

МАТЕМАТИЧЕСКИЕ МЕТОДЫ ОБРАБОТКИ ДАННЫХ

*Методические указания к лабораторным работам
для студентов бакалавриата направления 15.03.04*

**САНКТ-ПЕТЕРБУРГ
2020**

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное
учреждение высшего образования
Санкт-Петербургский горный университет

Кафедра автоматизации технологических процессов
и производств

МАТЕМАТИЧЕСКИЕ МЕТОДЫ ОБРАБОТКИ ДАННЫХ

*Методические указания к лабораторным работам
для студентов бакалавриата направления 15.03.04*

САНКТ-ПЕТЕРБУРГ
2020

УДК 519.242 (073)

МАТЕМАТИЧЕСКИЕ МЕТОДЫ ОБРАБОТКИ ДАННЫХ: Методические указания к лабораторным работам / Санкт-Петербургский горный университет. Сост.: *Н.В. Васильева, А.В. Бойков*. СПб. 2019. 76 с.

Изложены наиболее распространенные количественные и качественные методы обработки производственных и интерпретации результатов исследований.

Лабораторные работы позволяют студентам глубже освоить соответствующие разделы лекционного курса, знакомят студентов с практикой нахождения аппроксимирующих зависимостей с использованием различных методов и оценкой адекватности найденных зависимостей объекту исследования.

Методические указания предназначены для студентов бакалавриата направления 15.03.04 «Автоматизация технологических процессов и производств», а также могут быть использованы студентами других специальностей и слушателями курсов повышения квалификации при изучении основ планирования экспериментов и математической обработки производственной информации.

Научный редактор проф. *В.Ю. Бажин*

Рецензент канд. техн. наук *В.В. Васильев* (ООО «ТОМС инжиниринг»)

© Санкт-Петербургский
горный университет, 2020

МАТЕМАТИЧЕСКИЕ МЕТОДЫ ОБРАБОТКИ ДАННЫХ

***Методические указания к лабораторным работам
для студентов бакалавриата направления 15.03.04***

Сост.: *Н.В. Васильева, А.В. Бойков*

Печатается с оригинал-макета, подготовленного кафедрой
автоматизации технологических процессов и производств

Ответственный за выпуск *Н.В. Васильева*

Лицензия ИД № 06517 от 09.01.2002

Подписано к печати 29.01.2020. Формат 60×84/16.
Усл. печ. л. 4,4. Усл.кр.-отт. 4,4. Уч.-изд.л. 3,5. Тираж 75 экз. Заказ 50. С 17.

Санкт-Петербургский горный университет
РИЦ Санкт-Петербургского горного университета
Адрес университета и РИЦ: 199106 Санкт-Петербург, 21-я линия, 2

ВВЕДЕНИЕ

При исследовании технических систем могут использоваться теоретические и эмпирические методы познания. Каждое из этих направлений обладает относительной самостоятельностью, имеет свои достоинства и недостатки. В общем случае, теоретические методы в виде математических моделей позволяют описывать и объяснять взаимосвязи элементов изучаемой системы или объекта в относительно широких диапазонах изменения переменных величин. Однако при построении теоретических моделей неизбежно введение каких-либо ограничений, допущений, гипотез и т.п. Поэтому возникает задача оценки адекватности (достоверности) полученной модели реальному процессу или объекту. Для этого проводится экспериментальная проверка разработанных теоретических моделей.

Практика является решающей основой научного познания. В ряде случаев именно результаты экспериментальных исследований дают толчок к теоретическому обобщению изучаемого явления. Экспериментальное исследование дает более точное соответствие между изучаемыми параметрами. Однако не следует преувеличивать результаты экспериментальных исследований, которые справедливы только в пределах условий проведенного эксперимента.

Таким образом, теоретические и экспериментальные исследования дополняют друг друга и являются составными элементами процесса познания окружающего нас мира.

Как правило, результаты экспериментальных исследований нуждаются в определенной математической обработке. В настоящее время процедура обработки экспериментальных данных достаточно хорошо формализована, и исследователю необходимо только ее правильно использовать. Круг вопросов, решаемых при обработке результатов эксперимента, не так уж велик. Это вопросы подбора эмпирических формул и оценка их параметров, вопросы оценки истинных значений измеряемых величин и точности измерений, вопросы исследования корреляционных зависимостей и некоторые другие.

ЛАБОРАТОРНАЯ РАБОТА № 1. ОСНОВЫ РАБОТЫ В MATHCAD

В последнее время широкое распространение получили пакеты математических программ (или математические системы), которые можно использовать для различных вычислений (Mathematica, Derive, Statistica, MathCAD, MatLAB и др.). В этих системах процесс вычислений сильно автоматизирован, что позволяет экономить время и больше внимания уделять физическому смыслу получаемого результата. Выбор системы зависит от характера решаемых задач, от вкуса, от практики.

Mathcad чрезвычайно прост в использовании и легок в обучении. Большинство действий, необходимых для управления программой, являются интуитивно понятными.

Система Mathcad обладает следующими особенностями:

1. везде используется привычный способ математической записи. Если существует общепринятый способ изображения уравнения или математической операции, то Mathcad использует его;

2. используется принцип "То, что Вы видите, это то, что Вы получаете" (What you see is what you get). Не существует никакой скрытой информации, все показывается на экране. Результат вывода на печать выглядит в точности так же, как на экране дисплея;

3. простые выражения набираются на клавиатуре с использованием стандартных клавиш. Для специальных операторов (знаков сумм, интегралов, матриц и т.д.) предусмотрены специальные палитры;

4. удобная справочная система. Отметив указателем оператор, функцию или сообщение об ошибке и нажав [F1], можно отобразить на экране поясняющую информацию из справочной системы. Справки содержат пошаговые разъяснения по конкретной теме и иллюстрирующие примеры.

1. ОКНО MATHCAD

Двойной щелчок на пиктограмме Mathcad вызывает появление рекламной заставки, а через некоторое время окна Mathcad, показанного на рис. 1.1.

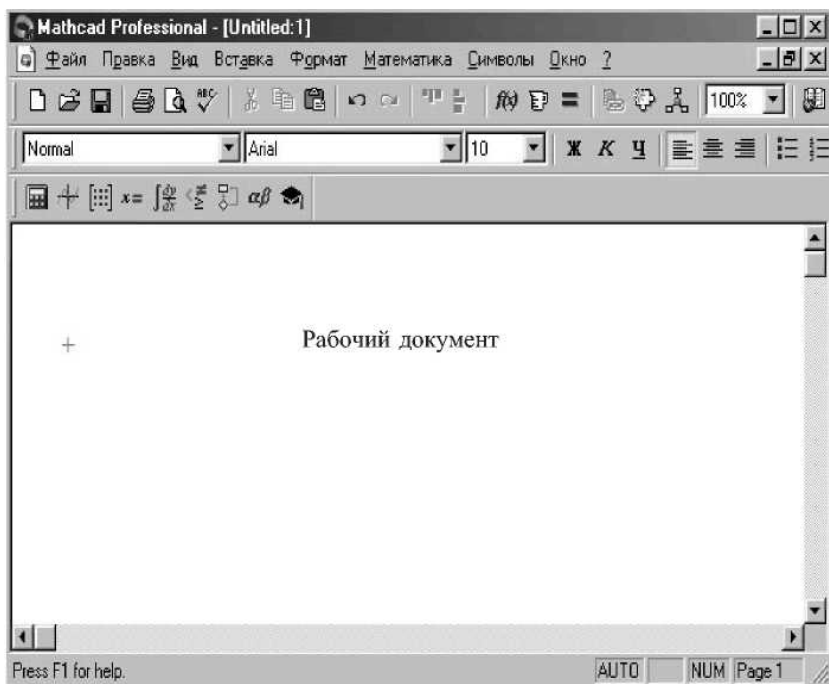


Рис. 1.1. Окно Mathcad

Для использования кнопок палитр символов нужно установить курсор в выбранное место рабочего документа и щелкнуть левой кнопкой мыши. В рабочем документе появится небольшой крестик. Затем установить курсор на нужной кнопке палитр символов (рис. 1.2) и снова нажать левую кнопку мыши и выбрать нужный элемент (знаки равенства, отношения, двух- или трехмерный график, интеграл, программную структуру и т.д.). Выбранный элемент появится на месте крестика в рабочем документе.

Для работы предназначены следующие панели инструментов (рис. 1.2):

Calculator (арифметика) – для ввода чисел, знаков типичных математических операций и часть употребляемых стандартных функций;

Graph (графики) – для построения графиков;

Matrix (матрицы) – для ввода векторов, матриц и матричных операций;

Evaluation (вычисления) – для ввода операторов вычисления;

Calculus (матанализ) – для задания операций, относящихся к математическому анализу;

Boolean (Булевы операторы) – для ввода знаков логических операций;

Programming (программирование), позволяющая вводить собственные программные конструкции;

Greek (греческий алфавит) – для ввода знаков логических операций;

Symbolic (символы) – для управления аналитическими преобразованиями.

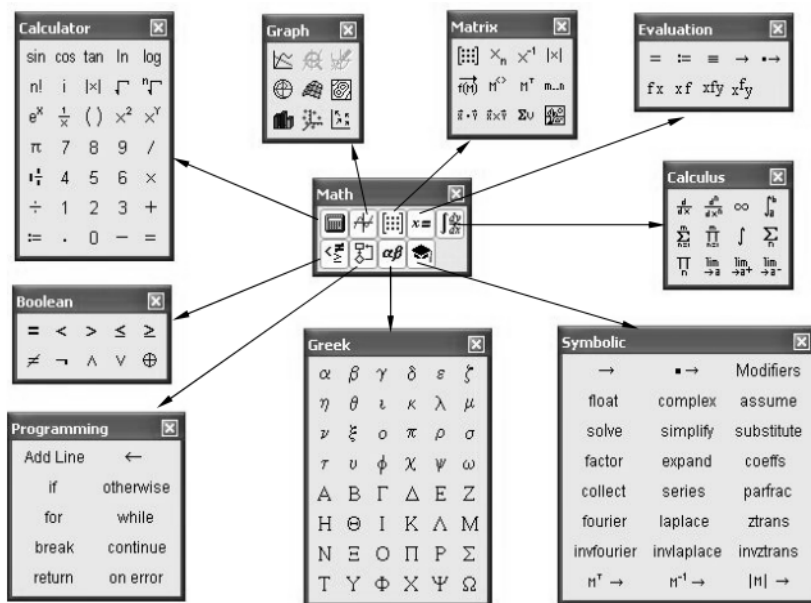


Рис. 1.2. Панели инструментов Mathcad для ввода формул

Ниже полосы кнопок палитр символов находятся кнопки панели инструментов, дублирующие основные команды меню. Если

указатель установить на кнопке, появляется текст, описывающий действие кнопки. Прямо под панелью инструментов располагается панель шрифтов, позволяющая изменять размер и другие характеристики шрифтов в формулах и тексте. В целях экономии места на экране каждая из этих компонент может быть выведена на экран, либо скрыта с помощью соответствующей команды из меню Окно.

2. ПРИМЕРЫ ПРОСТЫХ ДЕЙСТВИЙ

Алфавит системы Mathcad содержит: строчные и прописные латинские и греческие буквы; арабские цифры от 0 до 9; системные переменные; операторы; имена встроенных функций; специальные знаки; строчные и прописные буквы кириллицы (при работе с русифицированными документами); укрупненные элементы языка: типы данных, операторы, функции пользователя и управляющие структуры. К типам данных относятся числовые константы, обычные и системные переменные, массивы (векторы и матрицы) и данные файлового типа.

Числовые константы задаются с помощью арабских цифр, десятичной точки (а не запятой) и знака «-» (минус). Большинство вычислений система выполняет как с действительными, так и комплексными числами. Диапазон возможных значений чисел лежит в пределах от 10^{-307} до 10^{307} (машинный ноль и машинная бесконечность).

Имена переменных в системе Mathcad могут иметь практически любую длину, в них могут входить любые латинские и греческие буквы, а также цифры. Однако начинаться они могут только с буквы, например: x, x2, alpha, X. Кроме того, имя переменной не должно содержать пробелов. Строчные и прописные буквы в именах различаются.

Все формулы в Mathcad набираются только в латинском алфавите! Арифметические действия можно совершать, вводя числа и знаки операций с клавиатуры или с панели инструментов Calculator.

Если в процессе решения задачи происходит нарушение грамматики Mathcad, то место ошибки окрашивается в красный цвет и появляется разъяснение ошибки.

Рассмотрим некоторые примеры.

А. Вычислить логарифм натуральный от 25.

На панели инструментов Calculator нажмите кнопку «ln». На экране появится прямоугольная рамка, внутри которой расположен черный прямоугольник со знаком «ln». Введите с клавиатуры или с панели инструментов Calculator число 25. Нажав «=», мгновенно высветится ответ.

Аналогично вычисляются sin, cos, tg любого угла в радианах, десятичный логарифм log, модуль числа и др.

Б. Вычислить дробь.

$$\frac{5 \cdot |-6| + 2}{8^2}$$

При решении задачи знак модуля вводится с панели инструментов Calculator, а дробь – с клавиатуры.

$$\frac{5 \cdot |-6| + 2}{8^2} = 0.5$$

Этот пример демонстрирует особенности работы Mathcad. Mathcad отображает формулы в том виде, как их печатают в книгах – по всей площади экрана. Mathcad подбирает размеры для дробных черт, скобок и других математических символов, чтобы они выглядели на экране так, как их обычно пишут на бумаге.

Выражение на экране можно редактировать, устанавливая в нужном месте указатель и заменяя старые символы на новые. После установки указателя на свободное поле или другое выражение новый результат будет вычислен автоматически.

В. Вычислить значения функции $y=4x^2+5x+3$ для $x=1, 2, \dots, 10$.

Сначала задается диапазон значений x :

$$x := 1, 2 \dots 10$$

Здесь используется знак присвоения «:=», а не знак «=». При задании диапазона x набирается первое значение x , затем через запя-

тую второе его значение, чем задается шаг вычислений, и, наконец, последнее значение.

Две точки между 2 и 10 набираются нажатием кнопки **m..n** панели инструментов матрицы.

Затем набирается выражение для y (также через знак присвоения). Причем следует набрать $y(x)$, а не просто y .

$$y(x) := 4 \cdot x^2 + 5 \cdot x + 3$$

Mathcad выполняет команды слева направо и сверху вниз. Поэтому выражение для $y(x)$ должно быть расположено или справа, или ниже от выражения $x :=$.

После этого следует набрать « $x =$ » (равно, а не присвоить) и появится столбец со всеми значениями x . Также после нажатия « $y(x) =$ » ($y(x)$ равно) появится столбец вычисленных значений $y(x)$.

3. ГРАФИКИ

Почти ни один технический расчет не обходится без построения графиков. Графики являются удобным средством представления любой информации. Mathcad обладает обширным арсеналом средств для построения двумерных графиков в декартовых и полярных координатах, картин линий уровня, для изображения поверхностей и ряда других трехмерных графиков.

Рассмотрим создание простого двумерного графика функции $x(t) = \exp(-0.1t) \cdot \sin t$.

Для решения этой задачи необходимо сначала задать диапазон изменения величины t , потом набрать выражение для расчета функции $x(t)$, как это описано в предыдущем разделе. Обратите внимание, что при задании шага изменения аргумента, чем он будет меньше, тем более гладким получается график.

Чтобы создать график в Mathcad, нужно щелкнуть мышью на том свободном месте, где его нужно разместить, вызвать панель инструментов Graph (графики), нажать кнопку с изображением декартовых графиков. Появится пустой график с полями ввода для дан-

ных. В поле под серединой оси абсцисс нужно ввести имя переменной t . Теперь нужно щелкнуть в поле напротив середины оси ординат и ввести здесь $x(t)$. Остальные поля предназначены для ввода границ на осях – максимального и минимального значений, откладываемых на оси. Если оставить их пустыми Mathcad автоматически заполнит их при создании графика. После щелчка вне графика Mathcad вычисляет и строит точки графика, как показано на рис. 1.3.

