

ВВЕДЕНИЕ

Владение основами статистических методов необходимо специалистам, работающим в различных областях науки и техники. В настоящее время разработано большое количество статистических пакетов-программ, разделенных на две основных группы: специализированные пакеты и пакеты общего назначения. Среди универсальных статистических пакетов одним из лидеров является SPSS (Superior Performance Software System, что переводится, как «Система программного обеспечения высшей производительности») [1]. Пакет отличается гибкостью и мощностью применения для всех видов статистических расчетов. В России существует представительство компании SPSS, которое распространяет русифицированную версию пакета. На русском языке создан электронный учебник по применению данного пакета, изданы работы по его применению [2]. Среди универсальных систем статистического анализа данных широкое распространение получил пакет STATISTICA [3,4,5]. Пакет относится к числу базовых пакетов вузов России. Фирма разработчик имеет российское представительство [6], на сайте которого размещен электронный учебник по статистике, а также студенческая версия программы, распространяемая бесплатно. К числу достаточно мощных универсальных статистических пакетов относится также STATGRAPHICSPLUS [7], важнейшим достоинством пакета считаются хорошая интеграция математико-статистического аппарата обработки данных с современной интерактивной графикой и его динамичная эволюция с учетом развития компьютерных технологий.

Несмотря на разнообразие статистического программного обеспечения в России чаще всего используется программный комплекс (приложение) Microsoft Excel [4, 8, 9].

Это объясняется широким распространением русскоязычной версии данного ПО для персональных компьютеров. В программной среде MSOffice приложение MS Excel выполняет функции электронной таблицы с достаточно мощной математической поддержкой решения задач, в которой определенные статистические процедуры являются дополнительными встроенными формулами. Существует также макрос-дополнение XLSTAT-Pro [10, 11] для приложения MS Excel, включающее в себя более 50 статистических процедур.

Таким образом, благодаря современному уровню развития информационных технологий в распоряжении исследователей различных сфер науки и техники, экономики и производства, а также образовательных учреждений имеются доступные научные и статистические пакеты программ, удовлетворяющие разнообразные потребности пользователей.

ГЛАВА 1. ТЕХНОЛОГИИ СТАТИСТИЧЕСКОГО АНАЛИЗА ДАННЫХ

Общая технология статистического анализа данных с использованием типового статистического пакета включает в себя следующие основные этапы:

- ввод исходных данных в электронную таблицу и их предварительное преобразование перед анализом;
- визуализация данных при помощи того или иного типа графиков;
- определение подходящих методов статистической обработки;
- применение конкретного метода статистической обработки;

- вывод результата анализа в виде графиков и электронных таблиц с численной и текстовой информацией;
- подготовка, печать и сохранение отчета.

1.1. Определение целей и типа экспериментальных исследований

Согласно общепринятому определению [12], экспериментом называется постановка опытных исследований, включающая наблюдение за протекающим исследованием в точно учитываемых условиях. Эксперимент должен обладать свойством точного воссоздания его при повторении условий. Исходя из этого, можно сделать вывод, что само понятие «эксперимент» означает действие по созданию требуемых условий, чтобы осуществить конкретное явление, при минимизации сторонних воздействий на него со стороны других явлений.

Цель эксперимента заключается в определении свойств исследуемых объектов, проверке обоснованности гипотез и изучении темы научного исследования, основываясь на полученных данных. Постановка и организация экспериментального исследования определяется его назначением.

Различают разные типы экспериментов:

- по способу формирования условий (естественные и искусственно созданные условия);
- по цели исследования: преобразующие, констатирующие, поисковые, контролирующие, решающие;
- по организации проведения: лабораторные, производственные;

- по структуре изучаемых объектов и явлений: простые, сложные;
- по характеру внешних воздействий на объект исследования: вещественные, энергетические, информационные;
- по характеру взаимодействия средства экспериментального исследования с объектом исследования: обычный и модельный;
- по числу варьируемых факторов (однофакторный и многофакторный);
- по характеру изучаемых объектов или явлений (технологические, социометрические) и т.п.

Для классификации могут быть использованы и другие признаки, определяющиеся техническими условиями.

1.1.1. Обработка статистических данных при помощи MS Office Excel

Программа MS Office Excel применяется в тех случаях, когда требуется быстрая обработка больших объёмов данных. Она полезна операций статистической обработки данных, их последующего анализа, решения задач оптимизации, построения выходных диаграмм и графиков. Для такого рода задач применяют как основные средства программного продукта MS Office EXCEL, так и дополнительные надстройки.

Для расчетного анализа данных используются отдельные библиотеки модулей. Под модулем понимается внешняя процедура или программа на языке программирования высокого уровня, удовлетворяющая некоторым

дополнительным ограничениям, наиболее важными из которых являются:

- ограничения на способ аварийного завершения работы модуля;
- на способы связи по информации, например, на допустимость переменных внешнего типа и использование общей области памяти;
- на возможность передачи управления между модулями с помощью операторов вызова, расположенных в теле модуля;
- на использование операторов ввода-вывода.

Наиболее типовые расчетные модули современных статических пакетов условно делятся на три группы:

- описательная статистика и разведочный анализ исходных данных;
- статистическое исследование зависимостей;
- вспомогательные программы.

При помощи модуля описательной статистики и разведочного анализа исходных данных можно решить ряд следующих задач:

- анализ резко выделяющихся наблюдений;
- проверка статистической независимости рядов наблюдений;
- определение основных числовых характеристик и частотная обработка исходных данных (построение гистограмм, полигонов частот, вычисление выборочных средних, дисперсий и т.д.);
- расчет критериев однородности (средних, дисперсий, законов распределения и т.д.);

- определение критериев согласия (хи-квадрат, Колмогорова-Смирнова и др.);
- статистическое оценивание параметров;
- вычисление наиболее распространенных законов распределения вероятностей (нормального, Пуассона, хи-квадрат и некоторых других);
- визуализация анализируемых многомерных статистических данных.

Модуль статистического исследования зависимостей занимает достаточно объемную часть типового статистического пакета. Такой модуль включает решение следующих задач:

- корреляционно регрессионный анализ;
- дисперсионный анализ;
- планирование регрессионных экспериментов, выборочных обследований и др.

Вспомогательные программы расширяют возможности статистических пакетов и реализуют, в частности, оптимизационные алгоритмы, вычислительные процедуры, основанные на нейросетях, и генетических алгоритмах, задачи статистического моделирования на ЭВМ, которые являются полезными составными элементами компьютерных имитационных экспериментов, используемых при анализе сложных реальных систем.

1.1.2. Определение числовых характеристик экспериментальных данных при помощи пакета MS Excel

Необходимо выполнить следующую последовательность действий:

Запишите на рабочем листе Excel экспериментальные данные, согласно ячейкам: в столбец или в строку, после чего установите курсор на той ячейке, в которую будет занесено рассчитанное значение функции. Удобнее всего установить его в том же столбце, через ячейку ниже от введенных данных.

При помощи кнопки «Вставка функции» на стандартной панели инструментов вызовем диалоговое окно «Мастер функций» (рисунок 1).

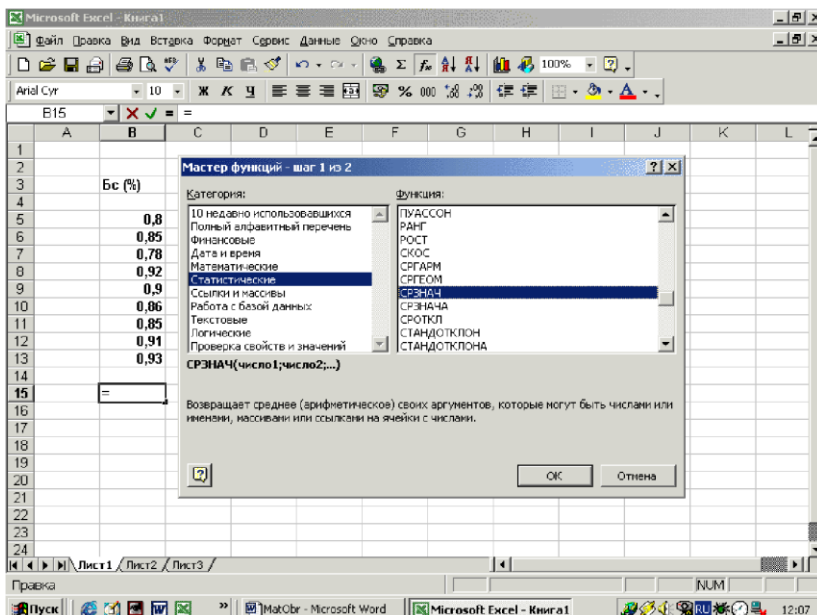


Рисунок 1 – Диалоговое окно «Мастер функций»

В диалоговом окне выберете категорию «Статистические».

Выберите необходимую функцию из списка, которая будет использоваться для последующей обработки данных. После подтверждения появится диалоговое окно (рисунок 2).

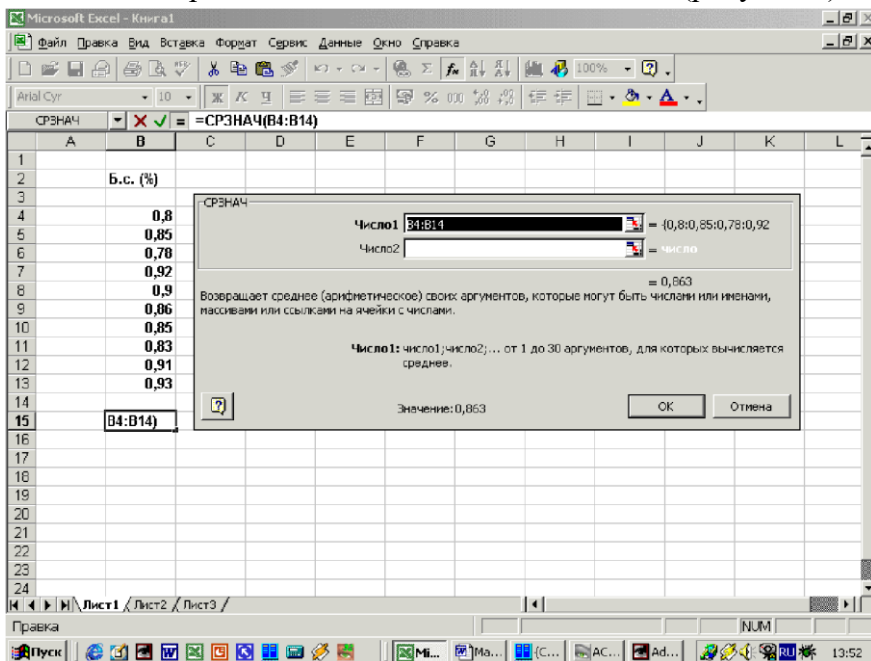


Рисунок 2 – Параметры настройки функции

В окне выбранной функции в нужный ряд чисел вводят данные. Для этого данные выделяют на листе, после чего те автоматически заносятся в требуемый ряд чисел.

