Таблица 9.1. Уровни варьирования факторов для плана B_2

Уровни	$X_1 = \text{const}$	X_2	X_3
1. Основной	20	20	25
2. Интервал варьирования	15	10	15
3. Нижний	5	10	10
4. Верхний	35	30	40

2. Строим матрицу планирования эксперимента. В качестве ядра плана берем план ПФЭ 2^2 (4 опыта). В качестве звездных точек выбираем пересечение вписанной окружности в факторное пространство с координатными осями (2n опытов). План-матрица B_2 приведена в таблице 9.2.

Таблица 9.2. План- матрица эксперимента B_2

	X_0	X_1	X_2	$X_1 X_2$	X_1^2	X_2^2	\bar{y}	S_j^2
1	+	+	+	+	+	+	8.199	0.001
2	+	-	+	-	+	+	8.037	0.045
3	+	+	-	-	+	+	8.54	0.085
4	+	-	-	+	+	+	8.481	0.004
5	+	+	0	0	+	0	8.379	0.068
6	+	-	0	0	+	0	8.263	0.032
7	+	0	+	0	0	+	8.174	0.005
8	+	0	-	0	0	+	8.228	0.019

3. Проведем имитационное моделирование. Проведение имитационного эксперимента по программе «Планирование эксперимента», (см. Рис. 6.1) осуществляется в том же порядке, что и в задании 6.

Результаты моделирования заносят в таблицу 9.2 :

- среднее [mean(y)] ;

- дисперсию [var(y)].

4. Дисперсионный анализ результатов эксперимента.

4.1. Для оценки наличия грубых ошибок в полученных данных проверяем гипотезу об однородности ряда дисперсий, для этого выбираем максимальное значение дисперсии из проведенных опытов и подсчитываем расчетное значение критерия Кохрена

$$G_p = \frac{S_{jmax}^2}{\sum S_j^2} = \frac{0,085}{0,259} = 0,328. \quad (9.4)$$

4.2. Критическое значения критерия находят по таблице 8 (приложение 1) для относительной ошибки $\alpha=0,05$, для 8 опытов и 3-х повторностей $n=3$

$$G_{кр} \equiv G_{\alpha=0.05;N=8;n-1=2} = 0,715. \quad (9.5)$$

4.3. Расчетное значение сравниваем с критическим и делаем вывод: так как $G_p < G_{кр}$ следовательно ряд дисперсий однороден и в серии опытов (повторности) не содержат грубых ошибок.

4.4. Определяем дисперсию воспроизводимости, оценивающую точность определения показателя y

$$S_g^2 = \frac{\sum S_j^2}{N} = \frac{0,259}{8} = 0,032. \quad (9.6)$$

Число степеней свободы для дисперсии воспроизводимости S_g^2 равно $f_g = N(n - 1) = 8 \cdot 2 = 16$.

4.5. Значения коэффициентов уравнения регрессии находим по матричному уравнению метода наименьших квадратов $B = (X^T X)^{-1} X^T Y$ (Программа В2Н, рис. 9.1.).

Для расчета вводим вектор-столбец средних значений функции отклика, полученные в эксперименте и в результате получим вектор-столбец B значений коэффициентов уравнения регрессии (9.2). Получим уравнение:

$$\hat{y} = 8,199 + 0,056x_1 - 0,14x_2 + 0,026x_1 \cdot x_2 + 0,113x_1^2 - 0,0067x_2^2.$$

4.6. Определяем дисперсию адекватности (по программе)

$$S_{ad}^2 = \frac{1}{f_{ad}} \cdot \sum_{j=1}^N (y_j - \hat{y}_j)^2 = 0,038 \quad (9.7)$$

Число степеней свободы равно $f_{ad} = N - 5 - 1 = 2$.

Для оценки адекватности полученного уравнения подсчитываем расчетное значение критерия Фишера

$$F_p = \frac{S_{ad}^2}{S_g^2} = \frac{0,038}{0,032} = 1,187. \quad (9.8)$$

План-матрица B2 оптимального плана

$$\begin{aligned}
 X &:= \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & -1 & 0 & 0 & 1 \end{pmatrix} & D := X^T \cdot X & Y := \begin{pmatrix} 8.7 \\ 8.75 \\ 8.34 \\ 8.5 \\ 8.52 \\ 8.54 \\ 8.53 \\ 8.51 \end{pmatrix} \\
 G &:= X \cdot B & C &:= \text{geninv}(D) & S &:= Y - G & S^2 &= \begin{pmatrix} 1.11 \times 10^{-3} \\ 3.799 \times 10^{-3} \\ 3.799 \times 10^{-3} \\ 1.11 \times 10^{-3} \\ 8.032 \times 10^{-4} \\ 8.032 \times 10^{-4} \\ 9.029 \times 10^{-3} \\ 9.029 \times 10^{-3} \end{pmatrix} \\
 B &= \begin{pmatrix} 8.477 \\ -0.038 \\ 0.105 \\ 0.027 \\ 0.053 \\ 0.043 \end{pmatrix} & G &= \begin{pmatrix} 8.667 \\ 8.688 \\ 8.402 \\ 8.533 \\ 8.492 \\ 8.568 \\ 8.625 \\ 8.415 \end{pmatrix} & a_0 &:= B_0 & a_1 &:= B_1 & a_2 &:= B_2 & a_{12} &:= B_3 & a_{11} &:= B_4 & a_{22} &:= B_5 \\
 \text{yp}(X1, X2) &:= a_0 + a_1 \cdot X1 + a_2 \cdot X2 + a_{12} \cdot X1 \cdot X2 + a_{11} \cdot X1^2 + a_{22} \cdot X2^2
 \end{aligned}$$

Рис. 9.1. Программа нахождения коэффициентов уравнения второго порядка

Критическое значение критерия находим по таблице 7 (приложения 1) для $\alpha=0.05$; $f_{a0} = 2$ и $f_b = 16$

$$F_{кр} \equiv F_{\alpha=0.05; f_{a0}=3; f_b=8} = 3,634. \quad (9.9)$$

Сравниваем расчетное значение критерия Фишера с критическим значением. Так, как расчетное значение меньше критического, то принимаем гипотезу об адекватности модели.