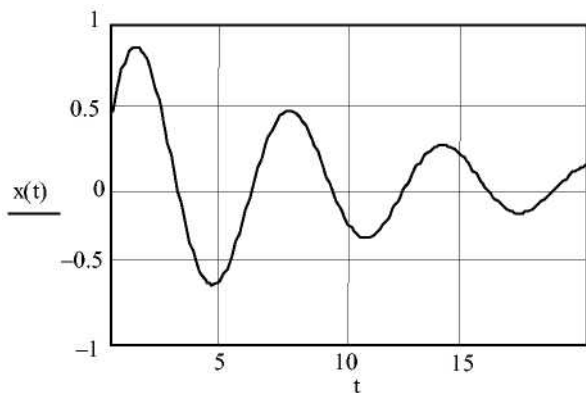


Рис. 1.3. Расчет и построение графика $x(t)$

Для вывода нескольких кривых на одном и том же чертеже необходимо ввести первое выражение для оси ординат, поставить запятую. Непосредственно под первым выражением появится пустое поле. Введите туда второе выражение, сопровождаемое запятой, и т.д. Максимально можно изобразить на одном графике 16 различных функций. Функции, построенные на одном графике, изображаются линиями различного цвета и типа. Цвета и типы линий можно изменить.

Для придания графику желаемого вида его следует отформатировать. Форматирование осуществляется следующим образом.

Дважды щелкнув по графику, вызовите соответствующее диалоговое окно, показанное на рис. 1.4.

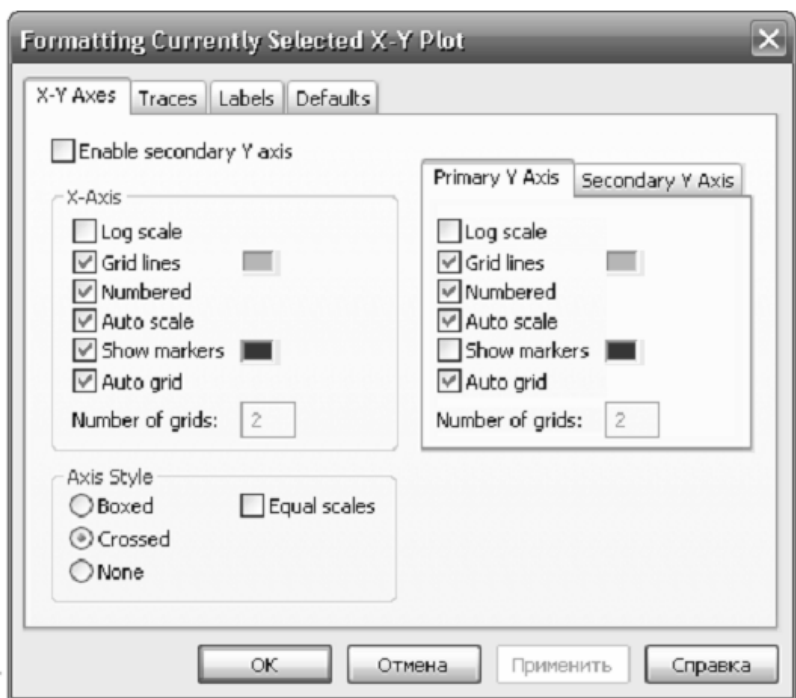


Рис.1.4. Первая страница окна форматирования графиков

В закладке «**X–Y Axes**» (**Оси X–Y**) расположен набор установок для каждой оси, позволяющий устанавливать равномерный или логарифмический масштаб, линии сетки и нумерацию линий сетки. Активизировать каждую установку можно, щелкнув левой кнопкой мыши по соответствующему квадратику. Об активизации установки свидетельствует появление в квадрате метки.

Установка «Auto grid» (Авто сетка) обеспечивает автоматический выбор числа интервалов сетки. Если квадратик не отмечен, можно установить число интервалов, набрав в поле «Number of grids» (Число интервалов) число от 2 до 99.

Наличие надписей «Enable secondary Y axis» (возможность второй оси) и «Secondary Y Axis» (вторая ось Y) дает возможность формировать графики различного масштаба для различных функций.

Нажав кнопку «Применить», можно увидеть перерисованный график, не закрывая диалоговое окно.

На рис. 1.5 представлена закладка «Traces» (След) окна форматирования графиков.

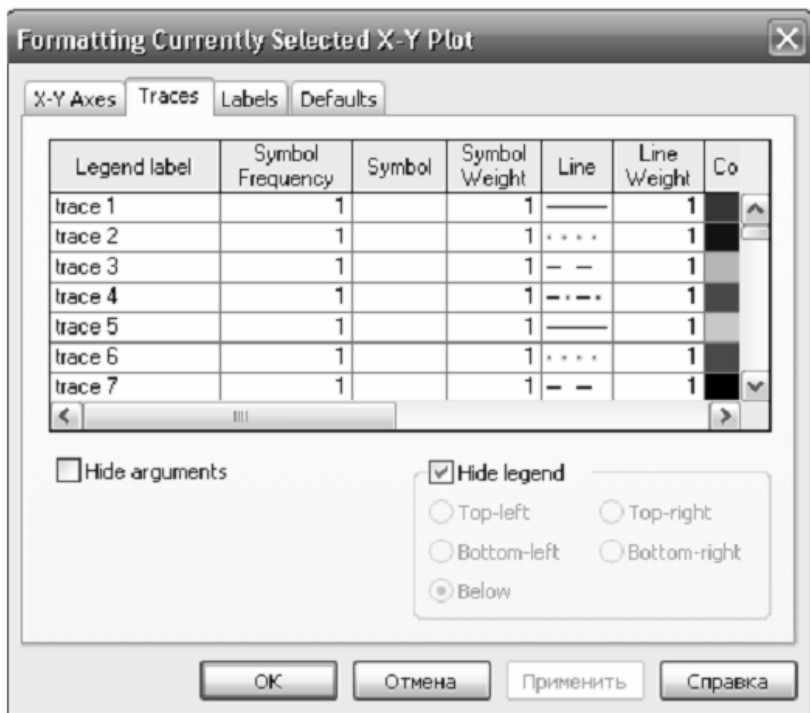


Рис.1.5. Вторая страница окна форматирования графиков

Из левого столбца (trace 1, trace 2 и т.д.) следует, что на одном графике можно наносить до 16 функций. Выбрать нужную кривую (если их на графике несколько) можно с использованием полосы прокрутки.

Чтобы изменить маркер кривой, тип линии, ее толщину, цвет, вид графика, необходимо нажать на стрелку около каждого соответствующего поля, чтобы увидеть раскрывающийся список свойств, и затем щелкнуть мышью на нужном месте.

Для подбора вида графика следует нажать «Применить», чтобы увидеть изменения, не закрывая диалоговое окно.

В закладке «**Labels**» (**Метки**) задается заголовок (Title), место его расположения (Above (сверху) или Below (снизу)), наименование осей (Axis Labels). Выбрав те ли иные требования к графику, можно закрыть диалоговое окно нажатием кнопки «ОК».

4. РЕШЕНИЕ СИСТЕМ ЛИНЕЙНЫХ И НЕЛИНЕЙНЫХ УРАВНЕНИЙ И НЕРАВЕНСТВ

Системы линейных и нелинейных уравнений и неравенств позволяет решать в Mathcad функция **Find** (Найти) в связке с ключевым словом **Given** (Дано). Конструкция **Given – Find** использует расчетную методику, основанную на поиске корня вблизи точки начального приближения, заданной пользователем.

В блоке **Given** записывается система уравнений и/или неравенств, подлежащих решению. Система уравнений и/или неравенств должна быть записана после или правее слова **Given**. Перед словом **Given** необходимо указывать начальные приближения для всех переменных.

Если мы хотим найти комплексный корень, следует задавать комплексное начальное приближение.

Признаком окончания системы служит функция **Find**.

Функция **Find** должна иметь столько же или меньше аргументов, сколько уравнений и неравенств содержит блок **Given**. Если окажется, что блок содержит слишком мало уравнений или неравенств, то его можно дополнить тождествами или повторяющимися выражениями.

В том случае, если решение не может быть найдено при заданном выборе начального приближения, появится сообщение в красной рамке **Did not find solution** – решение не найдено.

Рассмотрим некоторые примеры.

А. Решить уравнение $x^2 + 8 = e^x$.

Сначала необходимо задать начальное приближение. Затем ввести служебное слово **Given** и записать уравнение, используя знак **жирное равно**, которое применяется для записи уравнений и логических выражений (жирное (булево) равно, вводится с панели опе-

раторов Boolean). Написать функцию **Find** с неизвестной переменной в качестве параметра. В результате после знака равно выведется найденный корень. Если существует несколько корней, то их можно найти, меняя начальное приближение.

Листинг программы для решения этого уравнения приведен на рис. 1.6.

```
x := 2
given
x2 + 8 = ex
find(x) = 2.742
```

Рис. 1.6. Решение уравнения с помощью конструкции Given – Find

Б. Решить систему уравнений $\begin{cases} 3x_1 + 4x_2 = 5 \\ 5x_1 + 4x_2 = 3 \end{cases}$.

Решение данной системы с помощью вычислительной конструкции **Given – Find** приведено на рисунке 1.7.

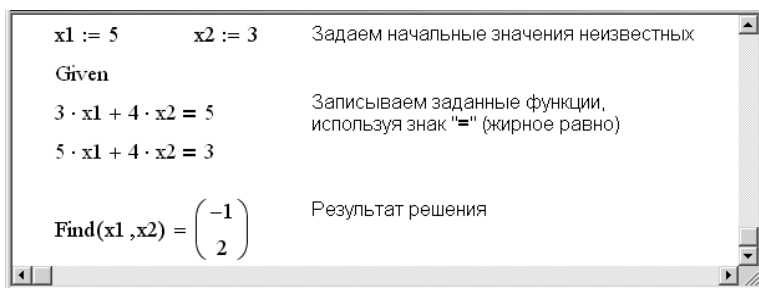


Рис. 1.7. Решение системы уравнений с помощью конструкции Given – Find


Совершенно аналогично решаются и системы неравенств.

5. ОПЕРАЦИИ С ВЕКТОРАМИ И МАТРИЦАМИ

Большим преимуществом системы Mathcad является возможность оперировать не только со скалярными величинами, но и с массивами. Mathcad поддерживает два вида массивов – одномерные (векторы) и двумерные (матрицы).

Элементами массива могут быть числа, константы, переменные, математические выражения и даже другие массивы. Соответственно массивы могут быть численными и символьными.

Матрице можно дать имя так же, как и скаляру.

Для ввода матриц и векторов необходимо нажать кнопку  на панели инструментов Matrix.

При этом появляется диалоговое окно (рис. 1.8), в котором надо задать число строк (Rows) и число столбцов (Columns) матрицы.

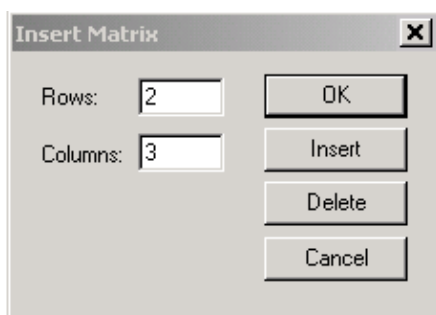


Рис. 1.8. Окно ввода матриц и векторов

Появится шаблон матрицы, в черные квадратики которого надо ввести значения элементов матрицы (рис. 1.9).

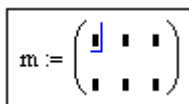


Рис. 1.9. Шаблон матрицы

Можно обращаться к отдельным элементам массива, используя нижние индексы. Например, имеется массив

$$V = \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}$$

Чтобы напечатать нижний индекс, используйте клавишу левой скобки. Например, напечатайте последовательно

$$V[0 = \quad V[1 = \quad V[2 =$$

В результате на экране появится

$$V_0 = 2 \quad V_1 = 3 \quad V_2 = 4$$

По умолчанию массивы Mathcad нумеруются с нулевого элемента.

Простейшие операции, которые можно проделывать с матрицами: сложение (вычитание), умножение на число, матричное перемножение.

Сложение и вычитание матриц

Складывать (вычитать) можно матрицы одного размера. При сложении (вычитании) матриц складываются (вычитаются) соответствующие элементы.

То есть при сложении матрицы A с элементами a_{ij} и матрицы B с элементами b_{ij} получается матрица C с элементами $c_{ij} = a_{ij} + b_{ij}$.

$$\begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} + \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 5 \end{pmatrix}$$

Умножение матрицы на число

При умножении матрицы $C = \{c_{ij}\}$ на число k все элементы матрицы умножаются на это число, то есть $c_{ij} = k \cdot a_{ij}$

$$v := \begin{pmatrix} 13 \\ -3 \\ 50 \end{pmatrix} \quad w := 2 \cdot v \quad w = \begin{pmatrix} 26 \\ -6 \\ 100 \end{pmatrix}$$

Основные операции для работы с векторами и матрицами собраны на панели инструментов Matrix: определение суммы эле-

ментов массива, вычисление определителя матрицы, скалярное и векторное умножение, транспонирование матриц.

Поэлементное умножение матриц

Поэлементное умножение матриц можно реализовать несколькими способами.


1-й способ основан на определении, что при умножении матрицы A с элементами a_{ij} и матрицы B с элементами b_{ij} получается матрица C с элементами $c_{ij} = a_{ij} \cdot b_{ij}$.

$$A := \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} \quad B := \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$$

$$C_i := A_i \cdot B_i$$

$$C = \begin{pmatrix} 0 \\ -3 \\ 4 \end{pmatrix}$$

При этом необходимо помнить, что предварительно необходимо задать диапазон изменения величины i (номер элемента массива).

2-й способ заключается в использовании специальной кнопки  на панели инструментов Matrix, позволяющей выполнять поэлементные операции с массивами. То есть $\overrightarrow{A \cdot B} = \{a_{ij} \cdot b_{ij}\}$

$$A := \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} \quad B := \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$$

$$C := \overrightarrow{A \cdot B}$$

$$C = \begin{pmatrix} 0 \\ -3 \\ 4 \end{pmatrix}$$

Кроме того в Mathcad имеется большое количество встроенных функций для действий над матрицами и векторами. Рассмотрим некоторые из них.

Вычисление максимального и минимального элемента матрицы или вектора производится с помощью встроенных функций **max(A)** и **min(A)**.

$$C := \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

$$\max(C) = 9$$

$$\min(C) = 1$$

Определение числа элементов в матрице:

$$V := \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}$$

$$\text{length}(V) = 3$$

Индекс последнего элемента:

$$V := \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}$$

$$\text{last}(V) = 2$$

СОДЕРЖАНИЕ ЗАДАНИЯ

В среде Mathcad решить и прокомментировать решение нижеследующих задач.

1. Постройте график функции $y(x) = \sin(ax + b) + \frac{4x + c}{5}$, если $a=2$; $b=3$; $c=4,1$.

2. Найдите значения функции $y(x) = 2x^2 + 3x - 1$ на промежутке $[-7; 3]$.

Постройте график функции $y(x)$. Примените для графика функции $y(x)$ следующее форматирование: пунктирная линия синего цвета.

Добавьте на эту же область построения график функции $g(x) = -4x + 2$. Примените для графика функции $g(x)$ следующее форматирование: штрихпунктирная линия черного цвета.

3. Вычислите $C = A + \frac{B}{1,3}$, если $A = \begin{pmatrix} 3 \\ 4,1 \\ 5 \end{pmatrix}$ и $B = \begin{pmatrix} 5,6 \\ 2,1 \\ 8 \end{pmatrix}$. Най-

дите сумму элементов массива C . Вычислите наибольшее и наименьшее значения элементов массива C .