В нижней части окна выводится рассчитанная числовая характеристика. При подтверждении она выводится на рабочий лист.

1.1.3. Построение экспериментального графика и диаграммы

Графическое представление позволяет облегчить восприятие и интерпретацию экспериментальных данных. Такое представление часто упрощает анализ и сравнение данных эксперимента. Результаты представлены диаграммами в виде полос, линий, столбиков, секторов, точек и в иной форме. Пакет Excel позволяет создавать диаграммы в виде внедрённых диаграмм и диаграммных страниц.

Внедрённые диаграммы – это диаграммы, наложенные на рабочий лист, содержащий таблицы с данными. Такие диаграммы сохраняются вместе с таблицей в одном файле. Для диаграммных страниц, создаются отдельные графические файлы. Внедрённые диаграммы создаются при помощи «Мастера диаграмм» (рисунок 3).

Мастер Диаграмм - это последовательность диалоговых окон, которая позволяет сделать все необходимые шаги для создания новой диаграммы или для изменения установок уже существующей диаграммы. Мастера Диаграмм вызывают с помощью кнопки расположенной на стандартной панели инструментов.

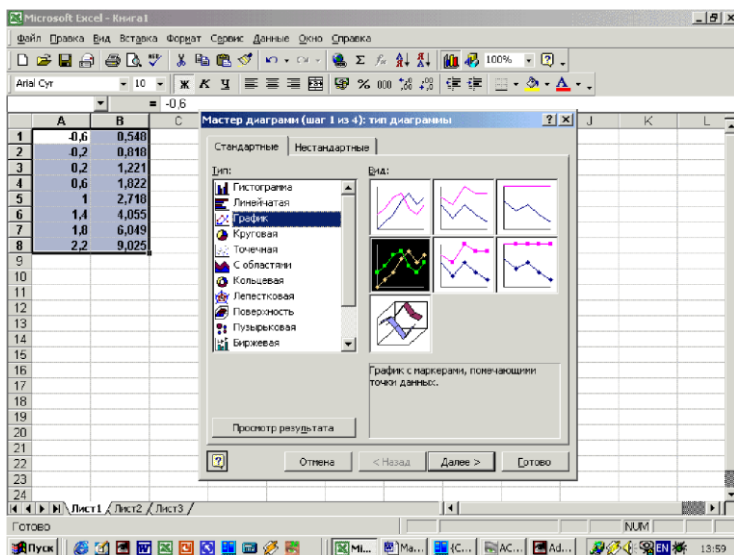


Рисунок 3 – Диалоговое окно «Мастер диаграмм»

Построение диаграмм включает последовательность от двух до пяти шагов, в зависимости от области выделенных данных. Если используется уже существующая диаграмма, то реализуются два шага. Если выделены данные на рабочем листе, то выполняются все пять шагов, поскольку в этом случае создаётся новая диаграмма.

На начальном этапе выбирают форму диаграммы. Доступные формы перечислены в списке «Тип» на вкладке «Стандартные» (Рисунок 3). Для выбранного типа диаграммы справа указываются несколько вариантов представления данных в палитре «Вид», из которых следует выбрать наиболее подходящий условиям эксперимента.

На следующем этапе выделяют диапазон данных, по которым будет строиться график. В случае если используется тип диаграммы «График», то для установления зависимости (Рисунок 4) следует установить параметр «Ряды в столбцах». Вкладка «Ряд» в диалоговом окне позволяет убрать

выделенный курсором ненужный ряд. Следует учитывать, что, например, при выборе типа диаграммы «Точечная» данные процедуры не нужны.

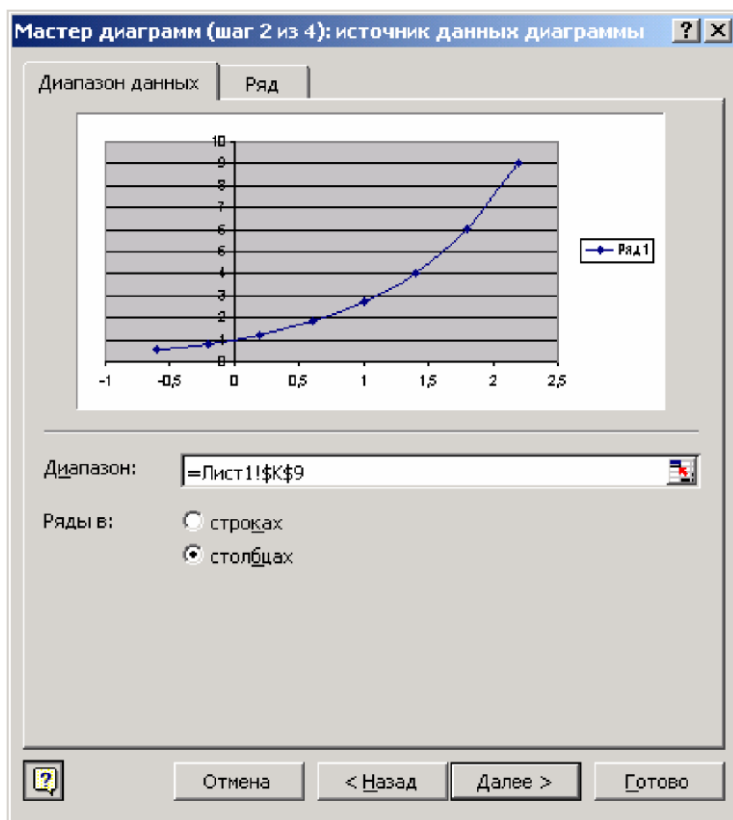


Рисунок 4 – Источник данных диаграммы

После этого необходимо перейти в окно «Параметры диаграммы» (Рисунок 5), щелкнув по кнопке «Далее». Вкладки данного окна позволяют оформить диаграмму.

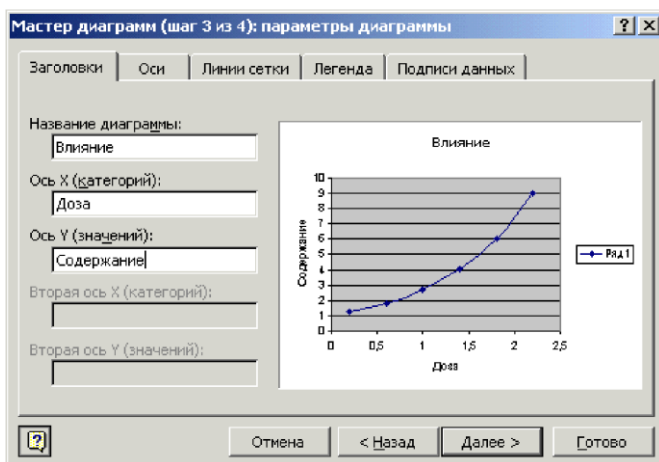


Рисунок 5 – Параметры диаграммы

Для установления функциональной зависимости экспериментальных данных необходимо выделить построенную диаграмму. При этом в меню приложения Excel появляется новый пункт – «Диаграмма». Затем необходимо вызвать команду «Диаграмма/Добавить линию тренда».

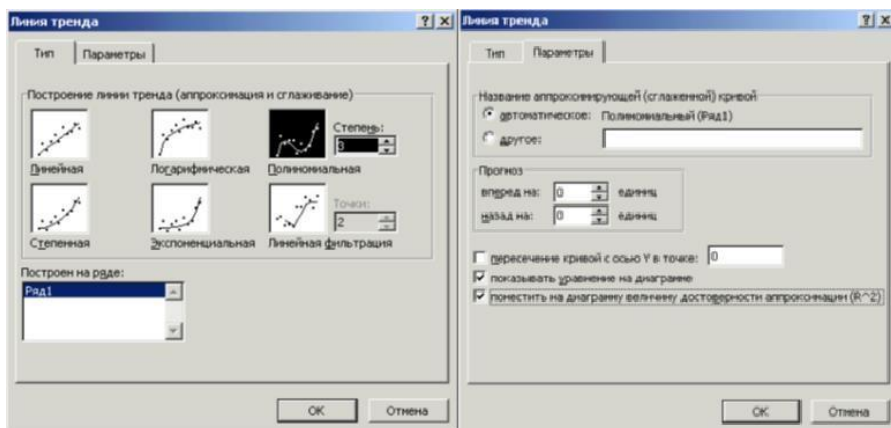


Рисунок 6 – Параметры линии тренда

Из предлагаемых кривых во вкладке «Тип», выбирается функция, которая наилучшим образом описывает исследуемый процесс. Для вывода уравнения функции во вкладке «Параметры» отмечаются пункты «Показать уравнение на диаграмме» и «Поместить величину достоверности аппроксимации R^2 ».

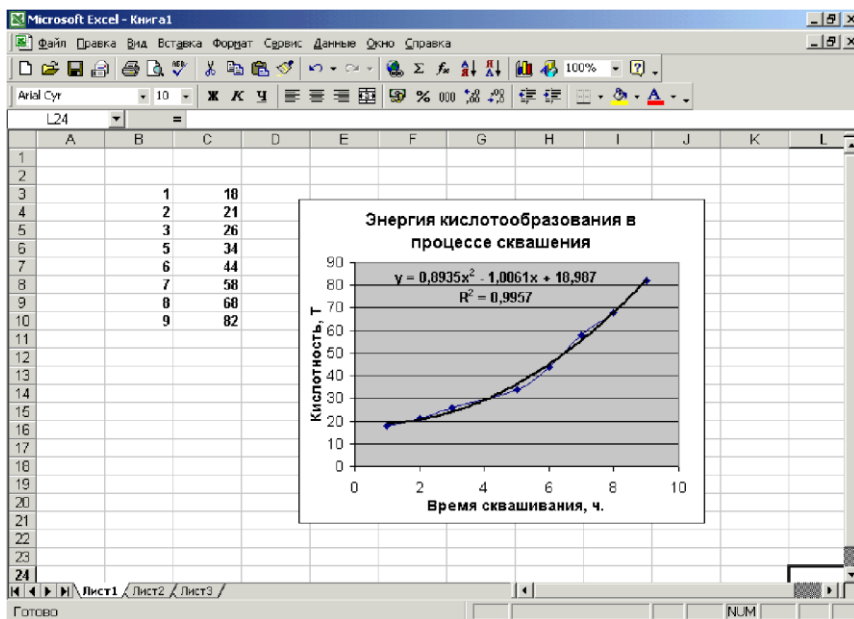


Рисунок 7 – Диаграмма на основе экспериментальных значений

Величина R^2 , должна иметь значение не менее 0,88 (при меньших значениях выбрать другой тип, приближающий R^2 к 1).