4.7. Оцениваем значимость коэффициентов уравнения регрессии. Для этого определяем среднее квадратическое отклонение для коэффициентов

$$S_{a_i} = \sqrt{\frac{S_b^2}{N \cdot n}} = \sqrt{\frac{0,038}{8 \cdot 3}} = 0,04 \quad . \quad (6.11)$$

Далее определяем расчетные значения критерия Стьюдента для каждого коэффициента при неизвестных

$$\begin{aligned}
 t_{p1} &= \frac{a_1}{S_{a_i}} = \frac{0,056}{0,04} = 1,4; \\
 t_{p2} &= \frac{a_2}{S_{a_i}} = \frac{0,14}{0,04} = 3,5; \\
 t_{p12} &= \frac{a_{12}}{S_{a_i}} = \frac{0,026}{0,04} = 0,65;
 \end{aligned}$$

$$t_{p11} = \frac{a_{11}}{S_{a_i}} = \frac{0,113}{0,04} = 2,82,;$$

$$t_{p22} = \frac{a_{22}}{S_{a_i}} = \frac{0,0067}{0,04} = 0,167.$$

По таблице 3 (приложения 1) определяем критическое значение критерия Стьюдента для $\alpha=0.05$, $k = N \cdot n = 8 \cdot 3 = 24$

$$t_{кр} \equiv t_{\alpha=0.05;k=24} = 2.064.$$

Сравниваем расчетные значения с критическим и делаем вывод: коэффициент a_{11} , a_{12} , a_{22} оказались незначимыми, так как t_{pi} меньше критического.

4.8. Определяем координаты максимума функции. Для этого находим две частные производные уравнения регрессии (п.4.5)

$$\frac{dy}{dx_1} = 0.056 + 0.026x_2 + 0.226x_1 = 0$$

$$\frac{dy}{dx_2} = -0.14 + 0.026x_1 - 0.007x_2 = 0$$

В ходе решения полученной системы уравнений получим значения координат экстремума функции в кодированном виде $x_{1\text{опт}} = 1.859$ и $x_{2\text{опт}} = -0.086$. Для определения координат оптимума необходимо воспользоваться формулой кодировки системы координат (6.2).

Контрольные вопросы:

1. По какому принципу формируется план-матрица для нахождения уравнений второго порядка?
2. На что влияет выбираемая величина интервала варьирования?
3. Как повысить точность определения коэффициентов уравнения регрессии ?
4. Какие виды планов второго порядка Вы знаете и опишите их особенности?
5. Как определить экстремум функции?

Таблица 1. Значения $U_{\alpha; n}$ критерия V (для отбрасывания грубых ошибок при измерениях), определяемые из условия $P(V > U_{\alpha; n}) = \alpha$

n	α			n	α		
	0,10	0,05	0,01		0,10	0,05	0,01
3	1,406	1,412	1,414	20	2,447	2,623	2,959
5	1,791	1,869	1,955	25	2,537	2,717	3,071
7	1,974	2,093	2,265	30	2,609	2,792	3,156
10	2,146	2,294	2,540	40	2,718	2,904	3,281
15	2,326	2,493	2,800	50	2,800	2,987	3,370

Таблица 2. Критические значения $\tau_{\alpha; k}$ критерия τ , определяемые из условия $P(\tau < \tau_{\alpha; k}) = \alpha$

n	α		n	α	
	0,05	0,01		0,05	0,01
4	0,390	0,256	10	0,531	0,376
5	0,410	0,269	12	0,564	0,414
6	0,445	0,281	14	0,591	0,447
7	0,468	0,307	16	0,614	0,475
8	0,491	0,331	18	0,633	0,499
9	0,512	0,354	20	0,650	0,520

Таблица 3. Критические точки распределения Стьюдента.

$k \setminus \alpha$	0,1	0,05	0,02	0,01	0,001
1	6,3138	12,7062	31,8205	63,6567	636,6192
2	2,9200	4,3027	6,9646	9,9248	31,5991
3	2,3534	3,1824	4,5407	5,8409	12,924
4	2,1318	2,7764	3,7469	4,6041	8,6103
5	2,0150	2,5706	3,3649	4,0321	6,8688
6	1,9432	2,4469	3,1427	3,7074	5,9588
7	1,8946	2,3646	2,9980	3,4995	5,4079
8	1,8595	2,3060	2,8965	3,3554	5,0413
9	1,8331	2,2622	2,8214	3,2498	4,7809
10	1,8125	2,2281	2,7638	3,1693	4,5869
11	1,7959	2,2010	2,7181	3,1058	4,4370
12	1,7823	2,1788	2,6810	3,0545	4,3178
13	1,7709	2,1604	2,6503	3,0123	4,2208
14	1,7613	2,1448	2,6245	2,9768	4,1405
15	1,7531	2,1314	2,6025	2,9467	4,0728
16	1,7459	2,1199	2,5835	2,9208	4,0150
17	1,7396	2,1098	2,5669	2,8982	3,9651
18	1,7341	2,1009	2,5524	2,8784	3,9216
19	1,7291	2,0930	2,5395	2,8609	3,8834
20	1,7247	2,0860	2,5280	2,8453	3,8495
21	1,7207	2,0796	2,5176	2,8314	3,8193
22	1,7171	2,0739	2,5083	2,8188	3,7921
23	1,7139	2,0687	2,4999	2,8073	3,7676
24	1,7109	2,0639	2,4922	2,7969	3,7454