4. Найдите вектор D , элементы которого получены почленным произведением соответствующих элементов векторов P и R ,

если $P = \begin{pmatrix} 12 \\ 2 \\ 5 \end{pmatrix}$; $R = \begin{pmatrix} 2 \\ 4 \\ 7 \end{pmatrix}$.

5. Решите систему уравнений:

$$\begin{cases} 5x + 2y = 7 \\ 2x - 4y = -2 \end{cases}$$

ЛАБОРАТОРНАЯ РАБОТА № 2. ОБРАБОТКА ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ ПО МЕТОДУ НАИМЕНЬШИХ КВАДРАТОВ

Цель работы – изучение практических приемов нахождения коэффициентов линейных и нелинейных регрессионных зависимостей и оценки точности аппроксимаций с использованием программной среды Mathcad.

ОСНОВЫ ТЕОРИИ И РАСЧЕТНЫЕ ФОРМУЛЫ

Метод наименьших квадратов является одним из наиболее распространенных и наиболее разработанных вследствие своей простоты и эффективности оценки параметров линейных моделей.

Рассмотрим процедуру оценки параметров линейной модели с помощью метода наименьших квадратов более подробно.

Пусть в результате эксперимента получена совокупность входных и выходных переменных в виде следующего набора:

x_{11}	x_{21}	...	x_{k1}	y_1
x_{12}	x_{22}	...	x_{k2}	y_2
.
.
.
x_{1N}	x_{2N}	...	x_{kN}	y_N

и задана функциональная зависимость в виде $y = \varphi(x_1, x_2, \dots, x_k)$.

Суть метода наименьших квадратов сводится к тому, чтобы сумма квадратов отклонений экспериментальных точек от сглаживающей кривой обращалась в минимум:

$$\sum_{i=1}^N [y_i - \varphi(x_i)]^2 = \min \quad (2.1)$$

где y_i и x_i – экспериментальные значения переменных в i -том опыте, N – число опытов, $\varphi(x)$ – искомая зависимость y от x .

Известно, что через N точек можно всегда провести кривую, аналитически выражаемую многочленом $(N-1)$ -й степени. Этот многочлен называют интерполяционным. А замену функции $y(x)$ на функцию $\varphi(x)$ так, что их значения совпадают в заданных точках $y(x_i) = \varphi(x_i); i=1,2,\dots,N$ называют интерполяцией.

Однако такое решение проблемы не является удовлетворительным, поскольку $y(x_i) = \varphi(x_i)$ из-за случайных ошибок измерения и влияния на измерения значений $y(x)$ различных помех.

Поэтому требуется провести кривую так, чтобы она в наименьшей степени зависела от случайных ошибок. Эта задача называется сглаживанием (аппроксимацией) экспериментальной зависимости и часто решается методом наименьших квадратов. Сглаживающую кривую называют аппроксимирующей.

Нахождение значений параметров, удовлетворяющих условию (2.1), сводится к отысканию экстремума функции многих переменных $a_j, j = \overline{0,s}$. Для этого необходимо взять производные по параметрам и приравнять их нулю, получив систему $s+1$ уравнений, решение которой дает возможность найти параметры a_j . Например при аппроксимации искомой зависимости полиномом S -того порядка $Y = a_0 + a_1X + a_2X^2 + \dots + a_sX^s$ система уравнений примет вид:

$$\begin{aligned}
 a_0 \cdot \sum_{i=1}^N x_i^0 + a_1 \cdot \sum_{i=1}^N x_i^1 + a_2 \cdot \sum_{i=1}^N x_i^2 + \dots + a_s \cdot \sum_{i=1}^N x_i^s &= \sum_{i=1}^N x_i^0 y_i \\
 a_0 \cdot \sum_{i=1}^N x_i^1 + a_1 \cdot \sum_{i=1}^N x_i^2 + a_2 \cdot \sum_{i=1}^N x_i^3 + \dots + a_s \cdot \sum_{i=1}^N x_i^{s+1} &= \sum_{i=1}^N x_i^1 y_i \\
 \dots & \\
 a_0 \cdot \sum_{i=1}^N x_i^s + a_1 \cdot \sum_{i=1}^N x_i^{s+1} + a_2 \cdot \sum_{i=1}^N x_i^{s+2} + \dots + a_s \cdot \sum_{i=1}^N x_i^{2s} &= \sum_{i=1}^N x_i^s y_i
 \end{aligned} \tag{2.2}$$

Эта система линейна относительно искомым параметрам a_j и носит название нормальной. Ее решение дает численные значения параметров a_j .

Пусть изучается зависимость одной физической величины Y от другой X . В результате опытов получены N пар чисел $(x_i; y_i)$. При этом зависимость $y = \varphi(x)$ может быть представлена в виде

$$y(x) = a_0 + a_1 \cdot x.$$

В соответствии с методом наименьших квадратов параметры этой зависимости должны удовлетворять условию

$$\sum_{i=1}^N [y_i - (a_0 + a_1 \cdot x_i)]^2 = \min \quad (2.3)$$

Взяв частные производные от выражения (2.3) по параметрам a_0 и a_1 и приравняв их нулю, получим систему нормальных уравнений:

$$\begin{aligned} a_0 \cdot \sum_{i=1}^N x_i^0 + a_1 \cdot \sum_{i=1}^N x_i &= \sum_{i=1}^N y_i \\ a_0 \cdot \sum_{i=1}^N x_i + a_1 \cdot \sum_{i=1}^N x_i^2 &= \sum_{i=1}^N y_i \cdot x_i \end{aligned} \quad (2.4)$$

Решая систему уравнений (2.4), находим неизвестные параметры a_0 и a_1 и тем самым полностью определяем функцию, которая наилучшим образом (в смысле наименьших квадратов отклонений расчетных значений от фактических данных) аппроксимирует (приближает) искомую функцию $y(x)$.

Как видно из рассмотренного примера, изменение количества параметров не приведет к искажению сущности самого подхода, изменится лишь количество уравнений в системе (2.4) (для N параметров соответственно будет записано N уравнений).

После вычисления коэффициентов аппроксимирующего полинома необходимо провести проверку пригодности найденного полинома изучаемому объекту. Такую проверку будем называть проверкой адекватности аппроксимирующего полинома. То есть, необходимо проверить, насколько результат вычислений соответствует данным проведенного эксперимента, и является ли найденная зависимость соответствующей изучаемому явлению. Проверка адекватности заключается в доказательстве факта, что точность результатов, полученных по модели, будет не хуже точности расчетов, произведенных на основании экспериментальных данных.

Адекватность аппроксимирующей функции определяется при помощи коэффициента детерминации, найденного как отношение суммы квадратов регрессии к общей сумме квадратов:

$$R^2 = \frac{SSf}{SSf + SSr} \quad (2.5)$$

где SSf – сумма квадратов регрессии, $SSf + SSr$ – общая сумма квадратов.

Здесь сумма квадратов регрессии находится по формуле:

$$SSf = \sum_{i=0}^{N-1} (y(x_i) - Y_{cp})^2 \quad (2.6)$$

где $y(x_i)$ – расчетное значение функции отклика, Y_{cp} – среднее значение исходных данных функции отклика.

Сумма квадратов остатков находится по формуле:

$$SSr = \sum_{i=0}^{N-1} (Y_i - y(x_i))^2 \quad (2.7)$$

где Y_i – исходные значения функции отклика.

Общая сумма квадратов находится по формуле:

$$SS = SSf + SSr = \sum_{i=0}^{N-1} (Y_i - Y_{cp})^2 \quad (2.8)$$

НАХОЖДЕНИЕ КОЭФФИЦИЕНТОВ АППРОКСИМИРУЮЩЕГО ПОЛИНОМА В СРЕДЕ MATHCAD

1. Использование встроенных функций *slope* и *intercept*, а также встроенной функции *line* для нахождения коэффициентов линейной зависимости

Применим метод наименьших квадратов для аппроксимации экспериментальных данных для нахождения коэффициентов линейной зависимости.

С этой целью удобно использовать встроенные функции *slope* и *intercept* для определения коэффициентов линейной регрессии (аппроксимация данных прямой линией).

Функция *slope* определяет угловой коэффициент прямой, а функция *intercept* – точку пересечения графика с осью ординат.

Для этого в командном окне Mathcad с клавиатуры следует набрать, предварительно задав исходные данные:

$$A := \text{intercept}(x, y) \qquad B := \text{slope}(x, y)$$

и вывести коэффициенты линейной регрессии –

$$A = \qquad B =$$

После набора знака равенства Mathcad вычисляет выражение и выводит результат.

Для этих же целей Mathcad предлагает также использовать встроенную функцию *line*:

$$\begin{pmatrix} b \\ a \end{pmatrix} := \text{line}(x, y)$$

После чего Mathcad вычисляет выражение, и можно вывести коэффициенты линейной регрессии –

$$a = \qquad b =$$

2. Нахождение коэффициентов линейной зависимости с использованием матричного решения

Матрицы дают возможность кратко записать систему линейных уравнений. Пусть дана система из 2-х уравнений с двумя неизвестными:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases}$$

Рассмотрим матрицу коэффициентов системы $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ и матрицы столбцы неизвестных и свободных членов $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ и $B = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$.

Найдем произведение

$$AX = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix}$$

т.е. в результате вычисления получаются левые части уравнений данной системы. Тогда, пользуясь определением равенства матриц, данную систему можно записать в виде:

$$\begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

или короче $AX = B$.

Здесь матрицы A и B известны, а матрица X неизвестна. Ее и нужно найти, т.к. ее элементы являются решением данной системы. Это уравнение называют матричным уравнением.

Пусть определитель матрицы отличен от нуля $|A| \neq 0$. Тогда матричное уравнение решается следующим образом. Умножим обе части уравнения слева на матрицу A^{-1} , обратную матрице A :

$$A^{-1} \cdot (AX) = A^{-1} \cdot B \quad \text{или} \quad (A^{-1} \cdot A)X = A^{-1} \cdot B.$$

Поскольку $A^{-1} \cdot A = E$ и $E \cdot X = X$ (E – единичная матрица), то получаем решение матричного уравнения в виде $X = A^{-1} \cdot B$.

Заметим, что поскольку обратную матрицу можно найти только для квадратных матриц, то матричным методом можно решать только те системы, в которых число уравнений совпадает с числом неизвестных.

Для использования данного метода в Mathcad необходимо предварительно задать исходные данные. Далее набрать с клавиатуры следующие команды для расчета матрицы коэффициентов системы A и матрицы столбцов свободных членов B :

$$i := 0..N - 1$$

$$m := 0..1$$

$$k := 0..1$$

$$A_{k,m} := \sum_{i=0}^{N-1} (X_i)^{k+m}$$

$$B_k := \sum_{i=0}^{N-1} Y_i \cdot (X_i)^k$$

После чего рассчитать коэффициенты прямой:

$$\begin{pmatrix} b \\ a \end{pmatrix} := A^{-1} \cdot B$$

и вывести их:

$$a = \qquad b =$$

3. Нахождение коэффициентов аппроксимирующей зависимости приведением ее к линейному виду

В ряде случаев в качестве зависимости берут функцию, в которую неопределенные коэффициенты входят нелинейно. При этом для упрощения расчетов иногда искомую функцию удается линеаризовать, т.е. свести к линейной.

К числу таких зависимостей относится, например, зависимость $y = A \cdot B^x$.

Прологарифмировав это выражение, получим:

$$\lg(y) = \lg(A) + x \cdot \lg(B)$$

Введем обозначения $z = \lg(y)$, $a = \lg(A)$ и $b = \lg(B)$. Тогда в новых обозначениях задача сводится к отысканию коэффициентов линейной функции $z = a + bx$.

Это нужно сделать для того, чтобы можно было воспользоваться встроенными функциями *slope* и *intercept*, которые определяют параметры линейной зависимости.

Находим

$$b := \text{slope}(x, z)$$

$$a := \text{intercept}(x, z)$$

Заметим, что аргументами функций *slope* и *intercept* являются введенные обозначения $z = \lg(y)$, $a = \lg(A)$ и $b = \lg(B)$.

Найденный коэффициент $a = \lg(A)$, отсюда $A := 10^a$. Аналогично $B := 10^b$. После определения коэффициентов A и B , по предлагаемой формуле можно найти $y(x) = A \cdot B^x$.

В табл. 2.1 приведены замены переменных для некоторых часто используемых нелинейных функций.

4. Нахождение коэффициентов аппроксимирующей зависимости с использованием функции линейного сглаживания *linfit*

При необходимости аппроксимации экспериментальных данных по методу наименьших квадратов линейной комбинацией произвольных функций в Mathcad рекомендуется использовать встроенную функцию *linfit*.

Функция *linfit* имеет три аргумента:

- вектор x – x -координаты заданных точек,
- вектор y – y -координаты заданных точек,

– функция F – содержит набор функций, который будет использоваться для построения линейной комбинации.

Таблица 2.1

Нелинейная функция	Результат логарифмирования	Замена переменных	Линейная функция
$y = a_0 \cdot x^{a_1}$	$\lg y = \lg a_0 + a_1 \lg x$	$z = \lg(y),$ $\xi = \lg(x),$ $b_1 = a_1,$ $b_0 = \lg(a_0)$	$z = b_0 + b_1 \xi$
$y = a_0 \cdot a_1^x$	$\lg y = \lg a_0 + x \lg a_1$	$z = \lg(y),$ $\xi = x,$ $b_0 = \lg(a_0),$ $b_1 = \lg(a_1)$	$z = b_0 + b_1 \xi$
$y = a_0 \cdot \exp(a_1 x)$	$\ln y = \ln a_0 + x a_1$	$z = \ln(y),$ $\xi = x,$ $b_0 = \ln(a_0),$ $b_1 = a_1$	$z = b_0 + b_1 \xi$

Например, необходимо найти функцию $y = a \frac{1}{x+1} + bx^2$.

Для нахождения неизвестных коэффициентов a и b сначала необходимо задать функцию множителей при соответствующих искомым коэффициентах. Для данной задачи функция F будет иметь вид:

$$F(x) := \left(\begin{array}{c} 1 \\ x+1 \\ x^2 \end{array} \right)$$

Результат использования встроенной функции **linfit** присвоим переменной s :

$$s := \text{linfit}(x, y, F)$$

После выведения рассчитанных коэффициентов командой:

s =

можно найти аппроксимирующую функцию $y(x) := F(x) \cdot s$.

5. Нахождение коэффициентов аппроксимирующей зависимости с использованием функции *minimize*

Нахождение коэффициентов аппроксимирующей функции также можно осуществить с использованием встроенной функции *minimize*, в которой реализуются алгоритмы оптимизации, основывающиеся на итерационных вычислениях и последовательном приближении к точке минимума.

Пусть необходимо найти функцию $y = a \frac{1}{x+1} + bx^2$.

Для поиска оптимальных (обеспечивающих минимум квадрата отклонения) значений аппроксимирующей функции удобно использовать встроенную функцию *minimize*:

Minimize (f, a₁, ..., a_n) – вектор значений аргументов, при которых функция **f** достигает минимума,

где:

f – функция, которую следует минимизировать;

a₁, ..., a_n – аргументы, по которым производится минимизация.

Всем аргументам функции **f** предварительно следует присвоить некоторые значения, причем для тех переменных, по которым производится минимизация, они будут восприниматься как начальные приближения, например:

a := 1 b := 1

Далее необходимо набрать с клавиатуры функцию, которую необходимо оптимизировать. В данном случае:

$$f(a,b) := \sum_{i=0}^{N-1} \left[y_i - \left(a \cdot \frac{1}{x_i + 1} + b \cdot x_i^2 \right) \right]^2$$

После этого необходимо задать ограничения для искомых коэффициентов аппроксимирующей функции, например:

Given

$$-100 < a < 100$$

$$-100 < b < 100$$

Введя команду на минимизацию функции $f(a, b)$:

$$\begin{pmatrix} a \\ b \end{pmatrix} := \text{minimize}(f, a, b)$$

можно рассчитать коэффициенты зависимости и вывести их командой:

$$a = \qquad b =$$

СОДЕРЖАНИЕ ЗАДАНИЯ

По данным эксперимента, приведенного в таблице 2.2:

1. Провести линейную аппроксимацию зависимости $y = f(x)$ предложенным способом.