1.2 Обработка статистических данных при помощи программного пакета «Statistica»

Statistica – программный пакет, предназначенный для обработки статистических данных. Данный продукт имеет ряд встроенных функций для расчёта основных статистических характеристик, что значительно облегчает работу над большими массивами данных.

Порядок работы с пакетом Statistica:

Внести в лист данных параметры вариационного ряда.

Во вкладке «Статистика» выбрать «Статистика данных блоков», «Столбцы блока» (рисунок 8):

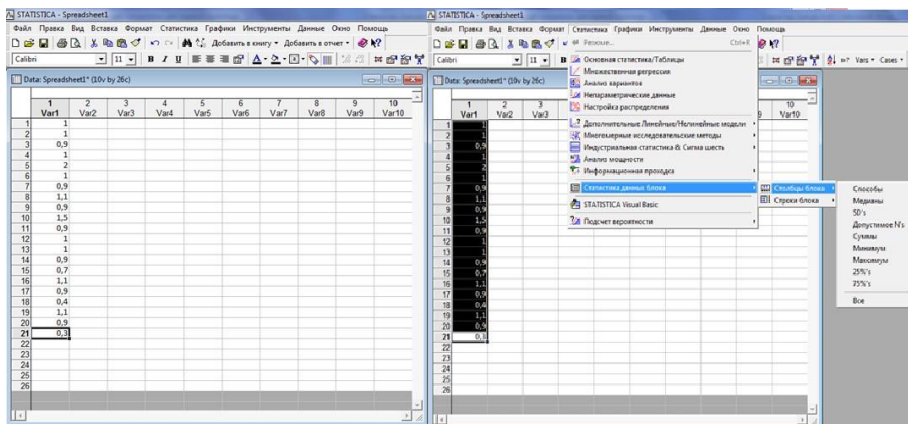


Рисунок 8 – Параметры вариационного ряда

Для расчёта медианы необходимо последовательно выбрать команды: «Статистика» / «Статистика данных блоков» / «Столбцы блока» / «Медианы». Для расчёта суммы

необходимо последовательно выбрать команды: «Статистика» / «Статистика данных блоков» / «Столбцы блока» / «Суммы» (рисунок 9).

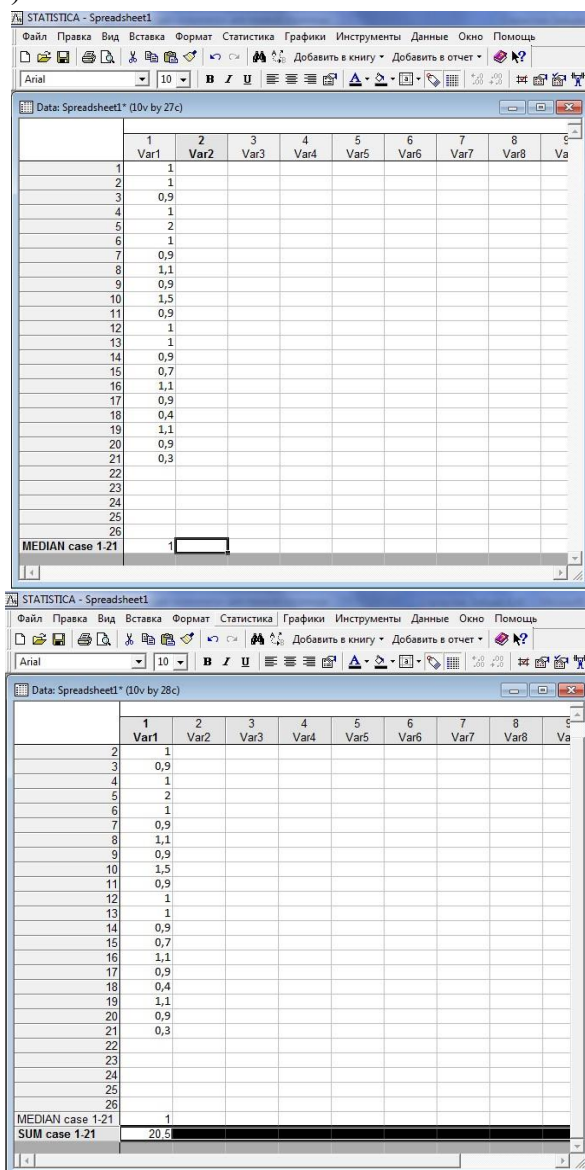


Рисунок 9 – Порядок расчета медианы вариационного ряда

Для расчёта минимального значения необходимо последовательно выбрать команды: «Статистика» / «Статистика данных блоков» / «Столбцы блока» / «Минимум».

Для расчёта максимального значения необходимо последовательно выбрать команды: «Статистика» / «Статистика данных блоков» / «Столбцы блока» / «Максимум» (Рисунок 10).

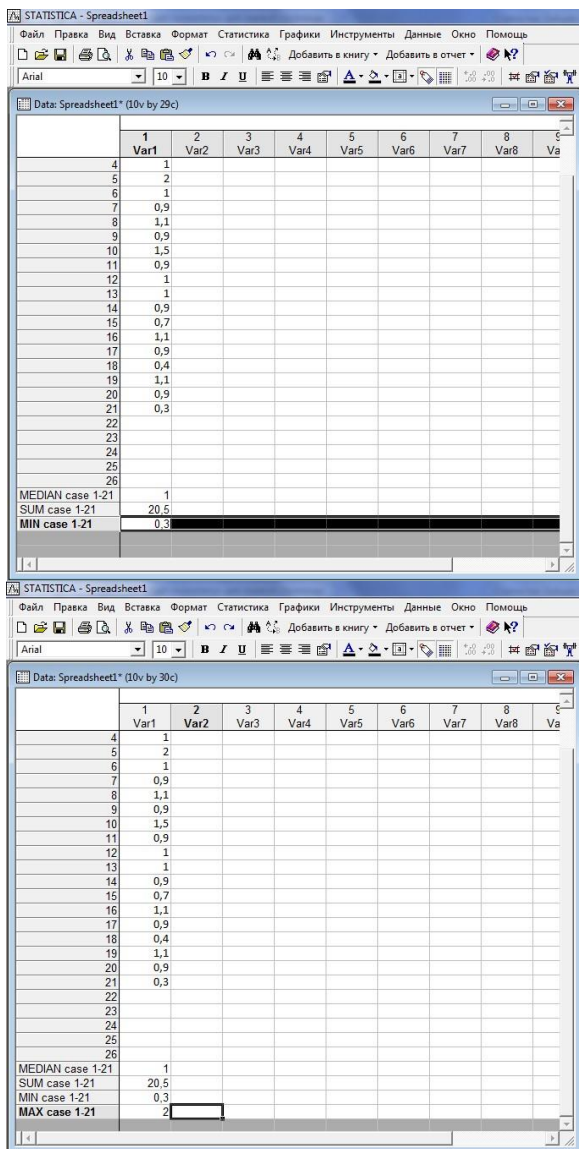


Рисунок 10 – Порядок расчета минимального и максимального значений вариационного ряда

Для расчёта и построения графика нормального распределения случайной величины необходимо последовательно выбрать команды: «Статистика» / «Подсчёт вероятностей» / «Распределения» / «Z-Normal» / «Создать график» / «Подсчёт».