25	1,7081	2,0595	2,4851	2,7874	3,7251
26	1,7056	2,0555	2,4786	2,7787	3,7066
27	1,7033	2,0518	2,4727	2,7707	3,6896
28	1,7011	2,0484	2,4671	2,7633	3,6739
29	1,6991	2,0452	2,4620	2,7564	3,6594
30	1,6973	2,0423	2,4573	2,7500	3,6460
35	1,6896	2,0301	2,4377	2,7238	3,5911
40	1,6839	2,0211	2,4233	2,7045	3,5510
45	1,6794	2,0141	2,4121	2,6896	3,5203
50	1,6759	2,0086	2,4033	2,6778	3,4960
55	1,6730	2,004	2,3961	2,6682	3,4764
60	1,6706	2,0003	2,3901	2,6603	3,4602
70	1,6669	1,9944	2,3808	2,6479	3,4350
80	1,6641	1,9901	2,3739	2,6387	3,4163
90	1,6620	1,9867	2,3685	2,6316	3,4019
100	1,6602	1,9840	2,3642	2,6259	3,3905
110	1,6588	1,9818	2,3607	2,6213	3,3812
120	1,6577	1,9799	2,3578	2,6174	3,3735
∞	1,6448	1,9600	2,3263	2,5758	3,2905

$$T = \frac{U\sqrt{n}}{\sqrt{X}}$$

Распределение t Стьюдента – это распределение случайной величины где случайные величины U и X независимы, U имеет распределение стандартное нормальное распределение $N(0,1)$, а X – распределение хи – квадрат с n степенями свободы. При этом n называется «числом степеней свободы» распределения Стьюдента.

Кривая плотности t -распределения похожа на кривую нормального распределения, но имеет более пологую форму. При количестве степеней свободы $k \geq 30$ t -распределение переходит в нормальное с параметрами $m_x = 0$; $\sigma_x = \sqrt{\frac{k}{k-2}}$.

Критические значения t -критерия Стьюдента можно подсчитать в Excel набрав команду: =СТЮДРАСПОБР(α ; ν), где α - уровень значимости, ν - число степеней свободы.

Таблица 4. Критические значения критерия Пирсона (χ^2)

Число степеней свободы	Уровень значимости α					
	0,01	0,05	0,1	0,90	0,95	0,99
1	6,6	3,8	2,71	0,02	0,004	0,0002
2	9,2	6,0	4,61	0,21	0,1	0,02
3	11,3	7,8	6,25	0,58	0,35	0,12
4	13,3	9,5	7,78	1,06	0,71	0,30
5	15,1	11,1	9,24	1,61	1,15	0,55
6	16,8	12,6	10,6	2,20	1,64	0,87
7	18,5	14,1	12,0	2,83	2,17	1,24
8	20,1	15,5	13,4	3,49	2,73	1,65
9	21,7	16,9	14,7	4,17	3,33	2,09
10	23,2	18,3	16,0	4,87	3,94	2,56
11	24,7	19,7	17,3	5,58	4,57	3,05
12	26,2	21,0	18,5	6,30	5,23	3,57
13	27,7	22,4	19,8	7,04	5,89	4,11
14	29,1	23,7	21,1	7,79	6,57	4,66
15	30,6	25,0	22,3	8,5	7,26	5,23
16	32,0	26,3	23,5	9,31	7,98	5,81
17	33,4	27,6	24,8	10,1	8,67	6,41
18	34,8	28,9	26,0	10,9	9,39	7,01
19	36,2	30,1	27,2	11,7	10,1	7,63
20	37,6	31,4	28,4	12,4	10,9	8,26
21	38,9	32,7	29,6	13,2	11,6	8,90
22	40,3	33,9	30,6	14,0	12,63	9,54
23	41,6	35,2	32,0	14,8	13,1	10,2
24	43,0	36,4	33,2	15,7	13,8	10,9
25	44,3	37,7	34,4	16,5	14,6	11,5
26	45,6	38,9	35,6	17,3	15,4	12,2
27	47,0	40,1	36,7	18,1	16,2	12,9
28	48,3	41,3	37,9	18,9	16,9	13,6
29	49,6	42,6	39,1	19,8	17,7	14,3
30	50,9	43,8	40,3	20,6	18,5	15,0

Распределение Пирсона χ^2 (хи - квадрат) – распределение случайной величины

$$X = X_1^2 + X_2^2 + \dots + X_n^2,$$

где случайные величины X_1, X_2, \dots, X_n независимы и имеют одно и тоже распределение $N(0,1)$. При этом число слагаемых, т.е. n , называется «числом степеней свободы» распределения хи – квадрат.

**Таблица 5 Значения критических значений критерия Колмогорова
А.Н.**

$$P\{\sqrt{n} \cdot D_n \geq \lambda\} = 1 - K(\lambda) = 1 - \sum_{k=-\infty}^{+\infty} (-1)^k e^{-k^2 \lambda^2} \quad . = \square$$