2. Выполнить проверку адекватности линейной регрессионной модели и нанести на график аппроксимирующий полином и экспериментальные данные.

3. Провести нелинейную аппроксимацию зависимости $y = f(x)$ предложенным способом.

4. Выполнить оценку адекватности нелинейной регрессионной модели и нанести на график аппроксимирующий полином и экспериментальные данные.

5. Сделать вывод: какая из предложенных видов зависимостей лучшим образом описывает экспериментальные данные.

Таблица 2.2

x_i	Таблица данных $y=f(x)$					Способ линейной аппроксимации	Нелинейные зависимости	
	0,1	0,5	1	2	4		Вид функции	Способ нелинейной аппроксимации
1	2	3	4	5	6	7	8	9
1	-0,1	0,16	0,27	0,38	0,49	a, b	$y = a+bx+cx^2$ $y=a+b \cdot \exp(x)$	e $ж$
2	0,98	0,60	0,33	0,10	0,01	$a, в$	$y = ab^x$ $y=a \cdot \ln(bx)$	∂ $ж$
3	0,78	0,40	0,20	0,70	5,30	a, z	$y = a+bx+cx^2$ $y=a \cdot \exp(bx)$	e $ж$
4	0,20	2,76	2,99	3,10	3,15	a, \bar{b}	$y = ab^x$ $y=a \cdot \ln(bx)$	∂ $ж$
5	0,34	0,26	0,18	0,09	0,02	$a, в$	$y = a-bx^{-1}+cx^{-2}$ $y = a+bx+cx^2$	e $ж$
6	0,06	1,35	1,90	2,45	3,01	a, z	$y = ab^{-x}$ $y=a+b \cdot \ln(x)$	∂ $ж$
7	0,78	0,43	0,21	0,05	0,00	a, b	$y = a-bx^{-1}-cx^{-2}$ $y =$ $= a+bx+cx^2$	e $ж$
8	1,53	1,69	1,77	1,86	1,95	$a, в$	$y = ab^x$ $y=a+b \cdot \ln(cx)$	∂ $ж$
9	0,78	0,72	0,65	0,54	0,36	a, z	$y = a-bx-cx^2$ $y = a+bsinx$	e $ж$
10	0,05	1,13	2,99	4,90	6,28	a, b	$y = ab^{-x}$ $y=a+b \cdot \exp(x)$	∂ $ж$
11	3,70	3,69	3,68	3,68	3,74	$a, в$	$y = a-bx-cx^2$ $y =$ $a+bx+cx^2$	e $ж$
12	2,59	3,87	4,60	5,47	6,51	a, z	$y = ab^{-x}$ $y=a+b \cdot \ln(cx)$	∂ $ж$
13	8,43	7,98	7,75	7,62	7,60	a, \bar{b}	$y = a+bx+cx^2$ $y=a+b \cdot \exp(-1/x)$	e $ж$
14	5,82	4,22	2,83	1,27	0,26	$a, в$	$y = ab^{-x}$ $y=a \cdot \ln(-bx)$	∂ $ж$
15	1,42	0,88	0,48	0,14	0,01	a, z	$y = a-bx^{-1}+cx^{-2}$ $y=a+b \cdot \ln(-cx)$	e $ж$
16	0,04	3,42	4,88	6,33	7,79	a, b	$y = ab^{-x}$ $y=a \cdot \ln(bx)$	∂ $ж$

1	2	3	4	5	6	7	8	9
17	6,58	6,18	5,90	6,10	9,50	<i>a, в</i>	$y = a+bx+cx^2$ $y = a+x^b$	<i>e</i> <i>ж</i>
18	0,85	1,08	1,20	1,33	1,48	<i>a, з</i>	$y = ab^x$ $y = a+bx+cx^2$	<i>д</i> <i>ж</i>
19	0,59	0,62	0,67	0,73	0,77	<i>a, б</i>	$y = a+bx+cx^2$ $y=a+b \cdot \exp(x)$	<i>e</i> <i>ж</i>
20	0,13	0,20	0,35	1,01	8,49	<i>a, в</i>	$y = ab^{-x}$ $y = a+bx+cx^2$	<i>д</i> <i>ж</i>
21	0,62	0,40	0,23	0,08	0,01	<i>a, з</i>	$y = a+bx-cx^2$ $y=a \cdot \exp(-bx)$	<i>e</i> <i>ж</i>
22	0,06	1,35	1,90	2,45	3,01	<i>a, б</i>	$y = ab^x$ $y=a+b \cdot \ln(x)$	<i>д</i> <i>ж</i>
23	0,78	0,43	0,21	0,05	0,00	<i>a, в</i>	$y = a-bx^{-1}+cx^{-2}$ $y = a+bx+cx^2$	<i>e</i> <i>ж</i>
24	1,53	1,69	1,77	1,86	1,95	<i>a, з</i>	$y = ab^x$ $y=a+b \cdot \ln(cx)$	<i>д</i> <i>ж</i>
25	0,78	0,72	0,65	0,54	0,36	<i>a, б</i>	$y = a-bx-cx^2$ $y=a \cdot \exp(-bx)$	<i>e</i> <i>ж</i>
26	0,05	1,13	2,99	4,90	6,28	<i>a, в</i>	$y = ab^x$ $y=a+b \cdot \exp(-1/x)$	<i>д</i> <i>ж</i>
27	3,70	3,69	3,68	3,68	3,74	<i>a, з</i>	$y = a-bx-cx^2$ $y = a+bx+cx^2$	<i>e</i> <i>ж</i>
28	0,85	1,08	1,20	1,33	1,48	<i>a, б</i>	$y = ab^{-x}$ $y = ax^b$	<i>д</i> <i>ж</i>
29	0,59	0,62	0,67	0,73	0,77	<i>a, в</i>	$y = a+bx^{-1}-cx^2$ $y=a \cdot \ln(bx)$	<i>e</i> <i>ж</i>
30	0,13	0,20	0,35	1,01	8,49	<i>a, з</i>	$y = ab^x$ $y = ax^b$	<i>д</i> <i>ж</i>

Способы линейной аппроксимации:

- a)* решением системы линейных уравнений с использованием конструкции *Given – Find*;
- б)* используя матричное решение;
- в)* используя функцию *line*;
- з)* используя встроенные функции *slope* и *intercept*.

Способы нелинейной аппроксимации:

- д)* приведением зависимости к линейному виду;
- е)* используя функцию линейного сглаживания *linfit*;
- ж)* используя функцию *minimize*.

ЛАБОРАТОРНАЯ РАБОТА № 3. ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТА

Цель работы – изучение приемов планирования эксперимента и обработки экспериментальных данных; нахождение параметров линейного регрессионного уравнения первой степени с учетом эффекта парного взаимодействия по результатам эксперимента.

ОСНОВЫ ТЕОРИИ И РАСЧЕТНЫЕ ФОРМУЛЫ

Планирование эксперимента состоит в выборе числа и условий проведения опытов, позволяющих получить необходимые знания об объекте исследования с требуемой точностью.

Важнейшим условием научно поставленного эксперимента является минимизация общего числа опытов, а следовательно, материальных и трудовых затрат и временных ресурсов, что, конечно, не должно существенно отражаться на качестве полученной информации.

Необходимо отметить еще одно важное обстоятельство. Применение методов планирования эксперимента ограничивается сложностью или невозможностью постановки экспериментов в реальных условиях. Однако методы моделирования позволяют проводить с помощью ЭВМ различные эксперименты с моделями объектов исследования. Принципиальных различий в планировании натурных экспериментов и исследований на ЭВМ нет, но в последнем случае больше возможностей изменять и стабилизировать любые переменные, которые учтены в модели. Функционирование реальной системы, связанное со случайным проявлением действия некоторых факторов, легко имитируется в машинных экспериментах с помощью специально разработанных программ.

Рассмотрим сущность планирования эксперимента по этапам:

1. Выбор входных и выходных переменных. Входные переменные X_i , $i = \overline{1, k}$, которые определяют состояние объекта, назовем влияющими факторами. Основное требование к ним – достаточная управляемость, под которой понимается возможность установить

нужный уровень фактора и стабилизировать его в течение всего опыта.

Выходная переменная Y – это реакция объекта на входные воздействия; она носит название функции отклика. Выбор этой функции определяется целью исследования, которая может представлять собой оптимизацию экономической (стоимость, производительность), технологической (точность, качество, быстродействие), конструктивной (габариты, надежность) или другой характеристики объекта.

2. *Выбор области экспериментирования*, т.е. области факторного пространства, изучение которой представляет интерес для исследования. Границы этой области по каждому фактору X_i обусловлены его минимальным и максимальным значениями, т.е. $X_{i \min} < X_i < X_{i \max}$. Область экспериментирования для случая учета двух факторов (X_1 и X_2) иллюстрируется рис. 3.1. При трех факторах такая область представляет собой параллелепипед. При большем числе факторов она ограничивается гиперплоскостями в k -мерном пространстве.

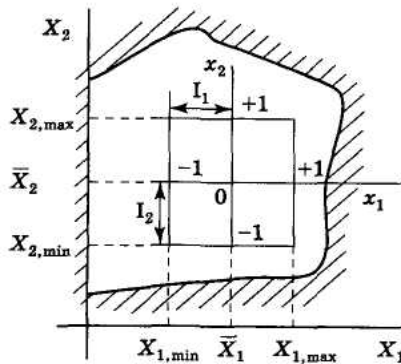


Рис. 3.1. Область экспериментирования для двух факторов

Оценка границ области определения (существования) факторов производится на основе принципиальных ограничений либо из других соображений.

Первый вид ограничений не может быть нарушен ни при каких обстоятельствах. Например, для температуры нижним пределом всегда будет абсолютный нуль.

При выборе ограничений второго вида исследователь руководствуется конкретными обстоятельствами, например: временем протекания процесса, стоимостью материала и т.п. Устанавливая область определения, необходимо выполнять также условие совместности факторов, т.е. значения факторов должны быть выбраны так, чтобы эксперимент можно было реализовать.

3. *Выбор математической модели объекта.* Если аналитическую зависимость, связывающую функцию отклика Y с влияющими факторами X_i , найти невозможно и вид функции $Y = f(X_1, X_2, \dots, X_k)$ априори неизвестен, то целесообразно использовать степенной ряд

$$Y = a_0 + \sum_{i=1}^k a_i X_i + \sum_{i < j} a_{ij} X_i X_j + \sum_{i=1}^k a_{ii} X_i^2 + \dots, \quad (3.1)$$

где k – число влияющих факторов.

Выражение (3.1) служит математической моделью исследуемого объекта. Так как, исходя из требований практики, число членов степенного ряда ограничивается, аппроксимирующая функция представляет собой полином некоторой степени.

Для определения коэффициентов аппроксимирующего полинома применяется наиболее универсальный метод наименьших квадратов. Как отмечалось выше, при его использовании необходимым условием получения статистических оценок является выполнение неравенства $N > s$, т.е. количество опытов N должно быть больше, чем число коэффициентов полинома s . Увеличить N можно повторением опыта в исходных точках эксперимента либо увеличением количества этих точек. Второй путь дает возможность не только учесть погрешности измерения, но и оценить адекватность аппроксимирующего полинома во всей области экспериментирования.

Для удобства обработки и интерпретации результатов эксперимента целесообразно все факторы представить в безразмерной форме, для чего производят операцию кодирования переменных. Ее сущность заключается в том, что начало координат факторного про-

пространства переносится в точку с координатами \bar{X}_i (рис. 3.1, центр эксперимента – точка 0), где $\bar{X}_i = 0,5(X_{i \max} + X_{i \min})$. Кроме того, интервал варьирования факторов $I_i = 0,5(X_{i \max} - X_{i \min})$ разбивается на ряд уровней, симметричных относительно центра эксперимента. В случае составления симметричных двухуровневых планов все k факторов изменяются на двух уровнях, при этом значениям $X_{i \max}$ отвечает кодированная переменная $x_i = +1$, а значениям $X_{i \min}$ соответствует $x_i = -1$. Для количественных факторов связь между физическими (X_i) и кодированными (x_i) значениями факторов определяется соотношением $x_i = (X_i - \bar{X}_i) / I_i$.

Описанные преобразования являются линейными. Поэтому в аппроксимирующей функции (3.1) изменяются только коэффициенты при факторах, т.е.

$$Y = b_0 + \sum_{i=1}^k b_i x_i + \sum_{i < j} b_{ji} x_i x_j + \sum_{i=1}^k b_{ji} x_i^2 + \dots, \quad (3.2)$$

где x_i – влияющие факторы в безразмерной форме (независимые переменные).

Заметим, что число членов полинома (3.2) при практической аппроксимации обычно ограничивается учетом линейного и квадратичного влияния факторов, а также эффекта парного их взаимодействия.

4. Составление плана эксперимента. Выбрав математическую модель объекта, определяют, какое значение должен принимать каждый из факторов в каждом из опытов. Таблица, составленная из значений факторов для каждого опыта, как независимых ($i = \overline{1, k}$) так и зависимых ($i = \overline{k+1, s-1}$), называется матрицей планирования, ее часть, которая включает в себя значения независимых переменных, и является планом эксперимента.

Эксперимент, в котором реализуются все возможные сочетания уровней факторов, называется полным факторным экспериментом (ПФЭ).

Если k факторов варьируются на двух уровнях, то число всех возможных сочетаний факторов равно 2^k . Тогда полный факторный эксперимент называется ПФЭ типа 2^k . Если число уровней факторов составляет n , необходим ПФЭ типа n^k .

В качестве примера рассмотрим составление плана ПФЭ 2^2 . Число опытов в этом случае $N = 2^2 = 4$. Соответствующая матрица планирования представлена таблице 3.1.

Таблица 3.1

Номер опыта	x_0	x_1	x_2	$x_3 = x_1x_2$	Y
1	+1	-1	-1	+1	Y_1
2	+1	+1	-1	-1	Y_2
3	+1	-1	+1	-1	Y_3
4	+1	+1	+1	+1	Y_4

Данный план соответствует модели вида

$$Y = b_0x_0 + b_1x_1 + b_2x_2 + b_{12}x_1x_2. \quad (3.3)$$

Первый столбец матрицы занимает фиктивный фактор $x_0 = +1$ при коэффициенте полинома b_0 . Столбцы матрицы x_1 и x_2 , обведенные рамкой (табл. 3.1), задают планирование: по ним непосредственно определяются условия опытов; столбец x_3 является самостоятельным и заполняется по данным столбцов x_1 и x_2 . Он, как и столбец x_0 , используется при расчетах.

По результатам эксперимента, проведенного в соответствии с представленным планом, можно определить все четыре коэффициента полинома (3.3). Однако в этом случае $s = 4$ и, следовательно, условие $N > s$ не выполняется, что не позволяет произвести статистических оценок аппроксимирующей зависимости. Для получения таких оценок нужно ограничиться линейной зависимостью без учета взаимодействия факторов (при этом $s = 3 < N = 4$) либо провести до-

полнительный опыт в нулевой точке $x_1 = x_2 = 0$ (тогда $s = 4 < N = 5$).

Если ошибка опыта априори неизвестна, для ее оценки опыты в каждой точке плана нужно повторить.