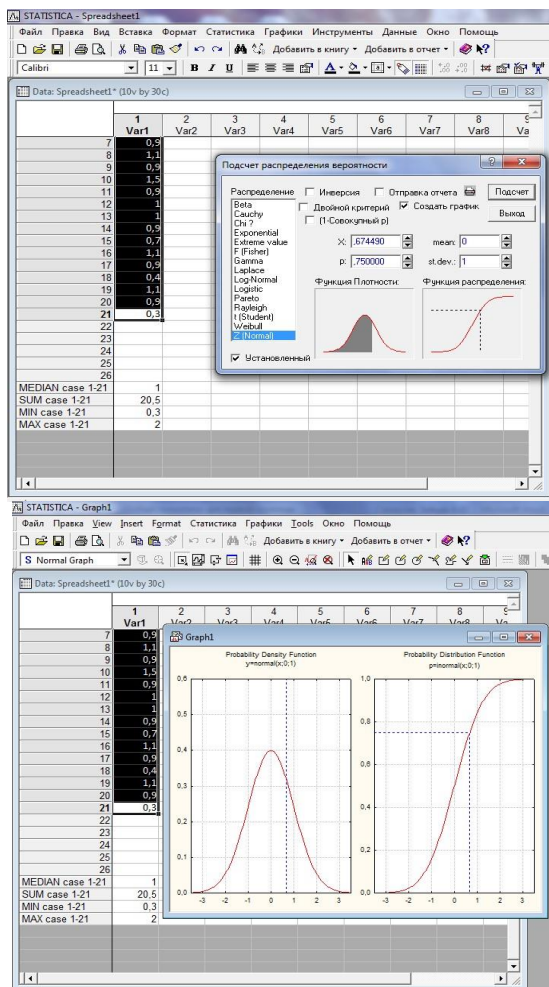
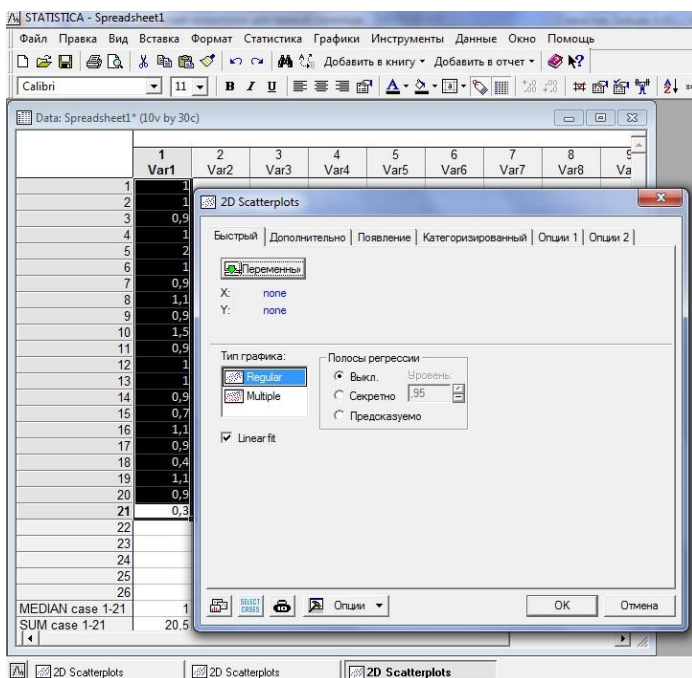
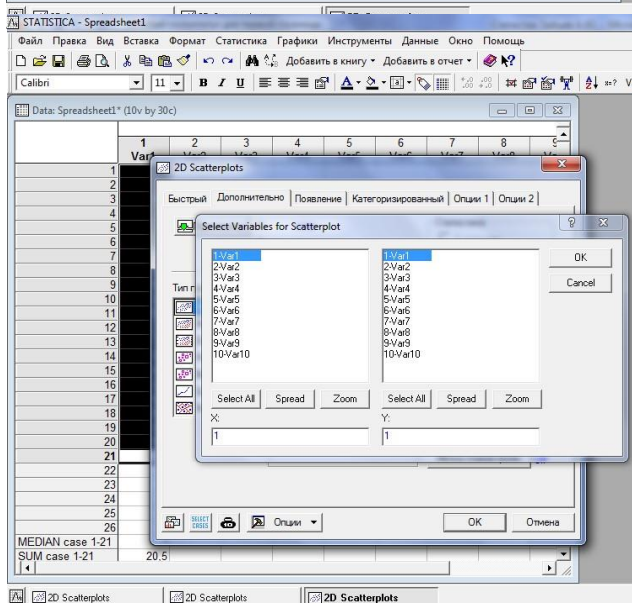
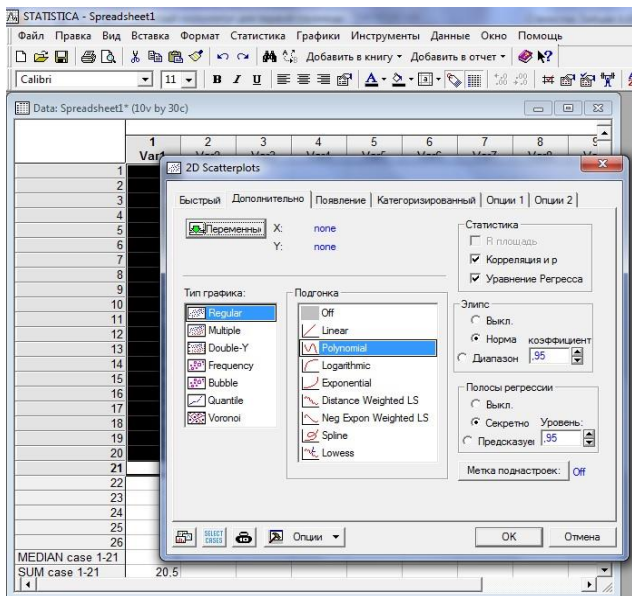


Рисунок 11 – Нормальное распределение случайной величины

Для расчёта и построения графика рассеяния случайной величины необходимо последовательно выбрать команды: «Графики» / «2-D Графики» / «Графики рассеяния» / «Дополнительно» / «Regular» / «Polinomial» / «Корреляция» / «Уравнение регрессии» / «Эллипс – норма» / «Полосы регрессии – секретно». Во всплывающем окне выбрать: «1-var1 справа и слева» / «Ок».

Результат построения представлен на Рисунке 12.





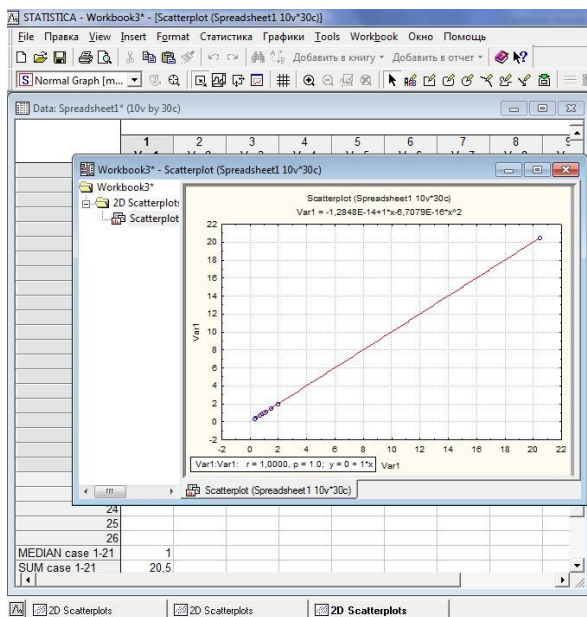


Рисунок 12 – Порядок построения графика рассеяния случайной величины

1.3. Обработка статистических данных при помощи программного пакета «Matlab»

Для статистических вычислений и статистической обработки данных при помощи программного пакета Matlab используется пакет прикладных программ Statistics and Machine Learning Toolbox, который реализует множество статистических функций:

- описательная статистика;
- распределения вероятностей;
- оценка параметров и аппроксимация;
- проверка гипотез;

- множественная регрессия;
- интерактивная пошаговая регрессия;
- моделирование Монте-Карло;
- аппроксимация на интервалах;
- статистическое управление процессами;
- планирование эксперимента;
- моделирование поверхности отклика;
- аппроксимация нелинейной модели;
- статистические графики.

Графический интерфейс пользователя приведен на рисунке 13.

Статистические вычисления в Matlab включают:

- исследовательский анализ данных;
- функции вероятности распределения;
- гипотетические тесты;
- регрессионный и дисперсионный анализы;
- машинное обучение;
- многомерный анализ данных;
- промышленная статистика.

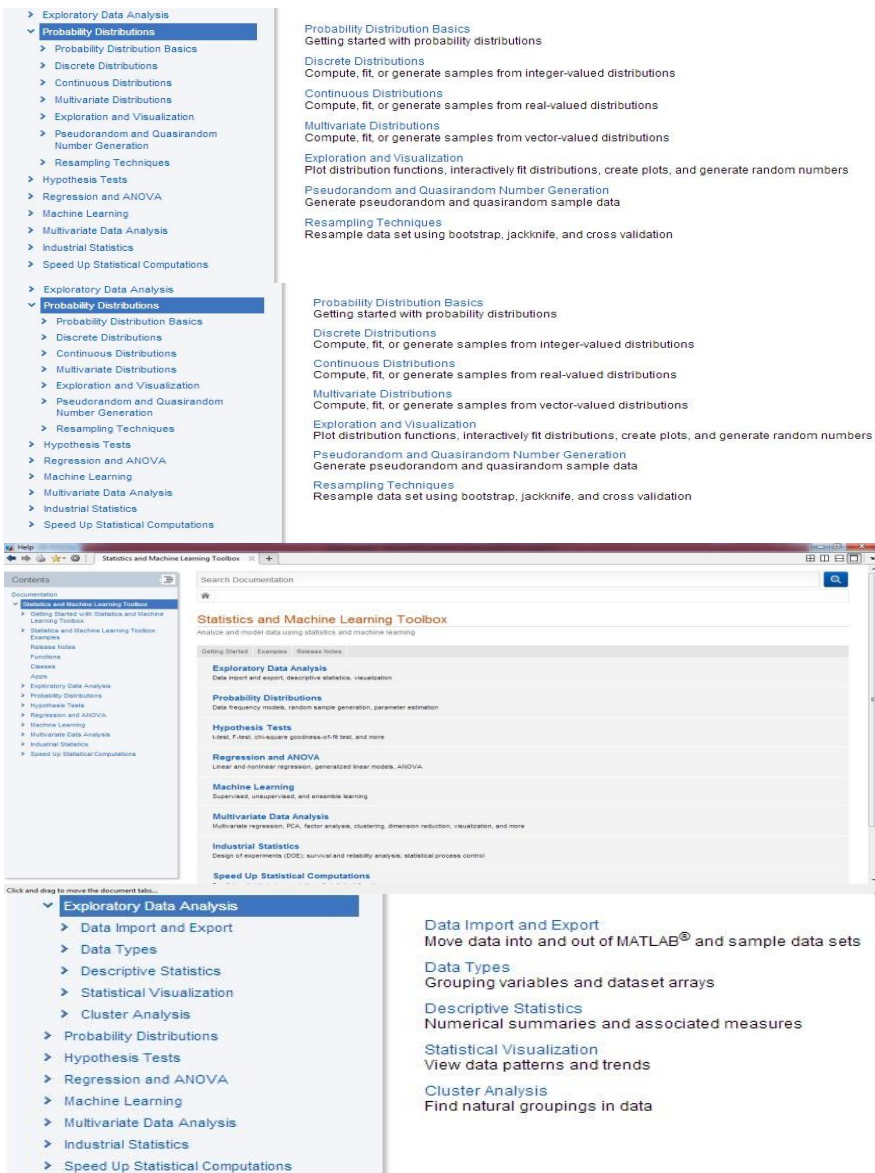


Рисунок 13 – Графический интерфейс пользователя.

1.3.1 Исследовательский анализ данных



Рисунок 14 – Исследовательский анализ данных

Statistics and Machine Learning Toolbox позволяет просматривать наши данные численно, создавая сводную статистику, включая измерения центральной тенденции, дисперсии, формы и корреляции.

Исследовательский анализ данных содержит в себе:

- импорт и экспорт данных;
- перемещение данных в и из MATLAB и наборов выборочных данных;

- группировка переменных и массивов наборов данных;
- описательная статистика;
- численные резюме и связанные с ними меры;
- статистическая визуализация;
- просмотр шаблонов данных и трендов;
- анализ кластеров;
- поиск естественных группировок в данных.