□	□	□	□	□	□	□	□	□	□	□	□
$\lambda \leq \square$ □□	□□□□ □□	0,50	0,9639	0,71	0,6945	0,92	0,3657	1,13	0,1555	1,34	0,0551
0,30	0,99999	0,51	0,9572	0,72	0,6777	0,93	0,3527	1,14	0,1486	1,35	0,0522
0,31	0,99998	0,52	0,9497	0,73	0,6609	0,94	0,3399	1,15	0,1420	1,36	0,0495
0,32	0,99995	0,53	0,9415	0,74	0,6440	0,95	0,3275	1,16	0,1356	1,37	0,0469
0,33	0,99991	0,54	0,9325	0,75	0,6272	0,96	0,3154	1,17	0,1294	1,38	0,0444
0,34	0,99993	0,55	0,9228	0,76	0,6104	0,97	0,3036	1,18	0,1235	1,39	0,0420
0,35	0,9997	0,56	0,9124	0,77	0,5936	0,98	0,2921	1,19	0,1177	1,40	0,0397
0,36	0,9995	0,57	0,9013	0,78	0,5770	0,99	0,2809	1,20	0,1122	1,41	0,0375
0,37	0,9992	0,58	0,8896	0,79	0,5605	1,00	0,2700	1,21	0,1070	1,42	0,0354
0,38	0,9987	0,59	0,8772	0,80	0,5441	1,01	0,2594	1,22	0,1019	1,43	0,0335
0,39	0,9981	0,60	0,8643	0,81	0,5280	1,02	0,2492	1,23	0,0970	1,44	0,0316
0,40	0,9972	0,61	0,8508	0,82	0,5120	1,03	0,2392	1,24	0,0924	1,45	0,0298
0,41	0,9960	0,62	0,8368	0,83	0,4962	1,04	0,2296	1,25	0,0879	1,46	0,0282
0,42	0,9945	0,63	0,8222	0,84	0,4806	1,05	0,2202	1,26	0,0836	1,47	0,0266
0,43	0,9926	0,64	0,8073	0,85	0,4653	1,06	0,2111	1,27	0,0794	1,48	0,0250
0,44	0,9903	0,65	0,7920	0,86	0,4503	1,07	0,2024	1,28	0,0755	1,49	0,0236
0,45	0,9874	0,66	0,7764	0,87	0,4355	1,08	0,1939	1,29	0,0717	1,50	0,0222
0,46	0,9840	0,67	0,7604	0,88	0,4209	1,09	0,1857	1,30	0,0681	1,51	0,0209
0,47	0,9800	0,68	0,7442	0,89	0,4067	1,10	0,1777	1,31	0,0646	1,52	0,0197
0,48	0,9753	0,69	0,7278	0,90	0,3927	1,11	0,1700	1,32	0,0613	1,53	0,0185
0,49	0,9700	0,70	0,7112	0,91	0,3791	1,12	0,1626	1,33	0,0582	1,54	0,0174

КОЛМОГорова КРИТЕРИЙ - статистический критерий, применяемый для проверки простой непараметрической гипотезы H_0 , согласно которой независимые одинаково распределенные случайные величины X_1, \dots, X_n имеют заданную непрерывную функцию распределения $F(x)$, причем

альтернативная гипотеза H_1 предполагается двусторонней: $|EF_n(x) - F(x)| > 0$, где

EF_n - математическое ожидание функции эмпирического распределения $F_n(x)$.

Таблица 6. Критические значения коэффициента корреляции Пирсона r_{xy} .

k	Уровень значимости для двустороннего критерия					k	Уровень значимости для двустороннего критерия				
	0,05	0,25	0,01	0,005	0,0005		0,05	0,25	0,01	0,005	0,0005
	Уровень значимости для одностороннего критерия						Уровень значимости для одностороннего критерия				
	0,1	0,05	0,02	0,01	0,001		0,1	0,05	0,02	0,01	0,001
1	0,98769	0,99692	0,9995	0,999877	0,9999988	21	0,352	0,413	0,482	0,526	0,640
2	0,90000	0,95000	0,980	0,990000	0,99900	22	0,344	0,404	0,472	0,515	0,629
3	0,8054	0,8783	0,934	0,95873	0,99116	23	0,337	0,396	0,462	0,505	0,618
4	0,7293	0,8114	0,882	0,91720	0,97406	24	0,330	0,388	0,453	0,496	0,607
5	0,6694	0,7545	0,833	0,8745	0,95074						
6	0,6215	0,7067	0,789	0,8343	0,92493	25	0,3233	0,3809	0,482	0,4869	0,5974
7	0,5822	0,6664	0,750	0,7977	0,8982	30	0,2960	0,3494	0,4487	0,4487	0,5541
8	0,5494	0,6319	0,715	0,7646	0,8721	35	0,2746	0,3246	0,4182	0,4182	0,5189
9	0,5214	0,6021	0,685	0,7348	0,8471	40	0,2573	0,3044	0,3932	0,3932	0,4896
10	0,4973	0,5760	0,658	0,7079	0,8233	45	0,2428	0,2875	0,3721	0,3721	0,4648
11	0,4762	0,5529	0,634	0,6835	0,8010	50	0,2306	0,2732	0,3541	0,3541	0,4433
12	0,4575	0,5324	0,612	0,6614	0,7800	60	0,2108	0,2500	0,3248	0,3248	0,4078
13	0,4409	0,5139	0,592	0,6411	0,7603	70	0,1954	0,2319	0,3017	0,3017	0,3799
14	0,4259	0,4973	0,574	0,6226	0,7420	80	0,1829	0,2172	0,2830	0,2830	0,3568
15	0,4124	0,4821	0,558	0,6055	0,7246	90	0,1726	0,2050	0,2673	0,2673	0,3375
						100	0,1638	0,1946	0,2540	0,2540	0,3211
16	0,4000	0,4683	0,542	0,5897	0,7084						
17	0,3887	0,4555	0,529	0,5751	0,6932	120	0,1500	0,1780	0,2100	0,2100	0,2940
18	0,3783	0,4438	0,515	0,5614	0,6787	∞	0,0730	0,0870	0,1030	0,1030	0,1460
19	0,3687	0,4329	0,503	0,5487	0,6652						
20	0,3598	0,4227	0,492	0,5368	0,6524						

$k = (N-2)$ – число степеней свободы.