В общем случае ПФЭ типа 2^k обладает следующими свойствами:

– *симметричностью относительно центра эксперимента*; при этом алгебраическая сумма элементов вектора-столбца для каждого фактора равна нулю:

$$\sum_{u=1}^N x_{iu} = 0, \quad i = \overline{1, k}, \quad (3.4)$$

где x_{iu} – значение i -го фактора в u -м опыте;

– *соответствием условиям нормировки* (сумма квадратов элементов каждого столбца равна числу опытов):

$$\sum_{u=1}^N x_{iu}^2 = N, \quad i = \overline{0, k} \quad (3.5)$$

Это условие следует из того обстоятельства, что значения факторов в матрице задаются в кодированном виде $(-1; +1)$;

– *условиям ортогональности*, при этом сумма почленных произведений любых двух векторов-столбцов матрицы равна нулю:

$$\sum_{u=1}^N x_{iu} x_{ju} = 0, \quad i < j; \quad i, j = \overline{1, k} \quad (3.6)$$

Ортогональность матриц планирования позволяет при обработке данных с помощью метода наименьших квадратов получить независимые друг от друга оценки коэффициентов.

При выполнении условий (3.4) ÷ (3.6) коэффициенты аппроксимирующего полинома рассчитываются по формулам:

$$b_i = \frac{1}{N} \sum_{u=1}^N Y_u x_{iu}; \quad b_{ij} = \frac{1}{N} \sum_{u=1}^N Y_u x_{iu} x_{ju} \quad (3.7)$$

Заметим, что модель типа (3.3) включает только линейные эффекты и эффекты парного взаимодействия факторов. Планы, соответствующие этой модели, называются планами первого порядка. В случае необходимости учета нелинейного влияния фактора аппроксимирующий полином должен содержать члены более высокого порядка. В частности, модель с квадратичными членами для двух факторов (полная квадрика) имеет вид

$$Y = b_0x_0 + b_1x_1 + b_2x_2 + b_{12}x_1x_2 + b_{11}x_1^2 + b_{22}x_2^2. \quad (3.8)$$

Для оценки коэффициентов полинома вида (3.8) следует пользоваться более сложными планами, чем планы первого порядка. Это обусловлено тем обстоятельством, что помимо информации о коэффициентах при линейных членах и членах, учитывающих эффект взаимодействия, необходимо оценить коэффициенты b_i при квадратичных членах аппроксимирующего полинома. Планы этого типа называются планами второго порядка.

5. *Обработка результатов эксперимента.* Для обработки результатов эксперимента обычно пользуются следующей схемой:

1) На основании данных параллельных наблюдений оценивается дисперсия воспроизводимости D_{y_i} для каждой строки плана по формуле:

$$D_{y_i} = \frac{1}{m-1} \sum (y_{ij} - \bar{y}_i)^2 \quad (3.9)$$

где m – число повторений i -того опыта, y_{ij} – экспериментальные значения функции отклика, \bar{y}_i – среднее значение функции отклика в i -том опыте.

Затем определяется критерий Кохрена G , который представляет собой отношение максимальной дисперсии воспроизводимости к сумме всех дисперсий, по формуле:

$$G = \frac{D_{y_{\max}}}{\sum D_{y_i}} \quad (3.10)$$

После чего осуществляется проверка однородности дисперсий (погрешности опыта). Для этого рассчитанное по (3.10) значение критерия Кохрена G при принятом уровне значимости α (α – вероятность отвергнуть правильную гипотезу при условии, что она верна, чаще всего $\alpha = 0,05$) сравнивается с табличным значением G_T .

Если $G < G_T$, то гипотеза о равнозначности дисперсий воспроизводимости не отвергается. При этом признается, что все опыты выполнены с равной погрешностью.

Если условие $G < G_T$ не выполняется, то делают вывод, что опыты выполнены не с равной погрешностью, и эксперимент необходимо переделать.

2) С помощью метода наименьших квадратов определяются коэффициенты аппроксимирующего полинома (3.3) по формулам (3.7).

3) Производится проверка гипотезы об адекватности модели изучаемому объекту, исходя из критерия Фишера F , найденного по формуле:

$$F = \frac{D_{y_a}}{D_{y_o}} \quad (3.11)$$

Здесь D_{y_a} – дисперсия адекватности, характеризующая рассеивание экспериментальных значений y относительно аппроксимирующей кривой, определяемая по формуле:

$$D_{y_a} = \frac{1}{N-s} \sum_{i=1}^N (y_{pi} - y_{\varepsilon i})^2 \quad (3.12)$$

где N – число опытов, s – число коэффициентов полинома, y_{pi} и $y_{\varepsilon i}$ – расчетные и экспериментальные значения y в i -том опыте соответственно.

D_{y_o} – дисперсия опыта, характеризующая рассеяние значений выходного параметра y в эксперименте, и рассчитывается по формуле:

$$D_{y_o} = \frac{1}{mN} \sum D_{y_i} \quad (3.13)$$

где D_{y_i} – дисперсия воспроизводимости опытов, найденная по формуле (3.9), mN – общее число опытов.

Для оценки качества регрессионной модели выполняется сравнение рассчитанного по (3.11) значения критерия Фишера F и табличного значения F_T .

Табличное значение критерия Фишера F_T – это максимально возможное значение критерия под влиянием случайных факторов при данных степенях свободы и уровне значимости.

Если $F < F_T$, то гипотеза об адекватности аппроксимирующей зависимости экспериментальным данным не отвергается (математическая модель признается адекватной объекту исследования).

Если неравенство $F < F_T$ не удовлетворяется, то гипотеза об адекватности аппроксимирующей зависимости экспериментальным данным отвергается (математическая модель признается неадекватной объекту исследования). В данном случае следовало бы либо увеличить число опытов эксперимента, либо поменять вид аппроксимирующей зависимости.

4) Проверяется значимость коэффициентов полинома (3.3) по критерию Стьюдента (отдельно по каждому коэффициенту), рассчитанному по формуле:

$$t_{b_i} = \frac{|b_i|}{\sigma_{b_i}} \quad (3.14)$$

где $|b_i|$ – абсолютное значение коэффициента аппроксимирующего полинома; σ_{b_i} – среднее квадратическое отклонение коэффициентов

аппроксимирующего полинома, $\sigma_{b_i} = \sqrt{D_{b_i}}$; D_{b_i} – дисперсия коэффициентов аппроксимирующего полинома b_i , $D_{b_i} = \frac{D_{y_0}}{N}$.

Значение t_{b_i} сравнивается с табличным t_T . Если $t_{b_i} < t_T$, то коэффициент b_i считается не значимым (т.е. можно принять $b_i = 0$) и соответствующее слагаемое исключается из уравнения регрессии. Если $t_{b_i} > t_T$, то коэффициент b_i считается значимым (существенно влияющим на функцию отклика).

После сравнения расчетных значений t_{b_i} и табличным t_T , из рассмотрения исключаются незначимые коэффициенты и осуществляется повторная проверка адекватности модели.

5) Анализ результатов эксперимента завершается интерпретацией модели в терминах объекта исследования. Прежде всего, выясняется, в какой мере каждый из факторов влияет на функцию отклика. Значения линейных коэффициентов служат количественной мерой, оценивающей влияние факторов: чем больше коэффициент b_i , тем сильнее это влияние. Знак коэффициента позволяет судить о характере зависимости функции отклика от соответствующих факторов.

Затем аналогично следует проанализировать эффект парных взаимодействий.

СОДЕРЖАНИЕ ЗАДАНИЯ

Рассмотрим объект, предназначенный для получения гальванического покрытия с минимальными внутренними напряжениями, при которых обеспечивается необходимая прочность покрытия. Внутренние напряжения могут быть измерены по деформации металла. В результате эксперимента необходимо выяснить, каким образом на внутренние напряжения, принятые в качестве функции отклика, влияют различные факторы.

Предварительная оценка показала, что наибольший интерес представляют следующие три фактора: плотность тока X_1 , температура раствора X_2 и концентрация в нем определенного вещества X_3 .

Уровни факторов и интервалы варьирования (табл. 3.2) выбраны на основе априорных сведений об объекте.

Таблица 3.2

Параметр	Фактор		
	Плотность тока X_1 , А/дм ²	Температура X_2 , °С	Концентрация X_3 , кг/м ³
Основной уровень \bar{X}_i	55	45	0,7
Интервал варьирования I_i	25	15	0,3
Верхний уровень $\bar{X}_{i \max}$	80	60	1,0
Нижний уровень $\bar{X}_{i \min}$	30	30	0,4

По данным эксперимента, приведенного в таблице 3.3:

1. Составить план эксперимента, учитывающий линейные эффекты и эффекты парного взаимодействия факторов.

2. Провести оценку равноточности дисперсий опыта по критерию Кохрена, определить коэффициенты аппроксимирующего полинома по методу наименьших квадратов.

3. Проверить адекватность модели исследуемому объекту по критерию Фишера.

4. Проверить значимость коэффициентов аппроксимирующего полинома по критерию Стьюдента, исключить из рассмотрения незначимые коэффициенты. Выполнить проверку адекватности для вновь полученной модели.

5. Дать интерпретацию результатов эксперимента в терминах объекта исследования.

Таблица 3.3

вар	1		2		3		4		5	
Y	Y'	Y''	Y'	Y''	Y'	Y''	Y'	Y''	Y'	Y''
1	3,50	3,70	3,30	3,99	3,20	3,87	2,45	4,15	3,35	4,05
2	-0,42	-0,62	-0,38	-0,59	-0,36	-0,56	-0,41	-0,64	-0,38	-0,61
3	2,80	3,50	2,60	1,69	2,50	1,55	2,73	1,87	2,64	2,55
4	2,45	3,27	2,25	3,04	2,15	2,92	2,41	3,20	2,30	2,87
5	2,30	3,42	2,10	3,19	2,00	3,06	2,25	3,35	2,15	2,56
6	-0,90	-1,22	-0,78	-1,05	-0,72	-1,04	-0,87	-1,20	-0,81	-1,14
7	0,64	0,77	0,56	0,87	0,52	0,83	0,62	0,92	0,58	0,65
8	0,64	0,55	0,56	0,37	0,53	0,32	0,60	0,42	0,59	0,21
вар	6		7		8		9		10	
Y	Y'	Y''	Y'	Y''	Y'	Y''	Y'	Y''	Y'	Y''
1	3,54	3,89	3,25	3,90	3,50	4,22	3,50	3,99	3,35	3,85
2	-0,43	-0,63	-0,37	-0,58	-0,44	-0,61	-0,36	-0,57	-0,41	-0,64
3	2,85	1,98	2,53	1,65	2,75	1,89	2,65	1,80	2,90	2,52
4	2,51	3,33	2,18	3,29	2,48	3,36	2,20	2,28	2,45	3,27
5	2,35	2,88	2,05	3,12	2,25	3,30	2,15	3,00	2,40	3,47
6	-0,91	-0,55	-0,75	-1,08	-0,96	-1,28	-0,72	-1,05	-0,87	-1,20
7	0,65	0,96	0,64	0,85	0,62	0,94	0,57	0,87	0,65	0,96
8	0,67	0,46	0,54	0,35	0,68	0,46	0,52	0,33	0,60	0,42
вар	11		12		13		14		15	
Y	Y'	Y''	Y'	Y''	Y'	Y''	Y'	Y''	Y'	Y''
1	3,20	3,87	3,45	4,15	3,35	4,05	3,54	4,29	3,25	3,90
2	-0,39	-0,60	-0,43	-0,63	-0,37	-0,58	-0,41	-0,64	-0,38	-0,61
3	2,55	2,65	2,73	1,87	2,64	1,75	2,90	2,02	2,50	1,55
4	2,30	3,15	2,57	3,37	2,15	2,92	2,45	3,27	2,25	3,04
5	2,10	3,07	2,30	3,42	2,10	3,19	2,40	3,47	2,00	3,06
6	-0,98	-1,05	-0,95	-1,28	-0,72	-1,04	-0,87	-1,20	-0,81	-1,14
7	0,52	0,83	0,62	0,92	0,68	0,98	0,65	0,96	0,64	0,85
8	0,58	0,59	0,67	0,16	0,60	0,42	0,60	0,42	0,59	0,37
вар	16		17		18		19		20	
Y	Y'	Y''	Y'	Y''	Y'	Y''	Y'	Y''	Y'	Y''
1	3,45	4,15	3,20	3,87	3,33	4,00	3,45	4,15	3,60	4,32
2	-0,43	-0,64	-0,38	-0,59	-0,36	-0,56	-0,43	-0,64	-0,41	-0,62
3	2,80	1,92	2,50	1,55	2,60	1,69	2,80	1,92	2,85	1,98
4	2,37	3,21	2,30	3,09	2,15	2,92	2,57	3,37	2,40	3,11
5	2,35	2,48	2,05	3,12	2,15	3,25	2,25	3,35	2,45	2,40
6	-0,87	-1,20	-0,81	-1,14	-0,75	-1,08	-0,91	-1,25	-0,90	-1,22
7	0,68	0,98	0,52	0,83	0,58	0,90	0,62	0,92	0,68	0,98
8	0,64	0,44	0,56	0,37	0,53	0,32	0,70	0,98	0,62	0,42

ЛАБОРАТОРНАЯ РАБОТА № 4. РАЗРАБОТКА МОДЕЛИ ВТОРОГО ПОРЯДКА

Цель работы – изучение приемов планирования эксперимента и обработки экспериментальных данных; нахождение параметров линейного регрессионного уравнения первой степени с учетом эффекта парного взаимодействия по результатам эксперимента.

ОСНОВЫ ТЕОРИИ И РАСЧЕТНЫЕ ФОРМУЛЫ

В области факторного пространства, в которой линейная модель становится неадекватной, доминирующим является взаимодействие факторов. Для адекватного описания этой области факторного пространства необходимо использовать нелинейные полиномы, поскольку функция отклика в этой области часто имеет экстремумы. Для описания этой области необходимо составить ортогональный центрально-композиционный план (ОЦКП), для чего обычно применяют полиномы 2-го порядка. Это связано с тем, что поверхности 2-го порядка легко поддаются систематизации и определению экстремума.

Полный факторный эксперимент типа 2^2 при планировании первого порядка представляют собой 4 опыта (точки 1-4), поставленные в вершинах квадрата (рис. 4.1).

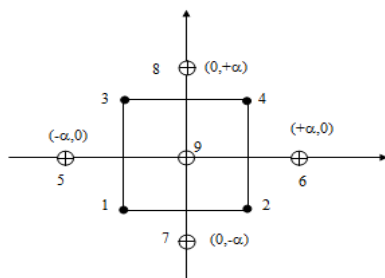


Рис. 4.1. Построение композиционного плана для двухфакторного эксперимента

В данном случае план ПФЭ типа 2^2 принимают за ядро или центр плана второго порядка и добавляют к нему симметрично расположенные точки факторного пространства (точки 5-8), называемые

мые звездными (с координатами $x_i = \pm\alpha; x_j = 0$), а также точку (точка 9) начала координат ($x_1 = x_2 = 0$) (табл. 4.1).

Таблица 4.1

	u	Факторы			
		x_0	x_1	x_2	$x_3 = x_1 \cdot x_2$
ПФЭ 2 ² Ядро плана	1	+1	-1	-1	+1
	2	+1	+1	-1	-1
	3	+1	-1	+1	-1
	4	+1	+1	+1	+1
Звездные точки	5	+1	$-\alpha$	0	0
	6	+1	$+\alpha$	0	0
	7	+1	0	$-\alpha$	0
	8	+1	0	$+\alpha$	0
Нулевая точка	9	+1	0	0	0

Выбор звездных точек следует проводить таким образом, чтобы выдерживались свойства симметрии (3.4) и ортогональности (3.6). В данном случае условия нормировки могут не выполняться.

Для выполнения условия симметричности относительно центра эксперимента (3.4) столбцы матрицы планирования, соответствующие квадратам факторов, необходимо подвергнуть простейшему преобразованию:

$$x_{iu}^* = x_{iu}^2 - d \quad (4.1)$$

где x_i^* – преобразованный фактор; d – постоянная величина, зависящая от числа факторов k и подлежащая определению из условий (3.4) и (3.6).