1.3.2. Функции вероятности распределений

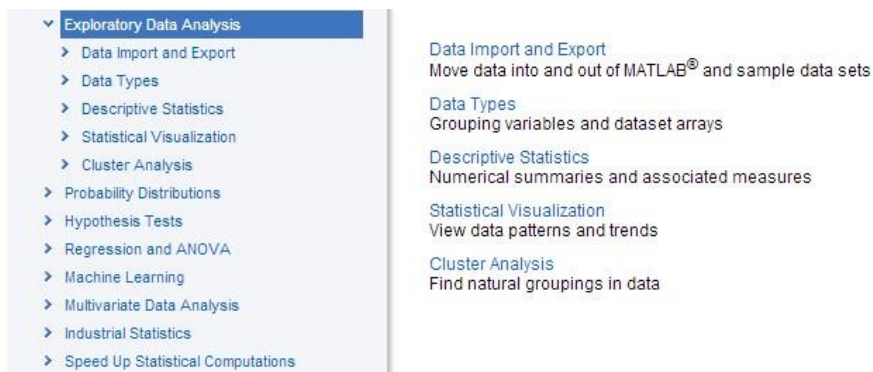


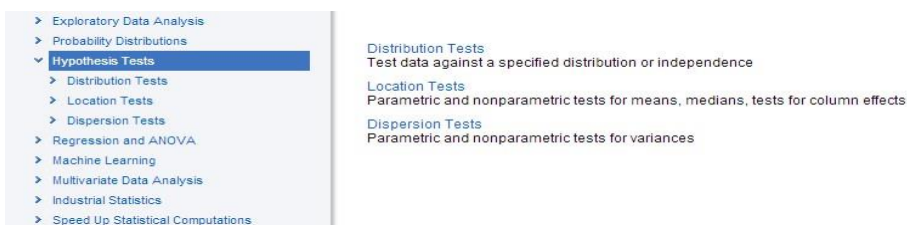
Рисунок 15 – Функции вероятности распределений

Statistics and Machine Learning Toolbox позволяет работать с дискретными, непрерывными и непараметрическими распределениями вероятностей.

Функции вероятности распределений содержат:

- основы распределения вероятностей: начало работы с распределением вероятностей;
- дискретные распределения: вычислить, сопоставить или сгенерировать образцы из целочисленных распределений;

- непрерывные распределения: вычислить, сопоставить или сгенерировать образцы из вещественных распределений;
- многомерные распределения: вычислить, сопоставить. или сгенерировать образцы из векторных распределений;
- исследование и визуализация: функции распределения графиков, интерактивное сопоставление распределений, создание графиков и генерация случайных чисел;
- генерация псевдослучайных и квазислучайных чисел:



генерировать псевдослучайные и квазислучайные данные выборки;

- методы повторной выборки.

Рисунок 16 – Гипотетические тесты

1.3.3 Гипотетические тесты

Statistics and Machine Learning Toolbox предоставляет множество параметрических и непараметрических тестов гипотез. Мы можем протестировать заданное распределение, среднее, медианное или дисперсионное значение (дисперсия).

Гипотетические тесты содержат в себе:

- тесты распределения: данные тестирования против определенного распределения или независимости;
- тесты местоположения: параметрические и непараметрические тесты для сред, тесты для эффектов столбцов;
- дисперсионные тесты: параметрические и непараметрические тесты на отклонения.

1.3.4 Регрессионный и дисперсионный анализы

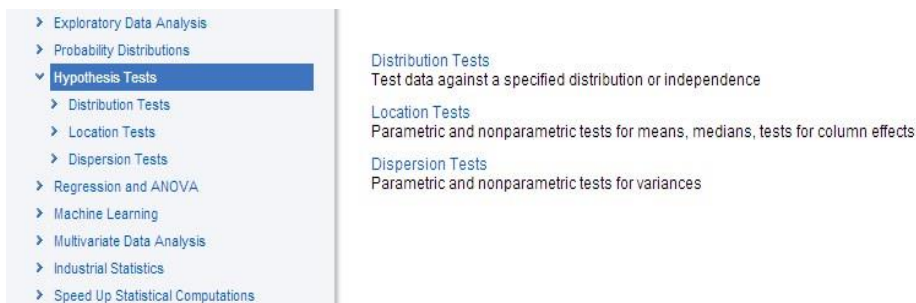


Рисунок 17 – Регрессионный и дисперсионный анализы

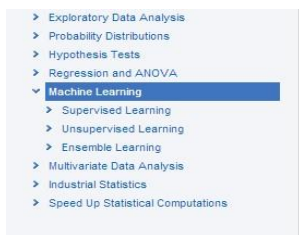
Модели регрессии описывают взаимосвязь между переменной ответа и одной или несколькими предикторными переменными.

Анализ дисперсии (ANOVA) - это процедура определения того, возникает ли изменение в переменной ответа внутри или среди разных групп населения.

Регрессионный и дисперсионный анализы содержат в себе:

- линейная регрессия: множественные, ступенчатые, многомерные модели регрессии и многое другое;

- обобщенные линейные модели: логистическая регрессия, многочленная регрессия, регрессия Пуассона и многое другое;
- нелинейная регрессия: нелинейные модели регрессии с фиксированным и смешанным эффектами;
- дисперсионный анализ: анализ дисперсии и



Supervised Learning
Regression, support vector machines, parametric and nonparametric classification, decision trees

Unsupervised Learning
Clustering, Gaussian mixture models, hidden Markov models

Ensemble Learning
Ensembles for boosting, bagging, or random subspace

ковариации, одномерный и многомерный дисперсионный анализ, повторные измерения дисперсионного анализа.

1.3.5 Машинное обучение

Рисунок 18 – Машинное обучение

Целью машинного обучения является построение модели, которая принимает решения на основе доказательств в условиях неопределенности. Поскольку адаптивные алгоритмы идентифицируют шаблоны в данных, компьютер «учится» из наблюдений. При использовании дополнительных наблюдений компьютер улучшает эффективность принятия решений.

Машинное обучение содержит в себе:

- контролируемое обучение (с учителем): регрессия, опорные векторные машины, параметрическая и непараметрическая классификация, деревья решений;

- неконтролируемое обучение (без учителя):

<ul style="list-style-type: none"> › Exploratory Data Analysis › Probability Distributions › Hypothesis Tests › Regression and ANOVA › Machine Learning ▼ Multivariate Data Analysis <ul style="list-style-type: none"> › Multivariate Regression › Repeated Measures and MANOVA › Principal Component Analysis (PCA) and Canonical Correlation › Factor Analysis › Nonnegative Matrix Factorization › Multidimensional Scaling › Cluster Analysis › Procrustes Analysis › Multivariate Visualization › Industrial Statistics › Speed Up Statistical Computations 	<p>Multivariate Regression Linear regression with multiple response variables</p> <p>Repeated Measures and MANOVA Analysis of variance, repeated measures modeling, and multiple comparisons for data with multiple responses</p> <p>Principal Component Analysis (PCA) and Canonical Correlation Transform data to a lower-dimensional space using rotation and projection</p> <p>Factor Analysis Model the covariance structure of multivariate data</p> <p>Nonnegative Matrix Factorization Approximate factorization of nonnegative matrix into nonnegative components</p> <p>Multidimensional Scaling Find a low-dimensional representation of data to match a distance matrix</p> <p>Cluster Analysis Find natural groupings in data</p> <p>Procrustes Analysis Fit one data set to another by rotation, translation, uniform scaling, and reflection</p> <p>Multivariate Visualization Visualize multidimensional data</p>
--	--

кластеризация, смеси гауссовских моделей, скрытые марковские модели;

- ансамбль обучений: ансамбли для ускорения, упаковки или случайного подпространства.

1.3.6 Многомерный анализ данных

Рисунок 19 – Многомерный анализ данных

Многовариантная статистика изучает взаимосвязь между множественными измерениями, наблюдаемыми у субъекта, и прогностическими переменными. Statistics and Machine Learning Toolbox предоставляет ряд методов для подгонки моделей к нескольким ответам, уменьшения размеров высокоразмерных данных и определения естественных группировок среди точек данных.

Многомерный анализ данных включает в себя:

- многомерная регрессия: линейная регрессия с несколькими переменными ответа;

- повторные измерения и MANOVA: анализ дисперсии, моделирование повторяющихся мер и множественное сравнение данных с несколькими ответами;
- анализ основных компонентов (PCA) и каноническая корреляция: преобразование данных в более низкое пространство с использованием вращения и проецирования;
- факторный анализ: модель ковариационной структуры многомерных данных;
- неотрицательная матричная факторизация: приближенная факторизация неотрицательной матрицы в неотрицательные компоненты;
- многомерное масштабирование: найти низкоразмерное представление данных в соответствии с матрицей расстояний;
- анализ кластеров: поиск естественных группировок в данных;
- прокрустов анализ: установите один набор данных в другой путем вращения, трансляции, равномерного масштабирования и отражения;
- многомерная визуализация: визуализировать многомерные данные.

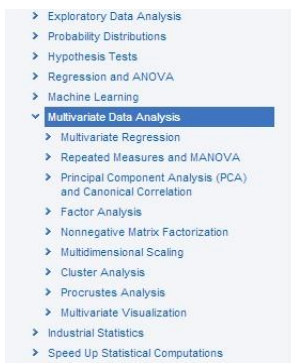
1.3.7 Промышленная статистика

Рисунок 20 – Промышленная статистика

Statistics and Machine Learning Toolbox предоставляет инструменты для проектирования экспериментов, анализа надежности и данных о выживании, контроля качества процесса и наблюдения данных.

Промышленная статистика включает в себя:

- разработка экспериментов: планирование экспериментов с систематическим сбором данных.
- анализ данных о продолжительности жизни: непараметрические и полупараметрические методы анализа данных о надежности и выживаемости.
- статистическое управление процессами: статистические методы контроля качества и контроля



Multivariate Regression
Linear regression with multiple response variables

Repeated Measures and MANOVA
Analysis of variance, repeated measures modeling, and multiple comparisons for data with multiple responses

Principal Component Analysis (PCA) and Canonical Correlation
Transform data to a lower-dimensional space using rotation and projection

Factor Analysis
Model the covariance structure of multivariate data

Nonnegative Matrix Factorization
Approximate factorization of nonnegative matrix into nonnegative components

Multidimensional Scaling
Find a low-dimensional representation of data to match a distance matrix

Cluster Analysis
Find natural groupings in data

Procrustes Analysis
Fit one data set to another by rotation, translation, uniform scaling, and reflection

Multivariate Visualization
Visualize multidimensional data

производственных процессов.

1.3.8 Контейнер table

The screenshot shows the MATLAB workspace with a variable named `carData` of type `2595x24 table`. The table is displayed in the Variables window, showing the first 14 rows. The columns are: Year, MfrName, CarLine, Car_Truck, EngDisp, Police, RatedHP, Transmission, and Drive.