Таблица 7 . Критические значения критерия F-Фишера

P=0,05

		Степени свободы для числителя k_1											
		1	2	3	4	5	6	7	8	10	12	24	?
K₂	3	10,128	9,552	9,277	9,117	9,013	8,941	8,887	8,845	8,785	8,745	8,638	8,527
	5	6,608	5,786	5,409	5,192	5,050	4,950	4,876	4,818	4,735	4,678	4,527	4,366
	7	5,591	4,737	4,347	4,120	3,972	3,866	3,787	3,726	3,637	3,575	3,410	3,231
	10	4,965	4,103	3,708	3,478	3,326	3,217	3,135	3,072	2,978	2,913	2,737	2,539
	11	4,844	3,982	3,587	3,357	3,204	3,095	3,012	2,948	2,854	2,788	2,609	2,406
	12	4,747	3,885	3,490	3,259	3,106	2,996	2,913	2,849	2,753	2,687	2,505	2,297
	13	4,667	3,806	3,411	3,179	3,025	2,915	2,832	2,767	2,671	2,604	2,420	2,208
	14	4,600	3,739	3,344	3,112	2,958	2,848	2,764	2,699	2,602	2,534	2,349	2,132
	15	4,543	3,682	3,287	3,056	2,901	2,790	2,707	2,641	2,544	2,475	2,288	2,067
	16	4,494	3,634	3,239	3,007	2,852	2,741	2,657	2,591	2,494	2,425	2,235	2,011
	18	4,414	3,555	3,160	2,928	2,773	2,661	2,577	2,510	2,412	2,342	2,150	1,918
	20	4,351	3,493	3,098	2,866	2,711	2,599	2,514	2,447	2,348	2,278	2,082	1,844
	30	4,171	3,316	2,922	2,690	2,534	2,421	2,334	2,266	2,165	2,092	1,887	1,624
	40	4,085	3,232	2,839	2,606	2,449	2,336	2,249	2,180	2,077	2,003	1,793	1,511
	50	4,034	3,183	2,790	2,557	2,400	2,286	2,199	2,130	2,026	1,952	1,737	1,440
	70	3,978	3,128	2,736	2,503	2,346	2,231	2,143	2,074	1,969	1,893	1,674	1,355
	100	3,936	3,087	2,696	2,463	2,305	2,191	2,103	2,032	1,927	1,850	1,627	1,286
	200	3,888	3,041	2,650	2,417	2,259	2,144	2,056	1,985	1,878	1,801	1,572	1,192
∞	3,843	2,998	2,607	2,374	2,216	2,100	2,011	1,940	1,833	1,754	1,519		

$$P > 0,01$$

		Степени свободы для числителя k_1											
		1	2	3	4	5	6	7	8	10	12	24	?
k_2	3	34,116	30,816	29,457	28,710	28,237	27,911	27,671	27,489	27,228	27,052	26,597	26,126
	5	16,258	13,274	12,060	11,392	10,967	10,672	10,456	10,289	10,051	9,888	9,466	9,022
	7	12,246	9,547	8,451	7,847	7,460	7,191	6,993	6,840	6,620	6,469	6,074	5,651
	10	10,044	7,559	6,552	5,994	5,636	5,386	5,200	5,057	4,849	4,706	4,327	3,910
	11	9,646	7,206	6,217	5,668	5,316	5,069	4,886	4,744	4,539	4,397	4,021	3,604
	12	9,330	6,927	5,953	5,412	5,064	4,821	4,640	4,499	4,296	4,155	3,780	3,362
	13	9,074	6,701	5,739	5,205	4,862	4,620	4,441	4,302	4,100	3,960	3,587	3,166
	14	8,862	6,515	5,564	5,035	4,695	4,456	4,278	4,140	3,939	3,800	3,427	3,005
	15	8,683	6,359	5,417	4,893	4,556	4,318	4,142	4,004	3,805	3,666	3,294	2,870
	16	8,531	6,226	5,292	4,773	4,437	4,202	4,026	3,890	3,691	3,553	3,181	2,754
	18	8,285	6,013	5,092	4,579	4,248	4,015	3,841	3,705	3,508	3,371	2,999	2,567
	20	8,096	5,849	4,938	4,431	4,103	3,871	3,699	3,564	3,368	3,231	2,859	2,422
	30	7,562	5,390	4,510	4,018	3,699	3,473	3,305	3,173	2,979	2,843	2,469	2,008
	40	7,314	5,178	4,313	3,828	3,514	3,291	3,124	2,993	2,801	2,665	2,288	1,806
	50	7,171	5,057	4,199	3,720	3,408	3,186	3,020	2,890	2,698	2,563	2,183	1,685
	70	7,011	4,922	4,074	3,600	3,291	3,071	2,906	2,777	2,585	2,450	2,067	1,542
	100	6,895	4,824	3,984	3,513	3,206	2,988	2,823	2,694	2,503	2,368	1,983	1,429
200	6,763	4,713	3,881	3,414	3,110	2,893	2,730	2,601	2,411	2,275	1,886	1,281	
∞	6,637	4,607	3,784	3,321	3,019	2,804	2,641	2,513	2,323	2,187	1,793		