Тогда условие (3.4) с учетом (4.1) примет вид:

$$\sum_{u=1}^N x_{iu}^* = \sum_{u=1}^N (x_{iu}^2 - d) = 0 \quad (4.2)$$

А условие (3.6) – вид:

$$\sum_{u=1}^N x_{iu}^* \cdot x_{ju}^* = \sum_{u=1}^N (x_{iu}^2 - d) \cdot (x_{ju}^2 - d) = 0 \quad (4.3)$$

Таким образом, используя условия (4.2) и (4.3) можно найти неизвестные d и α .

Перепишем выражения (4.2) и (4.3) в виде:

$$\begin{cases} \sum_{u=1}^N x_{iu}^2 - Nd = 0 \\ \sum_{u=1}^N x_{iu}^2 \cdot x_{ju}^2 + Nd^2 - d \sum_{i=1}^N x_{iu}^2 - d \sum_{j=1}^N x_{ju}^2 = 0 \end{cases} \quad (4.4)$$

Эта система решается для конкретного числа опытов и числа факторов. В нашем случае для двух факторов и девяти опытов.

$$\begin{cases} \sum_{u=1}^N x_{iu}^2 = \sum_{u=1}^4 x_{iu}^2 + \sum_{u=5}^8 x_{iu}^2 + \sum_{u=9} x_{iu}^2 = 4 + 2\alpha^2 + 0 \\ \sum_{u=1}^N x_{iu}^2 \cdot x_{ju}^2 = 4 \end{cases}$$

Тогда, подставляя полученные выражения в (4.4), получим:

$$\begin{cases} 4 + 2\alpha^2 - 9d = 0 \\ 4 + 9d^2 - 2d(4 + 2\alpha^2) = 0 \end{cases} \quad (4.5)$$

Решение системы (4.5) относительно d и α дает следующие результаты: $d = \frac{2}{3}$; $\alpha = 1$. С учетом найденных коэффициентов матрица плана ОЦКП примет следующий вид (табл. 4.2):

Таблица 4.2

	u	Факторы						Отклик
		x_0	x_1	x_2	x_3	x_4	x_5	Y
ПФЭ 2 ² Ядро плана	1	+1	-1	-1	+1	1/3	1/3	Y_1
	2	+1	+1	-1	-1	1/3	1/3	Y_2
	3	+1	-1	+1	-1	1/3	1/3	Y_3
	4	+1	+1	+1	+1	1/3	1/3	Y_4
Звездные точки	5	+1	-1	0	0	1/3	-2/3	Y_5
	6	+1	+1	0	0	1/3	-2/3	Y_6
	7	+1	0	-1	0	-2/3	1/3	Y_7
	8	+1	0	+1	0	-2/3	1/3	Y_8
Нулевая точка	9	+1	0	0	0	-2/3	-2/3	Y_9

где $x_3 = x_1 \cdot x_2$, $x_4 = x_1^2 - \frac{2}{3}$, $x_5 = x_2^2 - \frac{2}{3}$.

В силу ортогональности матрицы планирования все коэффициенты регрессии определяются независимо друг от друга по формуле:

$$b_i = \frac{\sum_{u=1}^N Y_u x_{iu}}{\sum_{u=1}^N x_{iu}^2} \quad (4.6)$$

где j – порядковый номер столбца в матрице планирования.

Уравнение регрессии для полинома второго порядка при $k = 2$ имеет вид:

$$Y = b_0 x_0 + b_1 x_1 + b_2 x_2 + b_{12} x_1 x_2 + b_{11} x_1^2 + b_{22} x_2^2. \quad (4.7)$$

Заканчивается расчет проверкой адекватности найденного полинома и интерпретацией модели в терминах объекта исследования.

СОДЕРЖАНИЕ ЗАДАНИЯ

Определить математическую модель объекта исследования, применяя методы планирования эксперимента. Для этого:

1. Составить простейший план типа ПФЭ 2^2 и по данным эксперимента, приведенным в таблице 4.3 (опыты (I-IV)), найти линейную модель объекта с учетом эффекта парного взаимодействия. Найти расчетные значения функции отклика.

2. Проверить адекватность полученной зависимости по критерию Фишера. Сделать вывод о возможности использования F -критерия в условиях данного эксперимента.

3. Проверить адекватность полученной зависимости, поставив дополнительный опыт в нулевой точке ($x_1 = x_2 = 0$) (опыт IX). Для этого найти расчетное значение функции отклика в нулевой точке ($x_1 = x_2 = 0$). Найти относительную ошибку в нулевой точке по формуле:

$$\Delta = \frac{|Y_{\text{э}} - Y_{\text{р}}|}{Y_{\text{э}}} \cdot 100\%$$

где $Y_{\text{э}}$ и $Y_{\text{р}}$ – экспериментальное и расчетное значение функции отклика соответственно.

В случае, если ошибка аппроксимации превышает 10 % от максимального значения функции отклика, составить ортогональный центрально-композиционный план (ОЦКП) (опыты I-IX).

4. Определить нелинейную модель объекта в виде полной квадратики, оценив адекватность полученной зависимости величиной относительной ошибки.

Таблица 4.3

вар	1	2	3	4	5	6	7	8	9	10
I	4,5	6,5	5,9	6,2	5,6	6,7	6,0	7,5	4,5	5,1
II	2,4	3,4	2,8	3,2	2,7	3,8	3,5	4,4	3,4	2,6
III	3,3	4,3	3,7	4,1	3,9	4,7	4,8	5,3	3,7	4,1
IV	6,8	7,8	6,8	7,2	6,5	7,9	7,1	8,0	6,2	6,5
V	4,2	5,2	4,8	5,1	4,7	5,3	6,0	6,2	5,1	5,2
VI	4,6	5,6	6,1	4,8	3,7	4,7	5,4	6,1	5,3	3,9
VII	0,7	1,7	1,4	1,8	2,3	2,3	1,8	2,7	2,0	1,6
VIII	2,3	3,3	2,7	2,6	3,4	3,1	3,1	4,0	3,1	2,9
IX	1,5	2,1	2,3	1,8	2,2	2,2	2,3	3,0	3,0	1,9
вар	11	12	13	14	15	16	17	18	19	20
I	6,5	5,5	6,5	6,0	6,7	7,5	6,7	5,6	6,2	5,9
II	2,8	2,8	7,3	3,8	2,7	4,4	3,5	3,8	2,7	2,6
III	3,6	4,3	4,4	3,9	4,1	4,8	5,3	4,8	4,7	4,1
IV	7,2	6,8	4,9	7,2	7,2	7,9	7,1	8,0	7,1	6,5
V	4,7	5,2	5,1	4,6	4,8	4,7	5,3	6,0	6,2	5,3
VI	4,7	6,1	6,5	6,1	5,6	4,8	5,3	5,4	5,4	5,4
VII	2,8	0,9	4,7	1,2	1,7	2,9	1,8	1,3	1,8	2,2
VIII	4,2	2,6	2,1	2,3	3,3	3,5	2,8	2,6	3,4	3,1
IX	2,8	2,3	3,2	1,6	2,2	2,1	1,9	2,3	2,8	2,8
вар	21	22	23	24	25	26	27	28	29	30
I	6,6	5,2	5,5	6,3	6,5	7,6	6,4	6,4	6,0	5,7
II	2,6	2,7	3,5	3,5	3,2	4,1	3,4	3,8	2,8	2,7
III	3,4	4,3	4,3	3,7	4,3	4,5	5,1	5,1	4,2	3,6
IV	6,9	7,1	7,4	6,9	6,5	7,6	7,1	8,3	6,8	6,2
V	4,1	4,8	4,9	4,5	4,9	4,7	5,6	6,1	6,2	5,3
VI	4,9	5,8	5,1	6,1	5,4	5,0	5,5	5,2	5,6	5,6
VII	2,9	0,6	2,2	1,9	2,8	3,6	1,8	2,6	2,6	2,2
VIII	4,0	2,5	4,3	2,3	2,9	3,3	2,7	4,3	3,6	3,3
IX	2,7	1,7	2,6	1,9	2,8	2,0	2,0	2,9	3,6	3,3

ЛАБОРАТОРНАЯ РАБОТА № 5. ПОСТРОЕНИЕ ГИСТОГРАММ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Цель работы – освоение практических приемов работы с числовыми характеристиками случайных величин, построение распределений случайных величин.

ОСНОВЫ ТЕОРИИ И РАСЧЕТНЫЕ ФОРМУЛЫ

Законом распределения случайной величины X называется любое правило, позволяющее находить вероятность P всевозможных событий, связанных с этой случайной величиной. Например, вероятность того, что случайная величина X не примет какое-то значение или попадет в определенный интервал. В последнем случае обычно пользуются так называемой плотностью распределения непрерывной случайной величины $f(x)$. При этом вероятность попадания этой величины в интервал $a < x < b$ будет

$$P\{a < x < b\} = \int_a^b f(x)dx \quad (5.1)$$

а интеграл от плотности распределения в пределах $(-\infty; +\infty)$ равен единице.

Предположим, на основе эксперимента имеем n наблюдений случайной величины X , которые разделим на k интервалов Δx_i . Подсчитаем количество значений этой величины m_i , приходящихся на каждый интервал. Тогда частота появления X в интервале Δx_i равна

$$p_i^* = \frac{m_i^*}{n} \quad (5.2)$$

Составим статистический ряд (табл. 5.1), в котором приведены интервалы в порядке их расположения по оси X , а также соответствующие им частоты.

Число интервалов, в которые, следует группировать статистический материал, не должно быть слишком большим (тогда ряд

распределения становится невыразительным, и частоты в нем обнаруживают незакономерные колебания); с другой стороны, оно не должно быть слишком малым (при малом числе разрядов свойства распределения описываются статистическим рядом слишком грубо). Практика показывает, что в большинстве случаев рационально выбрать порядка 10-20 интервалов.

Таблица 5.1

Статистический ряд

Δx_i	$[x_0, x_1]$	$[x_1, x_2]$...	$[x_i, x_{i+1}]$...	$[x_{k-1}, x_k]$
p_i^*	p_1^*	p_2^*	...	p_{i+1}^*	...	p_k^*

Статистический ряд часто оформляется графически в виде *гистограмм*. Для построения гистограммы по оси абсцисс откладывают интервалы Δx_i и на каждом из них, как на основании, строят прямоугольник, площадь которого равна p_i^* (рис. 5.1). Таким образом, из условия построения следует, что площадь гистограммы равна единице.

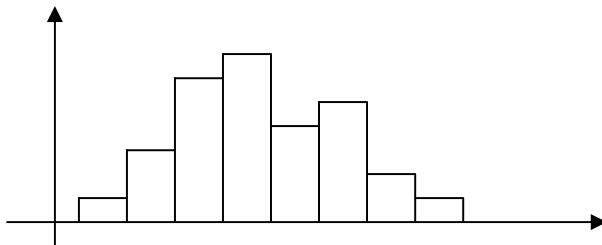


Рис. 5.1. Гистограмма распределения случайной величины

Имея гистограмму распределения, можно построить статистическую функцию распределения.

С помощью гистограммы распределения можно подобрать наиболее вероятную теоретическую кривую плотности распределения случайной величины X . Эта задача называется выравниванием статистических распределений. Наилучшее приближение статистической зависимости к имеющемуся статистическому материалу выбирается, исходя из соображений, связанных с физикой решаемой задачи или с внешним видом гистограммы.

Пользуясь данными статистического ряда, можно приближенно построить и статистическую функцию распределения $F^*(x)$ величины X . Для практики обычно достаточно построить статистическую функцию распределения по нескольким точкам. В качестве этих точек удобно взять границы интервалов x_0, x_1, \dots, x_k , которые фигурируют в статистическом ряде. Тогда,

$$\begin{aligned}
 F^*(x_0) &= 0; \\
 F^*(x_1) &= p_1^*; \\
 F^*(x_2) &= p_1^* + p_2^*; \\
 &\dots \\
 F^*(x_i) &= \sum_{i=0}^i p_i^*; \\
 &\dots \\
 F^*(x_k) &= \sum_{i=0}^k p_i^* = 1.
 \end{aligned}
 \tag{5.3}$$

Соединяя полученные точки ломаной линией или плавной кривой, получим приближенный график статистической функции распределения $F^*(x)$. Пример гистограммы распределения и статистической функции распределения приведены на рис. 5.2 и 5.3 соответственно.

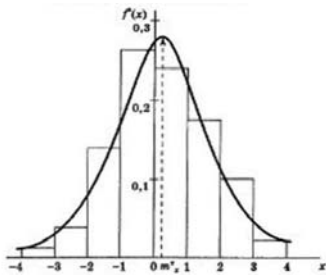


Рис. 5.2. Гистограмма распределения величины X

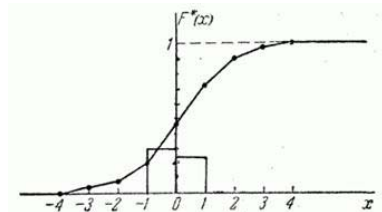


Рис. 5.3. Статистическая функция распределения

Во многих случаях при обработке экспериментальных данных предполагается, что распределение случайных величин подчиняется *нормальному закону*, или закону ошибок, введенному Гауссом. Следует отметить, что в большинстве случаев это предположение оправдано. Однако если априори закон распределения или его параметры неизвестны, то следует проверить правильность соответствующих гипотез.

Непрерывная случайная величина X называется распределенной по *нормальному закону*, если ее плотность распределения имеет вид:

$$f(x) = \frac{1}{\sigma_x \cdot \sqrt{2\pi}} \cdot \exp\left(-\frac{(x - m_x)^2}{2\sigma_x^2}\right) \quad (5.4)$$

где m_x и σ_x – математическое ожидание и среднее квадратическое отклонение случайной величины X соответственно.

Нормализованным нормальным распределением называется такое нормальное распределение, у которого $m_x = 0$ и $\sigma_x = 1$. Из нормализованного распределения можно получить любое другое нормальное распределение с заданными m_x и σ_x по формуле:
 $z = m_x + x \cdot \sigma_x$.

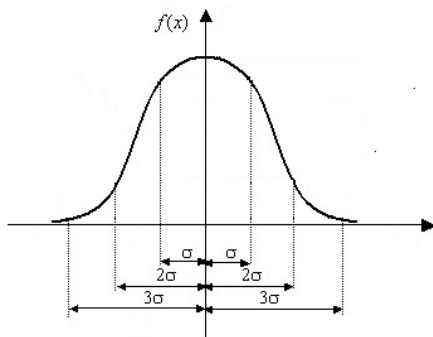


Рис. 5.4. Графический вид нормального закона распределения случайной величины X с параметрами $m_x = 0$ и $\sigma_x = 1$ (распределение нормализовано)

График плотности нормального распределения имеет вид колокола. На рис. 5.4 показано нормализованное нормальное распределение.

График (рис. 5.4) показывает, что в области $-\sigma < x < \sigma$ на графике сосредоточено 68 % площади распределения, в области $-2\sigma < x < 2\sigma$ – сосредоточено 95,4 % площади распределения, в области $-3\sigma < x < 3\sigma$ – сосредоточено 99,7 % площади распределения. Таким образом, практически все значения нормально распределенной случайной величины лежат в интервале $[m_x - 3\sigma; m_x + 3\sigma]$. Это правило называется «правилом трех сигм». Другими словами, если отклонение случайной величины X превосходит 3σ , то это свидетельствует либо о грубой ошибке в опыте (так называемый «промах»), либо о том, что случайная величина распределена не в соответствии с нормальным законом распределения. И в том, и в другом случаях объем эксперимента следует увеличить.

Свойства нормального распределения

Как видно из (5.4), нормальное распределение имеет два параметра: математическое ожидание m_x и среднеквадратичное отклонение σ_x величины X от этого математического ожидания m_x .

Изменение параметра нормального распределения m_x приводит к сдвигу кривой по оси x (рис. 5.5).