	1	2	3	4	5	6	7	8	9
	Year	MfrName	CarLine	Car_Truck	EngDisp	Police	RatedHP	Transmission	Drive
1	2007	Daimler...	'RAM 15...	'truck'	348 'N'		345 'L5'		'R'
2	2007	Daimler...	'RAM 15...	'truck'	348 'N'		345 'L5'		'R'
3	2007	Daimler...	'LIBERT...	'truck'	226 'N'		210 'L4'		'R'
4	2007	Daimler...	'LIBERT...	'truck'	226 'N'		210 'L4'		'R'
5	2007	Daimler...	'LIBERT...	'truck'	226 'N'		210 'L4'		4
6	2007	Daimler...	'LIBERT...	'truck'	226 'N'		210 'L4'		4
7	2007	Daimler...	'PT CR...	'truck'	148 'N'		220 'L4'		'F'
8	2007	Daimler...	'PT CR...	'truck'	148 'N'		220 'L4'		'F'
9	2007	Daimler...	'PT CR...	'truck'	148 'N'		220 'M4'		'F'
10	2007	Daimler...	'PT CR...	'truck'	148 'N'		220 'M4'		'F'
11	2007	Daimler...	'PT CR...	'truck'	148 'N'		220 'M4'		'F'
12	2007	Daimler...	'CARAV...	'truck'	148 'N'		150 'L4'		'F'
13	2007	Daimler...	'CARAV...	'truck'	148 'N'		150 'L4'		'F'
14	2007	Daimler...	'CARAV...	'truck'	148 'N'		150 'L4'		'F'

Рисунок 21 – Контейнер table

Statistics and Machine Learning Toolbox поддерживает все типы данных, которые поддерживает сам Matlab, кроме того он содержит свои типы данных.

Один из удобных контейнеров, которые можно использовать для статистических вычислений – это тип `table`, который может включать в себя любой тип данных. Он содержит численный тип данных, строковый и т.д.

1.3.9 Работа с распределениями

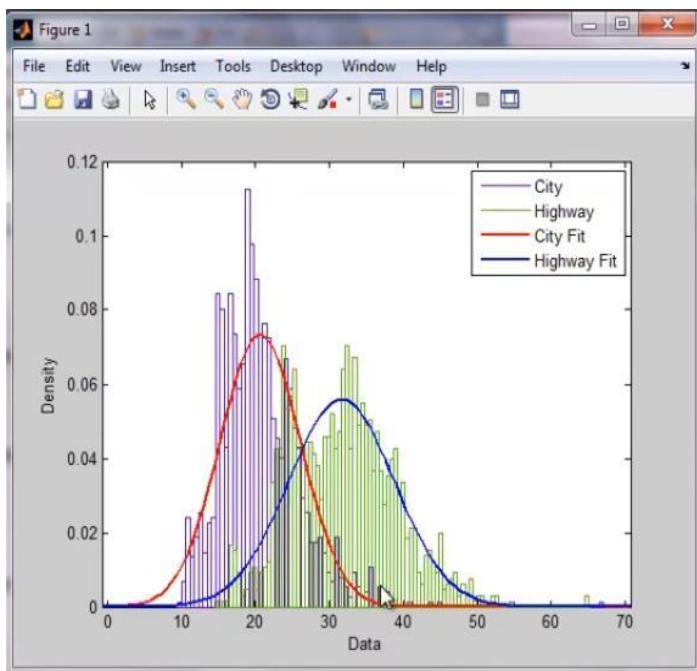


Рисунок 22 – Пример исследовательского анализа данных

На основе данных, которые у нас уже есть, проведем небольшой исследовательский анализ.

Обращение к любому столбцу и любым элементам происходит с помощью круглых скобок. Выделяем значения, которые нам нужны. Выделяем с 1 по 5 элементы колонки MPG.

Для получения сводной статистики из таблицы применяем команду `summary`. Из контейнера данных выделяем колонки, которые нас интересуют. Двоеточие обозначает, что мы выделяем все строки данных в этих колонках. Таким образом мы получаем нашу статистику. Здесь показывается сколько переменных, какие они: минимальные, средние – медиана и максимум.

1.3.10 Визуализация данных

Визуализация данных происходит путем обращения к колонкам данных из таблицы. Применяем команду `plotMPG`. На рисунке 23 показывается расход топлива в зависимости от мощности двигателя.

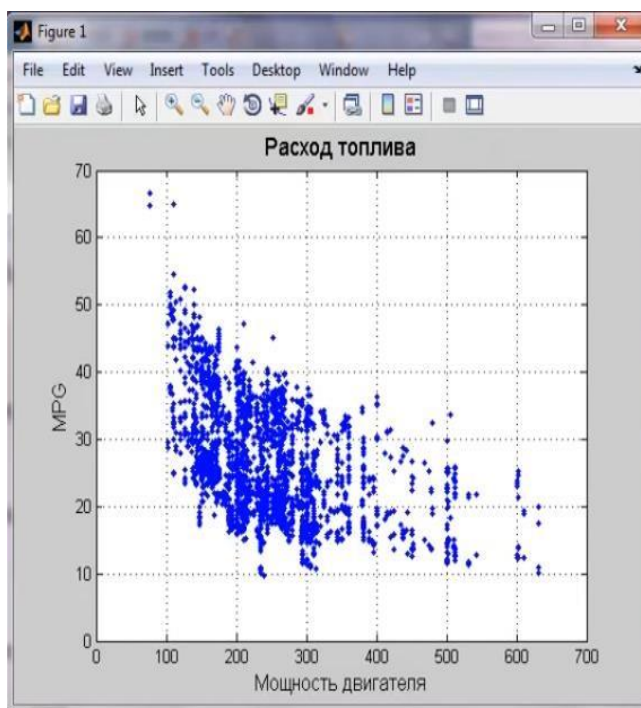


Рисунок 23 – Пример визуализации данных

1.4. Обработка статистических данных с помощью программного пакета «Statistica»

Statistica – программный пакет, предназначенный для обработки статистических данных [7, 13]. Данный продукт имеет ряд встроенных функций для расчёта основных статистических характеристик, что значительно облегчает работу над большими массивами данных.

Рассмотрим несколько примеров расчетов в пакете «Statistica». Для начала необходимо внести вариационный ряд, как показано на рисунке 24.

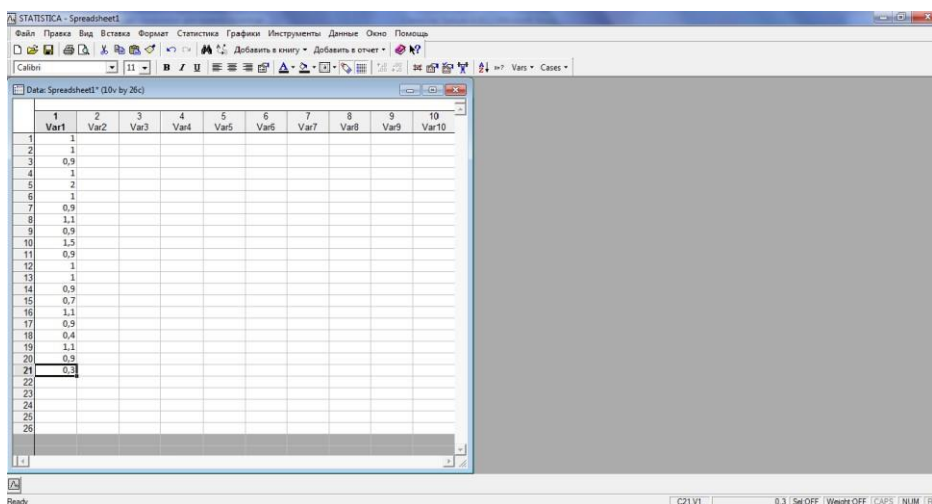


Рисунок 24 – Вариационный ряд

Во вкладке статистика выбрать статистика данных блоков, столбцы блока представлены на рисунке 25.

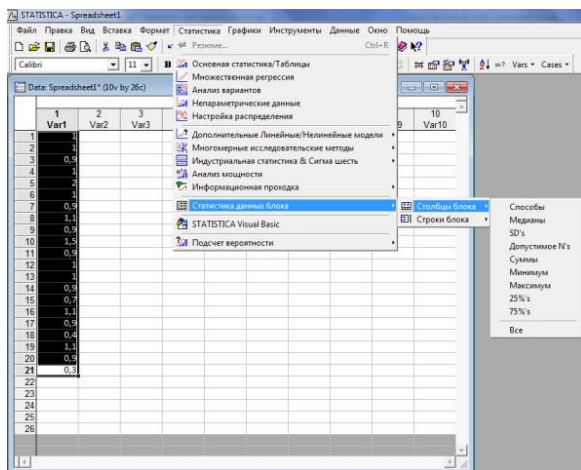


Рисунок 25 – Статистика данных блока

Для расчёта медианы необходимо выбрать: Статистика / Статистика данных блоков / Столбцы блока / Медианы на рисунке 26.

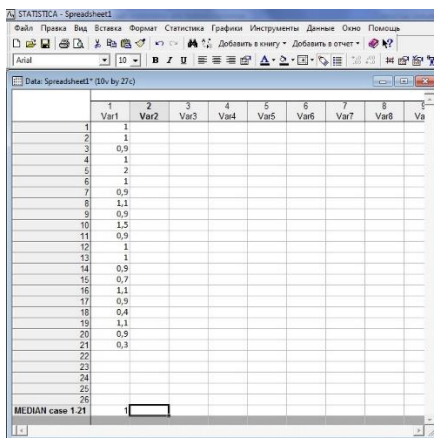


Рисунок 26 – Статистика / Статистика данных блоков / Столбцы блока / Медианы

Для расчёта суммы необходимо выбрать: Статистика / Статистика данных блоков / Столбцы блока / Суммы на рисунке 27.

STATISTICA - Spreadsheet1

Файл Правка Вид Вставка Формат Статистика Графики Инструменты Данные Окно Помощь

Добавить в книгу Добавить в отчет

Анал 10 B I U

Data: Spreadsheet* (10v by 28c)

	1	2	3	4	5	6	7	8	9
	Var1	Var2	Var3	Var4	Var5	Var6	Var7	Var8	Var9
2	1								
3	0,9								
4	1								
5	2								
6	1								
7	0,9								
8	1,1								
9	0,9								
10	1,5								
11	0,9								
12	1								
13	1								
14	0,9								
15	0,7								
16	1,1								
17	0,9								
18	0,4								
19	1,1								
20	0,9								
21	0,3								
22									
23									
24									
25									
26									
MEDIAN case 1-21	1								
SUM case 1-21	20,5								

Рисунок 27 – Статистика / Статистика данных блоков / Столбцы блока / Суммы

Для расчёта мин. значения необходимо выбрать: Статистика / Статистика данных блоков / Столбцы блока / Минимум (рисунок 28).