Таблица 8 - Критические значения для критерия Кохрена

<i>p</i>	<i>n</i> = 2		<i>n</i> = 3		<i>n</i> = 4		<i>n</i> = 5		<i>n</i> = 6	
	1 %	5 %	1 %	5 %	1 %	5 %	1 %	5 %	1 %	5 %
2	-	-	0,995	0,975	0,979	0,939	0,959	0,906	0,937	0,877
3	0,993	0,967	0,942	0,871	0,883	0,798	0,834	0,746	0,793	0,707
4	0,968	0,906	0,864	0,768	0,781	0,684	0,721	0,629	0,676	0,590
5	0,928	0,841	0,788	0,684	0,696	0,598	0,633	0,544	0,588	0,506
6	0,883	0,781	0,722	0,616	0,626	0,532	0,564	0,480	0,520	0,445
7	0,838	0,727	0,664	0,561	0,568	0,480	0,508	0,431	0,466	0,397
8	0,794	0,680	0,615	0,516	0,521	0,438	0,463	0,391	0,423	0,360
9	0,754	0,638	0,573	0,478	0,481	0,403	0,425	0,358	0,387	0,329
10	0,718	0,602	0,536	0,445	0,447	0,373	0,393	0,331	0,357	0,303
11	0,684	0,570	0,504	0,417	0,418	0,348	0,366	0,308	0,332	0,281
12	0,653	0,541	0,475	0,392	0,392	0,326	0,343	0,288	0,310	0,262
13	0,624	0,515	0,450	0,371	0,369	0,307	0,322	0,271	0,291	0,243
14	0,599	0,492	0,427	0,352	0,349	0,291	0,304	0,255	0,274	0,232
15	0,575	0,471	0,407	0,335	0,332	0,276	0,288	0,242	0,259	0,220
16	0,553	0,452	0,388	0,319	0,316	0,262	0,274	0,230	0,246	0,208
17	0,532	0,434	0,372	0,305	0,301	0,250	0,261	0,219	0,234	0,198
18	0,514	0,418	0,356	0,293	0,288	0,240	0,249	0,209	0,223	0,189
19	0,496	0,403	0,343	0,281	0,276	0,230	0,238	0,200	0,214	0,181
20	0,480	0,389	0,330	0,270	0,265	0,220	0,229	0,192	0,205	0,174
21	0,465	0,377	0,318	0,261	0,255	0,212	0,220	0,185	0,197	0,167
22	0,450	0,365	0,307	0,252	0,246	0,204	0,212	0,178	0,189	0,160
23	0,437	0,354	0,297	0,243	0,238	0,197	0,204	0,172	0,182	0,155
24	0,425	0,343	0,287	0,235	0,230	0,191	0,197	0,166	0,176	0,149
25	0,413	0,334	0,278	0,228	0,222	0,185	0,190	0,160	0,170	0,144
26	0,402	0,325	0,270	0,221	0,215	0,179	0,184	0,155	0,164	0,140
27	0,391	0,316	0,262	0,215	0,209	0,173	0,179	0,150	0,159	0,135
28	0,382	0,308	0,255	0,209	0,202	0,168	0,173	0,146	0,154	0,131
29	0,372	0,300	0,248	0,203	0,196	0,164	0,168	0,142	0,150	0,127
30	0,363	0,293	0,241	0,198	0,191	0,159	0,164	0,138	0,145	0,124
31	0,355	0,286	0,235	0,193	0,186	0,155	0,159	0,134	0,141	0,120
32	0,347	0,280	0,229	0,188	0,181	0,151	0,155	0,131	0,138	0,117
33	0,339	0,273	0,224	0,184	0,177	0,147	0,151	0,127	0,134	0,114
34	0,332	0,267	0,218	0,179	0,172	0,144	0,147	0,124	0,131	0,111
35	0,325	0,262	0,213	0,175	0,168	0,140	0,144	0,121	0,127	0,108

p - количество лабораторий для данного уровня;

Проверка однородности дисперсий включает вычисление доли максимальной дис-

персии среди всех дисперсий: $G = \frac{S_{\max}^2}{\sum_i S_i^2}$, которая затем сравнивается с критическим значением $G(p, m, f)$, где f - число степеней свободы каждой дисперсии (должно быть одинаковым у всех дисперсий), m - число дисперсий, p - доверительная вероятность.

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

Таблица 9 Значения плотности стандартного нормального распределения

x	0	1	2	3	4	5	6	7	8	9
0,0	0,398942	0,398922	0,398862	0,398763	0,398623	0,398444	0,398225	0,397966	0,397668	0,397330
0,1	0,396953	0,396536	0,396080	0,395585	0,395052	0,394479	0,393868	0,393219	0,392531	0,391806
0,2	0,391043	0,390242	0,389404	0,388529	0,387617	0,386668	0,385683	0,384663	0,383606	0,382515
0,3	0,381388	0,380226	0,379031	0,377801	0,376537	0,375240	0,373911	0,372548	0,371154	0,369728
0,4	0,36827	0,366782	0,365263	0,363714	0,362135	0,360527	0,358890	0,357225	0,355533	0,353812
0,5	0,352065	0,350292	0,348493	0,346668	0,344818	0,342944	0,341046	0,339124	0,337180	0,335213
0,6	0,333225	0,331215	0,329184	0,327133	0,325062	0,322972	0,320864	0,318737	0,316593	0,314432
0,7	0,312254	0,310060	0,307851	0,305627	0,303389	0,301137	0,298872	0,296595	0,294305	0,292004
0,8	0,289692	0,287369	0,285036	0,282694	0,280344	0,277985	0,275618	0,273244	0,270864	0,268477
0,9	0,266085	0,263688	0,261286	0,258881	0,256471	0,254059	0,251644	0,249228	0,246809	0,24439
1,0	0,241971	0,239551	0,237132	0,234714	0,232297	0,229882	0,227470	0,22506	0,222653	0,220251
1,1	0,217852	0,215458	0,213069	0,210686	0,208308	0,205936	0,203571	0,201214	0,198863	0,196520
1,2	0,194186	0,19186	0,189543	0,187235	0,184937	0,182649	0,180371	0,178104	0,175847	0,173602