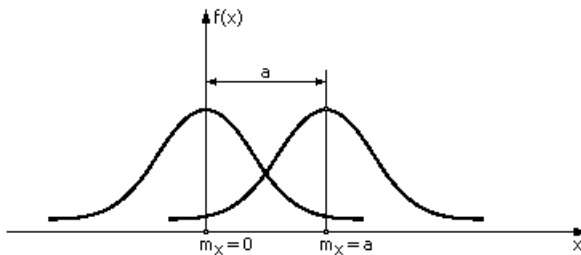


Рис. 5.5. Влияние параметра «математическое ожидание» на вид закона нормального распределения случайной величины X

Изменение параметра нормального распределения σ_x приводит к масштабированию формы по оси x (рис. 5.6). При этом в

Если задается интервал интегрирования функции Лапласа $[a, b]$, то:

$$\begin{aligned}
 P(a < x < b) &= \int_a^b \frac{1}{\sigma_x \cdot \sqrt{2\pi}} \cdot \exp\left(-\frac{(x - m_x)^2}{2\sigma_x^2}\right) dx = \\
 &= \int_{\frac{a - m_x}{\sigma_x}}^{\frac{b - m_x}{\sigma_x}} \frac{1}{\sigma_x \cdot \sqrt{2\pi}} \cdot \exp\left(-\frac{t^2}{2}\right) dt = \Phi\left(\frac{b - m_x}{\sigma_x}\right) - \Phi\left(\frac{a - m_x}{\sigma_x}\right)
 \end{aligned} \tag{5.6}$$

Выравнивание статистических рядов

Во всяком статистическом распределении неизбежно присутствуют элементы случайности, связанные с тем, что число наблюдений ограничено, что произведены именно те, а не другие опыты, давшие именно те, а не другие результаты. Только при очень большом числе наблюдений эти элементы случайности сглаживаются, и случайное явление обнаруживает в полной мере присущую ему закономерность. На практике мы почти никогда не имеем дела с таким большим числом наблюдений и вынуждены считаться с тем, что любому статистическому распределению свойственны в большей или меньшей мере черты случайности. Поэтому при обработке статистического материала часто приходится решать вопрос о том, как подобрать для данного статистического ряда теоретическую кривую распределения, выражающую лишь существенные черты статистического материала, но не случайности, связанные с недостаточным объемом экспериментальных данных. Такая задача называется задачей выравнивания (сглаживания) статистических рядов.

Задача выравнивания заключается в том, чтобы подобрать теоретическую плавную кривую распределения, с той или иной точки зрения наилучшим образом описывающую данное статистическое распределение (рис. 5.8).

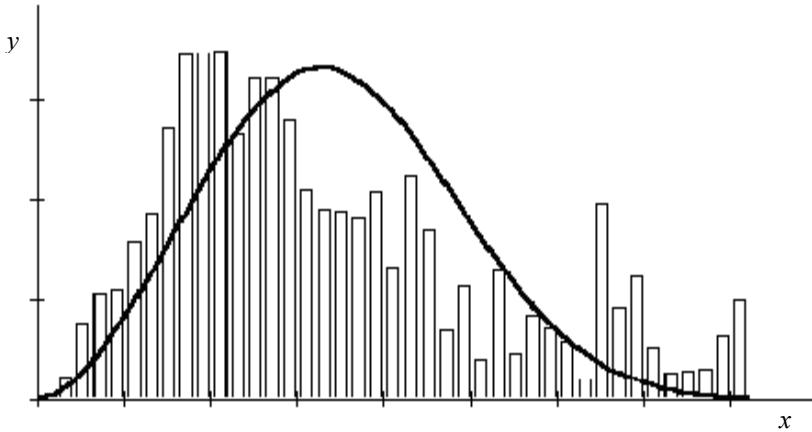


Рис. 5.8 Гистограмма и теоретическая кривая

Предположим, что исследуемая величина X есть ошибка измерения, возникающая в результате суммирования воздействий множества независимых элементарных ошибок, тогда из теоретических соображений можно считать, что величина X подчиняется нормальному закону 5.4:

$$f(x) = \frac{1}{\sigma_x \cdot \sqrt{2\pi}} \cdot \exp\left(-\frac{(x - m_x)^2}{2\sigma_x^2}\right)$$

и задача выравнивания переходит в задачу о рациональном выборе параметров m_x и σ_x .

Таким образом, требуется подобрать параметры m_x и σ_x так, чтобы функция $f(x)$ наилучшим образом описывала данный статистический материал.

Один из методов, применяемых для решения этой задачи, является метод моментов. Согласно методу моментов, параметры m_x и σ_x выбираются с таким расчетом, чтобы несколько важнейших числовых характеристик (моментов) теоретического распределения были равны соответствующим статистическим характеристикам. Таким образом, параметры m_x и σ_x выбираются так, чтобы

математическое ожидание m_x и среднее квадратическое отклонение σ_x теоретического распределения совпадали с соответствующими статистическими характеристиками m_x^* и σ_x^* . То есть, должны выполняться условия $m_x = m_x^*$ и $\sigma_x = \sigma_x^*$.

Статистическое математическое ожидание находится по формуле:

$$m_x^* = \sum_{i=1}^k m_{x_i} \cdot p_i^* \quad (5.7)$$

где m_{x_i} – среднее значение интервала, p_i^* – частота появления величины появления X в i -том интервале.

Статистическое среднее квадратическое отклонение находится в соответствии с выражением:

$$\sigma_x^* = \sqrt{D_x^*} \quad (5.8)$$

где D_x^* – дисперсия величины X , определяемая по формуле:

$$D_x^* = \alpha^* - (m_x^*)^2 \quad (5.9)$$

где α – второй начальный момент, определяемый выражением:

$$\alpha^* = \sum_{i=1}^k m_{x_i}^2 \cdot p_i^* \quad (5.10)$$

Критерий согласия Пирсона

Предположим, что гистограмма может быть аппроксимирована аналитической функцией $f(x)$. Как бы хорошо ни была подобрана теоретическая кривая, между нею и статистическим распределением неизбежны некоторые расхождения.

Естественно возникает вопрос: объясняются ли эти расхождения только случайными обстоятельствами, связанными с ограни-

ченный числом наблюдений, или они являются существенными и связаны с тем, что подобранная нами кривая плохо выравнивает данное статистическое распределение. Для ответа на этот вопрос может быть использован критерий согласия χ^2 (читается «хи-квадрат»), предложенный Пирсоном.

Предположим, что произведено n независимых опытов, в каждом из которых случайная величина X приняла определенное значение. Результаты опытов сведены в статистический ряд (табл. 5.1).

Требуется проверить, согласуются ли экспериментальные данные с гипотезой о том, что случайная величина X имеет данный закон распределения (заданный функцией распределения $F(x)$ или плотностью распределения $f(x)$). Назовем этот закон распределения «теоретическим».

Зная закон распределения, можно найти теоретические вероятности попадания случайной величины в каждый из интервалов: p_1, p_2, \dots, p_k .

Проверяя согласованность теоретического и статистического распределений, будем исходить из расхождений между теоретическими вероятностями p_i и наблюдаемыми частотами p_i^* .

В качестве меры расхождения между теоретическим и статистическим распределениями выбрана сумма квадратов отклонений между вероятностями, взятых с некоторыми «весами» c_i :

$$\chi^2 = \sum_{i=1}^k c_i (p_i^* - p_i)^2 \quad (5.11)$$

Коэффициенты c_i («веса» интервалов) вводятся потому, что в общем случае отклонения, относящиеся к различным разрядам, нельзя считать равноправными по значимости. Действительно, одно и то же по абсолютной величине отклонение $(p_i^* - p_i)$, может быть мало значимым, если сама вероятность p_i мала. Поэтому «веса» c_i

необходимо взять обратно пропорциональными вероятностям p_i попадания случайной величины в каждый из интервалов.

К. Пирсон показал, что если положить

$$c_i = \frac{n}{p_i} \tag{5.12}$$

то при большом числе опытов n закон распределения величины χ^2 обладает простыми свойствами: он практически не зависит от функции распределения $F(x)$ и от числа опытов n , а именно, этот закон при увеличении n приближается к так называемому «распределению χ^2 ».

Таким образом, с учетом (5.12) выражение (5.11) примет вид:

$$\chi^2 = \sum_{i=1}^k \frac{n}{p_i} (p_i^* - p_i)^2 \tag{5.13}$$

Учитывая, что $p_i^* = \frac{m_i}{n}$, где m_i – число значений в i -том интервале, получим:

$$\chi^2 = \sum_{i=1}^k \frac{(m_i - np_i)^2}{np_i} \tag{5.14}$$

Распределение χ^2 зависит от параметра r , называемого числом «степеней свободы» распределения. Число «степеней свободы» r равно числу интервалов k за вычетом числа независимых условий («связей»), наложенных на частоты p_i^* . Обычно на p_i^* накладывают следующие условия S :

$$\sum_{i=1}^k p_i^* = 1$$

Сумма частот должна равняться единице.

$$\sum_{i=1}^k x_i \cdot p_i^* = m_{x_i}$$

Средние значения теоретического и статистического распределений должны совпадать.

$$\sum_{i=1}^k (x_i - m_{x_i})^2 \cdot p_i^* = D_x$$

Должны совпадать теоретическая и статистическая дисперсии.

Для распределения χ^2 составлены таблицы (прил. 5). Пользуясь этими таблицами, можно для каждого значения χ^2 и числа степеней свободы r найти вероятность P того, что величина, распределенная по закону χ^2 , превзойдет это значение.

Если найденное значение P велико, то гипотезу о том, что функция $f(x)$ пригодна для аппроксимации полученной гистограммы, можно признать не противоречащей опытными данным. Если же вероятность P мала (менее 0,1), то эта гипотеза отбрасывается как неправдоподобная и следует либо выбрать другое распределение, либо провести дополнительные эксперименты.

Таким образом, схема применения критерия χ^2 к оценке согласованности теоретического и статистического распределений сводится к следующему:

- 1) Определяется величина χ^2 по формуле (5.14).
- 2) Определяется число степеней свободы $r = k - s$.
- 3) По r и χ^2 с помощью таблиц (прил. 5) определяется вероятность того, что величина, имеющая распределение χ^2 с r степенями свободы, превзойдет данное значение χ^2 . Если эта вероятность весьма мала (менее 0,1), то гипотеза отбрасывается как не-

правдоподобная. Если эта вероятность относительно велика, гипотезу можно признать не противоречащей опытными данным.

СПОСОБЫ ЗАДАНИЯ ГИСТОГРАММ В MATHCAD

В Mathcad имеется ряд функций для расчетов числовых характеристик случайных данных:

- $\text{mean}(x)$ – среднее значение, оценка математического ожидания выборки,
- $\text{var}(x)$ – дисперсия выборки,
- $\text{stdev}(x)$ – среднеквадратическое отклонение выборки,
- $\text{max}(x)$, $\text{min}(x)$ – максимальное и минимальное значение выборки соответственно.

Гистограмма в Mathcad может быть задана либо данными измерения, либо статистическим рядом распределения.

Рассмотрим оба варианта построения гистограмм на примере решения практической задачи.

А. Построение гистограммы распределения, заданной последовательностью измерений

Задача А. В результате титриметрического определения содержания меди в образце йодометрическим методом было отобрано 100 проб. В табл. 5.1 представлены результаты всех определений в порядке увеличения найденного содержания меди.

Таблица 5.1

Результаты определения количества меди

600,0	605,0	607,0	608,0	609,5	610,5	611,0	612,5	613,5	615,5
601,5	605,0	607,0	608,0	609,5	610,5	611,5	612,5	614,0	616,0
602,5	605,5	607,0	608,5	609,5	610,5	611,5	612,5	614,0	616,0
603,0	605,5	607,5	608,5	609,5	610,5	611,5	612,5	614,0	616,5
604,0	605,5	607,5	608,5	609,5	610,5	611,5	613,0	614,5	617,0
604,0	606,0	607,5	608,5	610,0	611,0	612,0	613,0	614,5	617,0
604,5	606,5	607,5	609,0	610,0	611,0	612,0	613,0	614,5	617,5
604,5	606,5	608,0	609,0	610,0	611,0	612,0	613,0	615,0	618,0
605,0	606,5	608,0	609,0	610,0	611,0	612,0	613,5	615,0	618,5
605,0	607,0	608,0	609,5	610,0	611,0	612,5	613,5	615,5	621,0

Запишем таблицу в виде текстового файла с именем «tit.txt». С помощью функции READPRN считаем данные из текстового файла в файл расчетов.

```
read := READPRN("tit.txt")  
data := stack(read(0), read(1), read(2), read(3), read(4))
```

Общее число определений меди

```
N := length(data)
```

```
N = 100
```

Поле рассеяния случайной величины

```
ω := max(data) – min(data)
```

```
ω = 21
```

```
max(data) = 621
```

```
min(data) = 600
```

Среднее значение

```
mx := mean(data)
```

```
mx=610.235
```

Дисперсия и среднеквадратическое отклонение

```
D:=var(data)
```

```
D=16.098
```

```
σ:=stdev(data)
```

```
σ=4.012
```

Задание интервала гистограммы и запись функции гистограммы

$$h := 2$$

$$\frac{\max(\text{data}) - \min(\text{data})}{h} = 10.5$$

$$n := 11$$

$$i := 0..n$$

$$x_i := 600 + i \cdot h$$

$$V := \text{hist}(x, \text{data})$$

Вероятность попадания в интервал

$$P := \frac{V}{N}$$

Построение гистограммы

	0
0	600
1	602
2	604
3	606
4	608
5	610
6	612
7	614
8	616
9	618
10	620
11	622

 $x =$

	0
0	2
1	2
2	11
3	12
4	18
5	20
6	16
7	10
8	6
9	2
10	1

 $V =$

	0
0	0.02
1	0.02
2	0.11
3	0.12
4	0.18
5	0.2
6	0.16
7	0.1
8	0.06
9	0.02
10	0.01

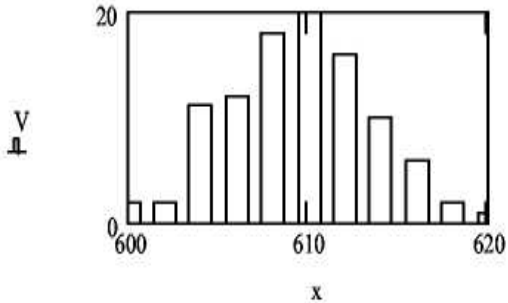
 $P =$


Рис. 5.9. Пример построения гистограммы в Mathcad

Построение графика нормальной плотности вероятности

$$G := \text{dnorm}(x, mx, \sigma)$$

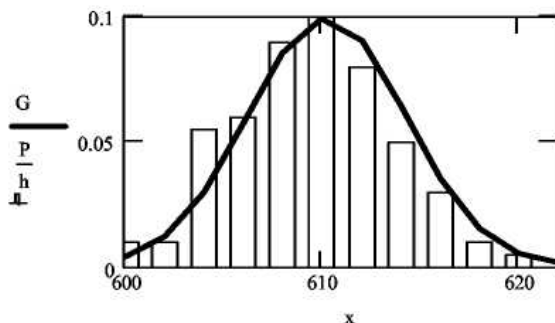


Рис. 5.10. Пример построения теоретического и статистического распределений в Mathcad

Б. Построение гистограммы распределения, заданной статистическим рядом распределения.

Задача Б. По условиям задачи А результаты измерений в порядке увеличения найденного содержания меди сведены в табл. 5.2, в которой q_i – количество меди, мг; m_i – число совпадающих результатов.