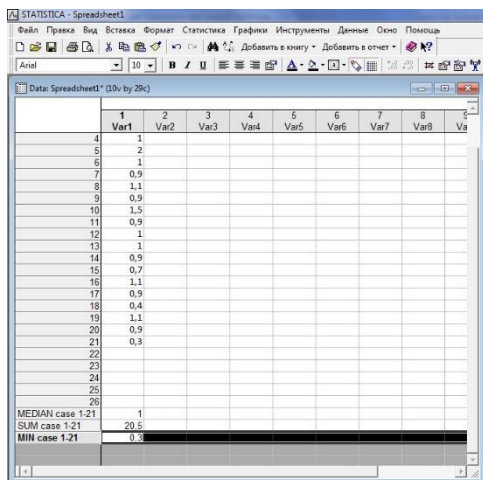


Рисунок 28 – Статистика/Статистика данных
блоков/Столбцы блока/Минимум

Для расчёта макс. значения необходимо выбрать:
Статистика/Статистика данных блоков / Столбцы блока /
Максимум (рисунок 29).

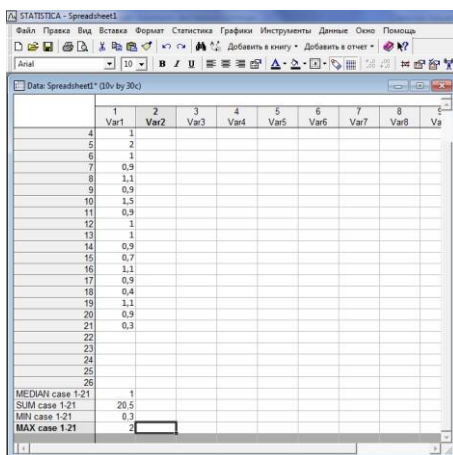


Рисунок 29 – Статистика / Статистика данных блоков /
Столбцы блока / Максимум

Расчёт и построение графика нормального распределения (Рисунок 30,31): Статистика / Подсчет вероятностей / Распределения / Z-Normal / Создать график (поставить галочку) / Подсчет.

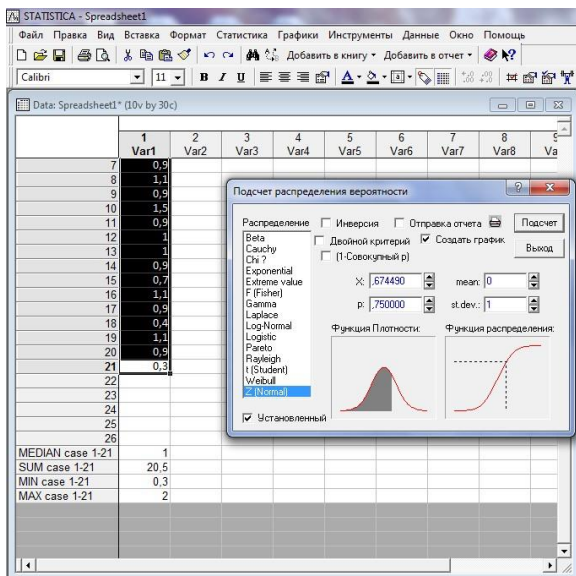


Рисунок 30 – Статистика / Подсчёт вероятностей / Распределения / Z-Normal / Создать график (поставить галочку) / Подсчёт

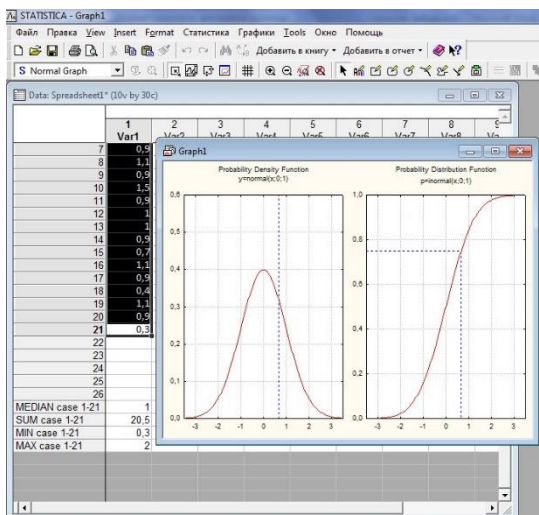


Рисунок 31 – Статистика / Подсчёт вероятностей / Распределения / Z-Normal / Создать график (поставить галочку) / Подсчёт

Построение графика рассеяния (рисунок 32): Графики / 2-D Графики / Графики рассеяния:

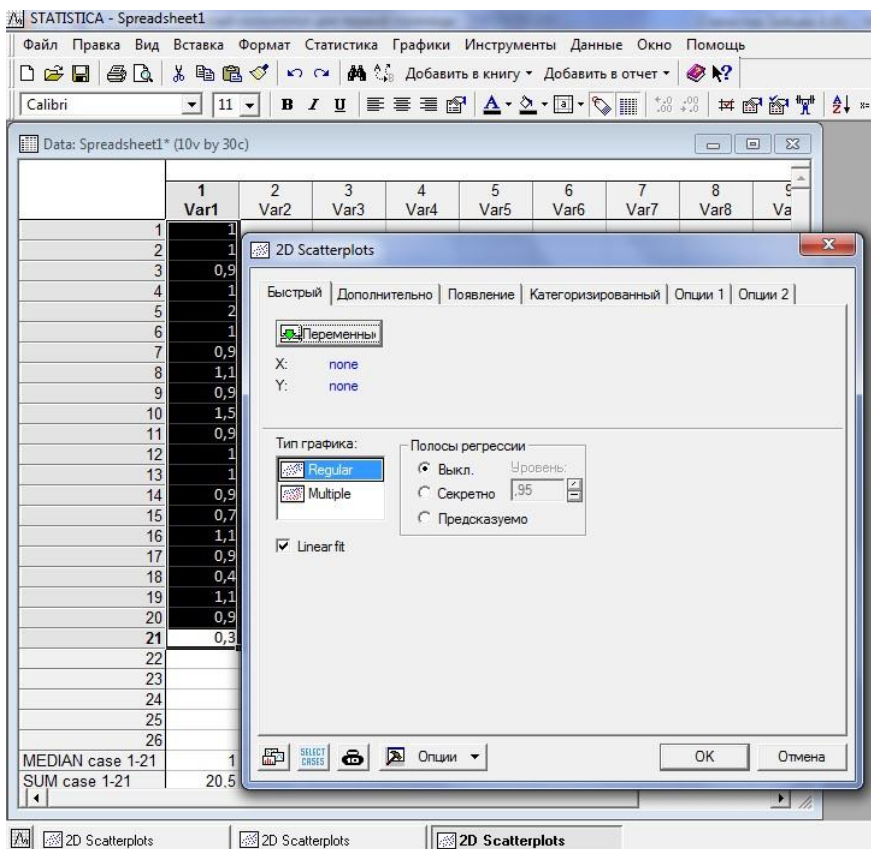


Рисунок 32 – Графики / 2-D Графики / Графики рассеяния

... / Дополнительно / Regular / Polinomial / Корреляция, Уравнение регрессии (поставить галочки) / Элипс – норма; Полосы регрессии – секретно / Ок (рисунок 33) / ...

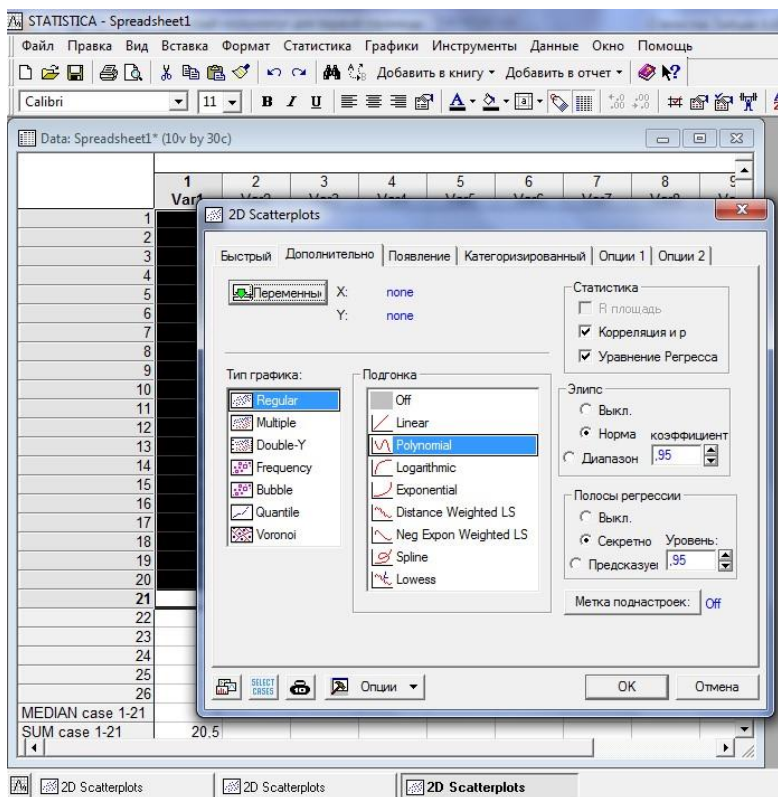


Рисунок 33 – ... / Дополнительно / Regular / Polinomial /
Корреляция, Уравнение регрессии (поставить галочки) /
Эллипс – норма; Полосы регрессии – секретно / Ок

Во всплывающем окне выбрать: 1-var1 справа и слева
(рисунок11) / Ок.