1,3	0,171369	0,169147	0,166937	0,164740	0,162555	0,160383	0,158225	0,15608	0,153948	0,151831
1,4	0,149727	0,147639	0,145564	0,143505	0,14146	0,139431	0,137417	0,135418	0,133435	0,131468
1,5	0,129518	0,127583	0,125665	0,123763	0,121878	0,120009	0,118157	0,116323	0,114505	0,112704
1,6	0,110921	0,109155	0,107406	0,105675	0,103961	0,102265	0,100586	0,098925	0,097282	0,095657
1,7	0,094049	0,092459	0,090887	0,089333	0,087796	0,086277	0,084776	0,083293	0,081828	0,08038
1,8	0,07895	0,077538	0,076143	0,074766	0,073407	0,072065	0,070740	0,069433	0,068144	0,066871
1,9	0,065616	0,064378	0,063157	0,061952	0,060765	0,059595	0,058441	0,057304	0,056183	0,055079
2,0	0,053991	0,052919	0,051864	0,050824	0,04980	0,048792	0,047800	0,046823	0,045861	0,044915
2,1	0,043984	0,043067	0,042166	0,041280	0,040408	0,039550	0,038707	0,037878	0,037063	0,036262
2,2	0,035475	0,034701	0,033941	0,033194	0,03246	0,031740	0,031032	0,030337	0,029655	0,028985
2,3	0,028327	0,027682	0,027048	0,026426	0,025817	0,025218	0,024631	0,024056	0,023491	0,022937
2,4	0,022395	0,021862	0,021341	0,020829	0,020328	0,019837	0,019356	0,018885	0,018423	0,017971
2,5	0,017528	0,017095	0,016670	0,016254	0,015848	0,015449	0,015060	0,014678	0,014305	0,01394
2,6	0,013583	0,013234	0,012892	0,012558	0,012232	0,011912	0,011600	0,011295	0,010997	0,010706
2,7	0,010421	0,010143	0,009871	0,009606	0,009347	0,009094	0,008846	0,008605	0,00837	0,00814
2,8	0,007915	0,007697	0,007483	0,007274	0,007071	0,006873	0,006679	0,006491	0,006307	0,006127

2,9	0,005953	0,005782	0,005616	0,005454	0,005296	0,005143	0,004993	0,004847	0,004705	0,004567
3,0	0,004432	0,004301	0,004173	0,004049	0,003928	0,003810	0,003695	0,003584	0,003475	0,00337
3,1	0,003267	0,003167	0,00307	0,002975	0,002884	0,002794	0,002707	0,002623	0,002541	0,002461
3,2	0,002384	0,002309	0,002236	0,002165	0,002096	0,002029	0,001964	0,001901	0,001840	0,001780
3,3	0,001723	0,001667	0,001612	0,001560	0,001508	0,001459	0,001411	0,001364	0,001319	0,001275
3,4	0,001232	0,001191	0,001151	0,001112	0,001075	0,001038	0,001003	0,000969	0,000936	0,000904
3,5	0,000873	0,000843	0,000814	0,000785	0,000758	0,000732	0,000706	0,000681	0,000657	0,000634
3,6	0,000612	0,00059	0,000569	0,000549	0,000529	0,000510	0,000492	0,000474	0,000457	0,000441
3,7	0,000425	0,000409	0,000394	0,000380	0,000366	0,000353	0,000340	0,000327	0,000315	0,000303
3,8	0,000292	0,000281	0,000271	0,000260	0,000251	0,000241	0,000232	0,000223	0,000215	0,000207
3,9	0,000199	0,000191	0,000184	0,000177	0,000170	0,000163	0,000157	0,000151	0,000145	0,000139
4,0	0,000134	0,000129	0,000124	0,000119	0,000114	0,000109	0,000105	0,000101	0,000097	0,000093

В таблице приведены значения плотности стандартного нормального распределения с параметрами $\mu=0$ (мат. ожидание) и $\sigma=1$ (ср. кв. отклонение). В Excel значения этой функции можно вычислить с помощью формулы =НОРМРАСП(x;0;1;0). Если рассматриваемое вами распределение отлично от стандартного ($\mu \neq 0$ или $\sigma \neq 1$), предварительно величину нужно нормализовать: $x^* = (x-\mu)/\sigma$, а потом уже смотреть в таблице, или воспользоваться формулой в Excel =НОРМРАСП(x;μ;σ;0).

Таблица 10 Значение функции Лапласа

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{z^2}{2}} dz .$$

x	Φ(x)	x	Φ(x)	x	Φ(x)	x	Φ(x)	x	Φ(x)	x	Φ(x)
0,00	0,00000	0,50	0,19146	1,00	0,34134	1,50	0,43319	2,00	0,47725	3,00	0,49865
0,01	0,00399	0,51	0,19497	1,01	0,34375	1,51	0,43448	2,02	0,47831	3,05	0,49886
0,02	0,00798	0,52	0,19847	1,02	0,34614	1,52	0,43574	2,04	0,47932	3,10	0,49903
0,03	0,01197	0,53	0,20194	1,03	0,34849	1,53	0,43699	2,06	0,48030	3,15	0,49918
0,04	0,01595	0,54	0,20540	1,04	0,35083	1,54	0,43822	2,08	0,48124	3,20	0,49931
0,05	0,01994	0,55	0,20884	1,05	0,35314	1,55	0,43943	2,10	0,48214	3,25	0,49942
0,06	0,02392	0,56	0,21226	1,06	0,35543	1,56	0,44062	2,12	0,48300	3,30	0,49952
0,07	0,02790	0,57	0,21566	1,07	0,35769	1,57	0,44179	2,14	0,48382	3,35	0,49960
0,08	0,03188	0,58	0,21904	1,08	0,35993	1,58	0,44295	2,16	0,48461	3,40	0,49966
0,09	0,03586	0,59	0,22240	1,09	0,36214	1,59	0,44408	2,18	0,48537	3,45	0,49972
0,10	0,03983	0,60	0,22575	1,10	0,36433	1,60	0,44520	2,20	0,48610	3,50	0,49977
0,11	0,04380	0,61	0,22907	1,11	0,36650	1,61	0,44630	2,22	0,48679	3,55	0,49981
0,12	0,04776	0,62	0,23237	1,12	0,36864	1,62	0,44738	2,24	0,48745	3,60	0,49984