Таблица 5.2

Данные о титриметрическом определении меди

q_i	m_i	q_i	m_i	q_i	m_i	q_i	m_i	q_i	m_i
600,0	1	605,5	3	609,0	3	612,5	5	616,0	2
601,5	1	606,0	1	609,5	6	613,0	4	616,5	1
602,5	1	606,5	3	610,0	5	613,5	3	617,0	2
603,0	1	607,0	4	610,5	5	614,0	3	617,5	1
604,0	2	607,5	4	611,0	6	614,5	3	618,0	1
604,5	2	608,0	5	611,5	4	615,0	2	618,5	1
605,0	4	608,5	4	612,0	4	615,5	2	621,0	1

Требуется оценить характер распределения случайной величины.

Решение.

Вводим последовательно исходные данные

$i := 0..50$

data1 $i :=$

600
600.5
601
601.5
602
602.5
603
603.5
604
604.5
605
605.5
606
606.5
607
607.5
608
608.5
609
609.5
610
610.5
611
611.5

data2 $i :=$

1
0
0
1
0
1
1
0
2
2
4
3
1
3
4
4
5
4
3
6
5
5
6
4

612
612.5
613
613.5
614
614.5
615
615.5
616
616.5
617
617.5
618
618.5
619
619.5
620
620.5
621

4
5
4
3
3
3
2
2
2
1
2
1
1
1
0
0
0
0
0
1

data $\langle 0 \rangle :=$ data1

data $\langle 1 \rangle :=$ data2

Рис. 5.11. Ввод исходных данных

Находим общее число определений меди и общее число опытов

$$n := \text{length}(\text{data}^{(0)})$$

$$n = 43$$

$$N := \sum_{i=0}^{n-1} (\text{data}^{(1)})$$

$$N = 100$$

Поле рассеяния случайной величины

$$\omega := \max(\text{data}^{(0)}) - \min(\text{data}^{(0)})$$

$$\omega = 21$$

Среднее значение

$$mx := \frac{\sum_{i=0}^{n-1} (\text{data}^{(0)})_i \cdot (\text{data}^{(1)})_i}{N}$$

$$mx = 610.235$$

Дисперсия и среднеквадратическое отклонение

$$D := \frac{\sum_{i=0}^{n-1} \left[\left[(\text{data}^{(0)})_i - mx \right]^2 \right] \cdot (\text{data}^{(1)})_i}{N - 1}$$

$$D = 16.098$$

$$\sigma := D^{0.5}$$

$$\sigma = 4.012$$

Задание интервала гистограммы

$$h := 0.5$$

$$i := 0.. \frac{\max(\text{data}^{(0)}) - \min(\text{data}^{(0)})}{h}$$

$$x_i := \min(\text{data}^{(0)}) + i \cdot h$$

Определение вероятности попадания в интервал

$$P := \frac{\text{data}^{(1)}}{N}$$

Построение гистограммы

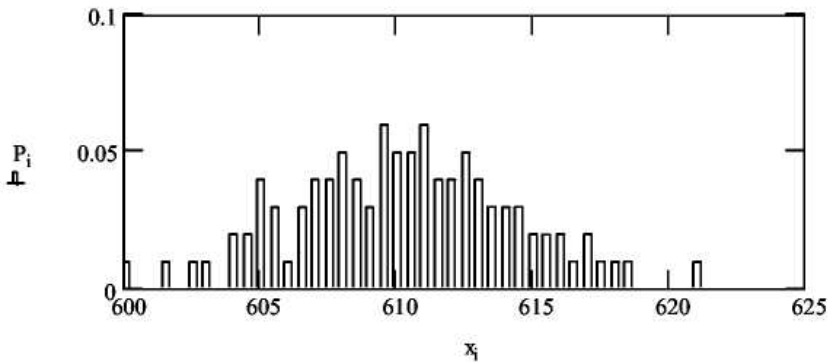


Рис. 5.12. Пример построения гистограммы в Mathcad

Построение графика нормальной плотности вероятности

$$X := \min(x), \min(x) + h.. \max(x)$$

$$G(X) := \frac{\exp\left[-\frac{(X - mx)^2}{2 \cdot \sigma^2}\right]}{\sqrt{2 \cdot \pi \cdot \sigma}}$$

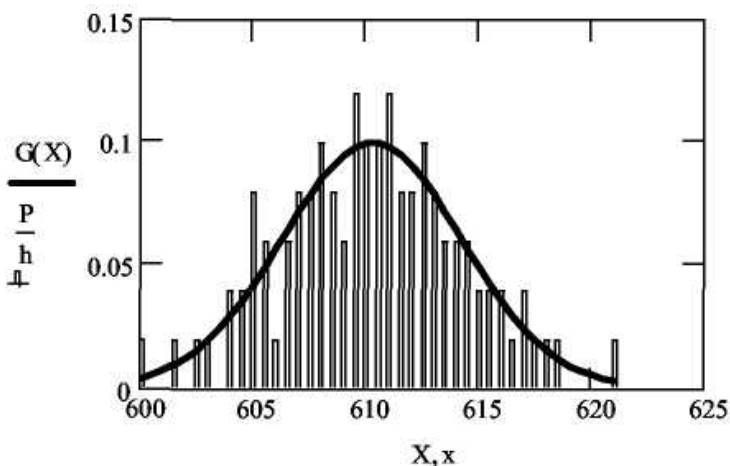


Рис. 5.13. Пример построения теоретического и статистического распределений в Mathcad

В. Создание и форматирование графика гистограммы

Для того чтобы создать график в виде гистограммы:

1. Постройте двумерный график, задайте переменные по осям и пределы оси абсцисс.
2. Войдите в диалоговое окно **Formatting Currently Selected Graph (Форматирование)** выбранного графика (например, двойным щелчком мыши) и перейдите на вкладку **Traces (Графики)**.
3. Установите для серии данных гистограммы в поле **Type (Тип)** элемент списка **bar (столбцы)** или **solidbar (гистограмма)**.
4. Нажмите кнопку ОК.

СОДЕРЖАНИЕ ЗАДАНИЯ

Решить задачи А и Б, изменив исходные данные, приведенные в таблице 5.1 путем умножения результатов определения количества меди на соответствующий множитель (табл. 5.3) в соответствии с заданным вариантом.

Таблица 5.3

вариант	множитель	вариант	множитель	вариант	множитель
1	0,987	11	0,997	21	1,008
2	0,988	12	0,998	22	1,009
3	0,989	13	0,999	23	1,010
4	0,990	14	1,001	24	1,011
5	0,991	15	1,002	25	1,012
6	0,992	16	1,003	26	1,013
7	0,993	17	1,004	27	1,014
8	0,994	18	1,005	28	1,015
9	0,995	19	1,006	29	1,016
10	0,996	20	1,007	30	1,017

ПРИЛОЖЕНИЕ 1

Таблица 6.1

Табличное значение критерия Кохрена

<i>N</i>	<i>m</i> – 1								
	1	2	3	4	5	7	9	16	36
2	0,998	0,975	0,939	0,906	0,877	0,833	0,801	0,734	0,660
3	0,967	0,871	0,798	0,746	0,707	0,653	0,617	0,547	0,475
4	0,906	0,768	0,684	0,629	0,589	0,536	0,502	0,437	0,372
5	0,841	0,684	0,598	0,544	0,506	0,456	0,424	0,364	0,307
6	0,781	0,616	0,532	0,480	0,445	0,398	0,368	0,314	0,261
7	0,727	0,561	0,480	0,431	0,397	0,354	0,326	0,276	0,228
8	0,680	0,516	0,438	0,391	0,360	0,318	0,293	0,246	0,202
9	0,638	0,478	0,403	0,358	0,329	0,290	0,266	0,223	0,182
10	0,602	0,445	0,373	0,331	0,303	0,267	0,244	0,203	0,166
15	0,471	0,335	0,276	0,242	0,220	0,191	0,174	0,143	0,114
20	0,389	0,270	0,220	0,192	0,174	0,150	0,136	0,111	0,088
30	0,293	0,198	0,159	0,138	0,124	0,106	0,096	0,077	0,060
60	0,174	0,113	0,090	0,076	0,068	0,058	0,052	0,041	0,032

ПРИЛОЖЕНИЕ 2

Таблица 6.2

Табличное значение критерия Фишера

<i>N</i> (<i>m</i> – 1)	<i>N</i> – <i>s</i>								
	1	2	3	4	5	6	8	10	15
2	18,51	19,00	19,16	19,25	19,30	19,33	19,37	19,40	19,43
3	10,13	9,55	9,28	9,12	9,01	8,94	8,85	8,79	8,70
4	7,71	6,94	6,59	6,39	6,26	6,16	6,04	5,96	5,86
5	6,61	5,79	5,41	5,19	5,05	4,95	4,82	4,74	4,62
6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4,06	3,94
7	5,59	4,74	4,35	4,12	3,97	3,87	3,73	3,64	3,51
8	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,35	3,22
9	5,12	4,26	3,86	3,63	3,48	3,37	3,23	3,14	3,01
10	4,96	4,10	3,71	3,48	3,33	3,22	3,07	2,98	2,85
12	4,75	3,89	3,49	3,26	3,11	3,00	2,85	2,75	2,62
14	4,60	3,74	3,34	3,11	2,96	2,85	2,70	2,60	2,46
16	4,49	3,63	3,24	3,01	2,85	2,74	2,59	2,49	2,35
18	4,41	3,55	3,16	2,93	2,77	2,66	2,51	2,41	2,27
20	4,35	3,49	3,10	2,87	2,71	2,60	2,45	2,35	2,20

ПРИЛОЖЕНИЕ 3

Таблица 6.3

Табличное значение критерия Стьюдента

$N(m-1)$	t_T	$N(m-1)$	t_T
1	12,7060	11	2,2010
2	4,3020	12	2,1788
3	3,1820	13	2,1604
4	2,7760	14	2,1448
5	2,5700	15	2,1314
6	2,4460	16	2,1190
7	2,3646	17	2,1098
8	2,3060	18	2,1009
9	2,2622	19	2,0930
10	2,2281	20	2,0860

ПРИЛОЖЕНИЕ 4

Таблица 6.4

Табличное значение функции Лапласа

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0,00	0,00000	1,00	0,34134	2,00	0,47725	3,00	0,49865	4,00	0,49997
0,05	0,01994	1,05	0,35314	2,05	0,47982	3,05	0,49886	4,05	0,49997
0,10	0,03983	1,10	0,36433	2,10	0,48213	3,10	0,49903	4,10	0,49998
0,15	0,05962	1,15	0,37493	2,15	0,48422	3,15	0,49919	4,15	0,49998
0,20	0,07926	1,20	0,38493	2,20	0,48610	3,20	0,49932	4,20	0,49999
0,25	0,09871	1,25	0,39435	2,25	0,48778	3,25	0,49928	4,25	0,49999
0,30	0,11791	1,30	0,40320	2,30	0,48928	3,30	0,49952	4,30	0,49999
0,35	0,13683	1,35	0,41149	2,35	0,49062	3,35	0,49960	4,35	0,49999
0,40	0,15542	1,40	0,41924	2,40	0,49180	3,40	0,49967	4,40	0,49999
0,45	0,17364	1,45	0,42647	2,45	0,49286	3,45	0,49972	4,45	0,50000
0,50	0,19146	1,50	0,43319	2,50	0,49379	3,50	0,49977	4,50	0,50000
0,55	0,20884	1,55	0,43943	2,55	0,49462	3,55	0,49981	4,55	0,50000
0,60	0,22575	1,60	0,44520	2,60	0,49534	3,60	0,49984	4,60	0,50000
0,65	0,24215	1,65	0,45053	2,65	0,49598	3,65	0,49987	4,65	0,50000
0,70	0,25804	1,70	0,45543	2,70	0,49654	3,70	0,49989	4,70	0,50000
0,75	0,27337	1,75	0,45994	2,75	0,49702	3,75	0,49991	4,75	0,50000
0,80	0,28814	1,80	0,46407	2,80	0,49745	3,80	0,49993	4,80	0,50000
0,85	0,30234	1,85	0,46784	2,85	0,49782	3,85	0,49994	4,85	0,50000
0,90	0,31594	1,90	0,47128	2,90	0,49814	3,90	0,49995	4,90	0,50000
0,95	0,32894	1,95	0,47441	2,95	0,49841	3,95	0,49996	4,95	0,50000

ПРИЛОЖЕНИЕ 5*Таблица 6.5***Табличное значение распределения χ^2**

r	P	0,99	0,95	0,90	0,80	0,50	0,20	0,10	0,05	0,01
1		0,00	0,01	0,02	0,06	0,45	1,64	2,71	8,84	6,64
3		0,12	0,35	0,58	1,01	2,37	4,64	6,26	7,82	11,8
5		0,65	1,14	1,61	2,84	4,35	7,29	9,24	11,1	15,1
7		1,24	2,17	2,38	3,82	6,35	9,80	12,0	14,1	18,5
9		2,09	3,32	4,17	5,38	8,34	12,2	14,7	16,9	21,7
12		3,57	5,23	6,30	7,81	11,3	16,8	18,5	21,0	28,2
16		5,23	7,26	8,55	10,8	14,8	19,8	22,8	25,0	34,6
18		7,02	9,39	10,9	12,9	17,8	22,8	26,0	28,9	34,8

СПИСОК РЕКОМЕНДОВАННОЙ ЛИТЕРАТУРЫ

Основная литература

1. *Власов, Константин Петрович.* Методы научных исследований и организации эксперимента: Учеб. пособие / Ред. А.А. Гальнбек; С.-Петерб. гос. горн. ин-т. Каф. печей, контроля и автоматизации металлургического производства. СПб: СПГГИ, 2000.

2. *Горелов, С.В.* Основы научных исследований: учебное пособие / С.В. Горелов, В.П. Горелов, Е.А. Григорьев; под ред. В.П. Горелова. - 2-е изд., стер. Москва; Берлин: Директ-Медиа, 2016

3. *Клинов, А.В.* Математическое моделирование химико-технологических процессов: учебное пособие / А.В. Клинов, А.Г. Мухаметзянова; Федеральное агентство по образованию, Государственное образовательное учреждение высшего профессионального образования "Казанский государственный технологический университет". - Казань: Казанский государственный технологический университет, 2009

4. Планирование научного эксперимента: Учебник / В.А.Волосухин, А.И.Тищенко, 2-е изд. М.: ИЦ РИОР, НИЦ ИНФРА-М, 2016

Дополнительная литература

1. *Диков, А.В.* Математическое моделирование и численные методы: учебное пособие / А.В. Диков, С.В. Степанова ; под ред. Г.В. Сугрובה. - Пенза : ПГПУ, 2000

2. Планирование и организация эксперимента [Электронный ресурс]: практикум / Новосиб. гос. аграр. ун-т. Биолого-технолог. фак; сост. И.А. Ленивкина. Новосибирск, 2012

3. Статистические методы обработки экспериментальных данных с использованием пакета MathCad: Учебное пособие/Ф.И.Карманов, В.А.Острейковский М.: КУРС, НИЦ ИНФРА-М, 2015

СОДЕРЖАНИЕ

Введение	1
Лабораторная работа № 1. Основы работы в Mathcad	4
1. Окно Mathcad	4
2. Примеры простых действий	7
3. Графики	9
4. Решение систем линейных и нелинейных уравнений и неравенств	13
5. Операции с векторами и матрицами	14
Содержание задания	19
Лабораторная работа № 2. Обработка экспериментальных данных по методу наименьших квадратов	20
Основы теории и расчетные формулы	20
Нахождение коэффициентов аппроксимирующего полинома в среде Mathcad	24
Содержание задания	30
Лабораторная работа № 3. Планирование эксперимента	33
Основы теории и расчетные формулы	33
Содержание задания	42
Лабораторная работа № 4. Разработка модели второго порядка	45
Основы теории и расчетные формулы	45
Содержание задания	49
Лабораторная работа № 5. Построение гистограмм распределения случайных величин	51
Основы теории и расчетные формулы	51
Способы задания гистограмм в MATHCAD	63
Содержание задания	70
Приложение 1	72
Приложение 2	72
Приложение 3	73
Приложение 4	73
Приложение 5	74
Список рекомендованной литературы	75