0,13	0,05172	0,63	0,23565	1,13	0,37076	1,63	0,44845	2,26	0,48809	3,65	0,49987
0,14	0,05567	0,64	0,23891	1,14	0,37286	1,64	0,44950	2,28	0,48870	3,70	0,49989
0,15	0,05962	0,65	0,24215	1,15	0,37493	1,65	0,45053	2,30	0,48928	3,75	0,49991
0,16	0,06356	0,66	0,24537	1,16	0,37698	1,66	0,45154	2,32	0,48983	3,80	0,49993
0,17	0,06749	0,67	0,24857	1,17	0,37900	1,67	0,45254	2,34	0,49036	3,85	0,49994
0,18	0,07142	0,68	0,25175	1,18	0,38100	1,68	0,45352	2,36	0,49086	3,90	0,49995
0,19	0,07535	0,69	0,25490	1,19	0,38298	1,69	0,45449	2,38	0,49134	3,95	0,49996
0,20	0,07926	0,70	0,25804	1,20	0,38493	1,70	0,45543	2,40	0,49180	4,00	0,49997
0,21	0,08317	0,71	0,26115	1,21	0,38686	1,71	0,45637	2,42	0,49224	4,05	0,49997
0,22	0,08706	0,72	0,26424	1,22	0,38877	1,72	0,45728	2,44	0,49266	4,10	0,49998
0,23	0,09095	0,73	0,26730	1,23	0,39065	1,73	0,45818	2,46	0,49305	4,15	0,49998
0,24	0,09483	0,74	0,27035	1,24	0,39251	1,74	0,45907	2,48	0,49343	4,20	0,49999
0,25	0,09871	0,75	0,27337	1,25	0,39435	1,75	0,45994	2,50	0,49379	4,25	0,49999
0,26	0,10257	0,76	0,27637	1,26	0,39617	1,76	0,46080	2,52	0,49413	4,30	0,49999
0,27	0,10642	0,77	0,27935	1,27	0,39796	1,77	0,46164	2,54	0,49446	4,35	0,49999
0,28	0,11026	0,78	0,28230	1,28	0,39973	1,78	0,46246	2,56	0,49477	4,40	0,49999

0,29	0,11409	0,79	0,28524	1,29	0,40147	1,79	0,46327	2,58	0,49506	4,45	0,50000
0,30	0,11791	0,80	0,28814	1,30	0,40320	1,80	0,46407	2,60	0,49534	4,50	0,50000
0,31	0,12172	0,81	0,29103	1,31	0,40490	1,81	0,46485	2,62	0,49560	4,55	0,50000
0,32	0,12552	0,82	0,29389	1,32	0,40658	1,82	0,46562	2,64	0,49585	4,60	0,50000
0,33	0,12930	0,83	0,29673	1,33	0,40824	1,83	0,46638	2,66	0,49609	4,65	0,50000
0,34	0,13307	0,84	0,29955	1,34	0,40988	1,84	0,46712	2,68	0,49632	4,70	0,50000
0,35	0,13683	0,85	0,30234	1,35	0,41149	1,85	0,46784	2,70	0,49653	4,75	0,50000
0,36	0,14058	0,86	0,30511	1,36	0,41309	1,86	0,46856	2,72	0,49674	4,80	0,50000
0,37	0,14431	0,87	0,30785	1,37	0,41466	1,87	0,46926	2,74	0,49693	4,85	0,50000
0,38	0,14803	0,88	0,31057	1,38	0,41621	1,88	0,46995	2,76	0,49711	4,90	0,50000
0,39	0,15173	0,89	0,31327	1,39	0,41774	1,89	0,47062	2,78	0,49728	4,95	0,50000
0,40	0,15542	0,90	0,31594	1,40	0,41924	1,90	0,47128	2,80	0,49744	5,00	0,50000
0,41	0,15910	0,91	0,31859	1,41	0,42073	1,91	0,47193	2,82	0,49760		
0,42	0,16276	0,92	0,32121	1,42	0,42220	1,92	0,47257	2,84	0,49774		
0,43	0,16640	0,93	0,32381	1,43	0,42364	1,93	0,47320	2,86	0,49788		
0,44	0,17003	0,94	0,32639	1,44	0,42507	1,94	0,47381	2,88	0,49801		

0,45	0,17364	0,95	0,32894	1,45	0,42647	1,95	0,47441	2,90	0,49813		
0,46	0,17724	0,96	0,33147	1,46	0,42785	1,96	0,47500	2,92	0,49825		
0,47	0,18082	0,97	0,33398	1,47	0,42922	1,97	0,47558	2,94	0,49836		
0,48	0,18439	0,98	0,33646	1,48	0,43056	1,98	0,47615	2,96	0,49846		
0,49	0,18793	0,99	0,33891	1,49	0,43189	1,99	0,47670	2,98	0,49856		

Для отрицательных значений $\Phi(-X)=-\Phi(X)$. Для $x>5$ $\Phi(X)=0,5$. В Excel значение функции Лапласа можно вычислить с помощью формулы =НОРМСТРАСП(число) - 0,5. По своему логическому содержанию функция Лапласа $\Phi(X)$ близка к интегральной функции нормального распределения $F(X)$, их взаимосвязь: $\Phi(X) = F(X) - 0,5$

Учебное издание

Составители:

**Левшин Александр Григорьевич
Левшин Андрей Александрович
Бутузов Антон Евгеньевич
Майстренко Николай Александрович**

«Планирование и организация эксперимента»

Учебное пособие

Корректор *Т.Н. Куклева*

Компьютерная верстка, о-макет – *А.С Лаврова*

Обложка — *Н. В. Савина*

Подписано в печать 02.12.2015 г. Формат 60x84 1/8

Усл. печ. л. 3. Уч. изд. л. 2,5. Усл. Кр.-отт.

Тираж 100 экз. Изд. № . Зак. .

Издательство РГАУ-МСХА
127550, Москва, Тимирязевская ул. 44
Тел. 8 (499) 977-00-12; 977-40-64