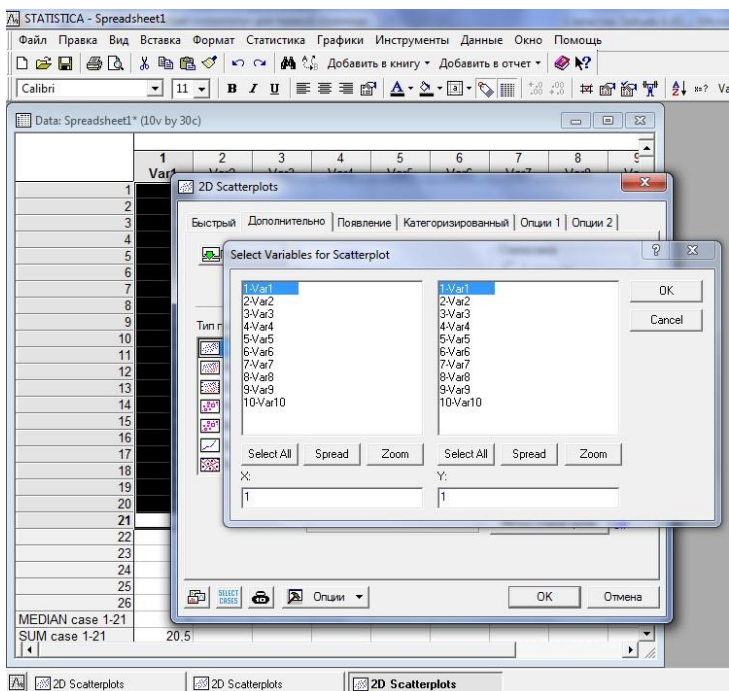


Рисунок 34 – Всплывающее окно

Результат построения представлен на рисунке 35.

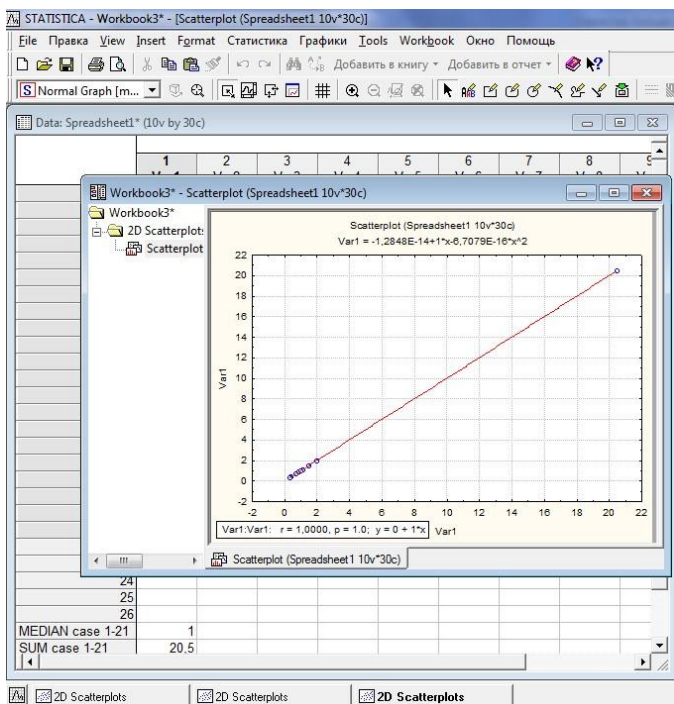


Рисунок 35 – Результат построения

1.4.1 Многомерная групповая визуализация

Визуализация на различных осях графиков, также различных данных. Применяется команда `gplotmatrix`.

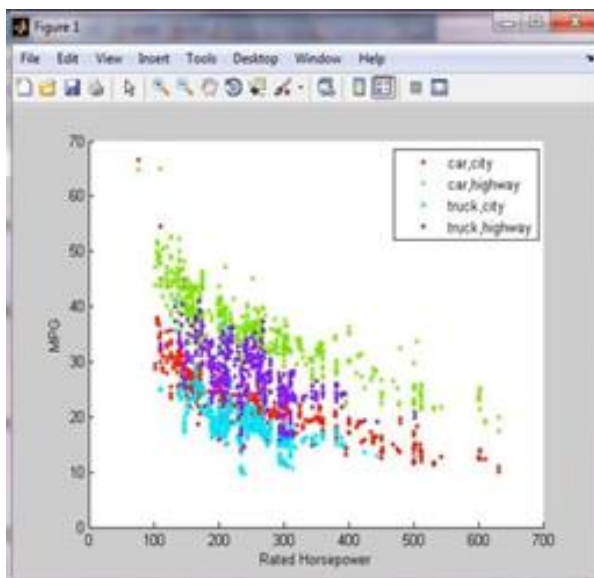


Рисунок 36 – Пример графика

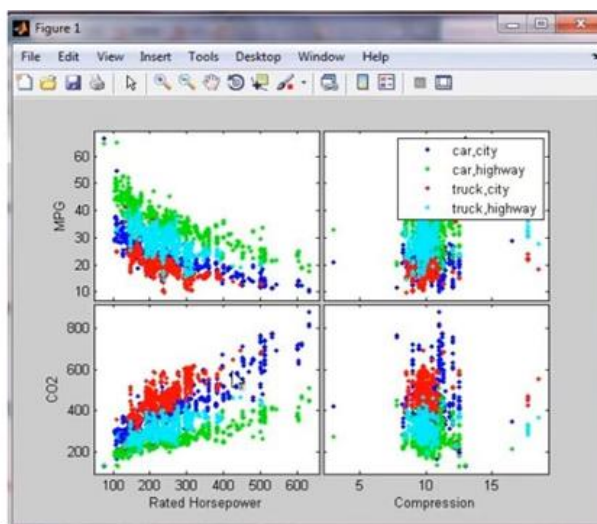


Рисунок 37 – Пример графика

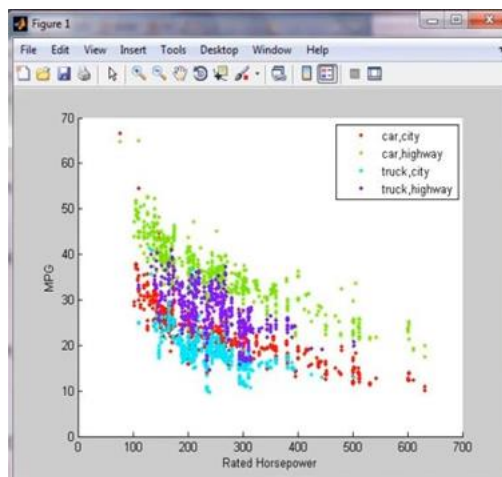


Рисунок 38 – Пример графика

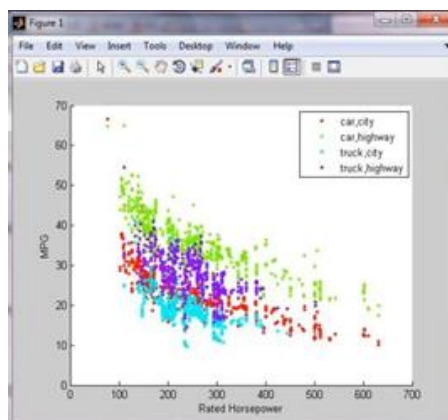


Рисунок 39 – Пример графика

1.4.2 Исследование многомерных данных

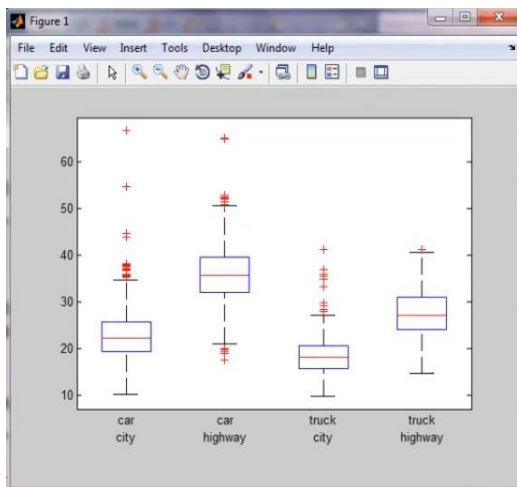


Рисунок 40 – График статистики для разных переменных

Применяется команда `boxplot`, которая строит график статистики для разных переменных. Она отражает для конкретной переменной значения квантилей: 25% и 75%, среднее значение и крайне значение, которое отображает значение 3σ .

1.4.2 Работа с распределениями

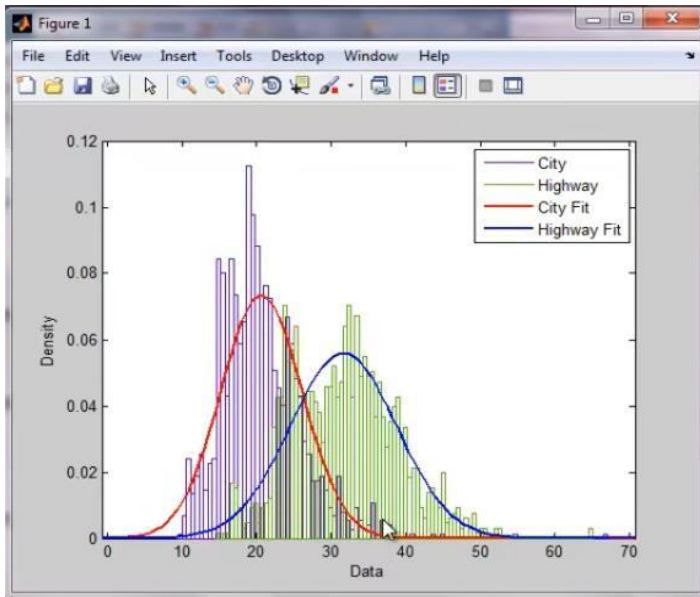


Рисунок 41 – График распределения

Различные возможности для работы с распределениями позволяют подобрать из Statistics and Machine Learning Toolbox нужные распределения под имеющиеся данные. Для вывода графика применяем команду `mpg Distribution`.

ГЛАВА 2. ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТА И ОБРАБОТКА РЕЗУЛЬТАТОВ

2.1 Проверка статистических гипотез

2.1.1 Сущность задачи проверки статистических гипотез

Проверка статистических гипотез складывается из следующих этапов [14]:

- формулируется в виде статистической гипотезы задача исследования;
- выбирается статистическая характеристика гипотезы;
- выбираются нулевая H_0 и альтернативная H_1 гипотезы на основе анализа возможных ошибочных решений и их последствий;
- выбирается приемлемый уровень значимости α ;
- выбирается критерий проверки гипотезы H_0 ;
- вычисляется фактическое значение статистического критерия;
- определяется критическое значение статистического критерия по соответствующей таблице;
- проверяется нулевая гипотеза на основе сравнения фактического и критического значений критерия, в зависимости от результатов проверки гипотеза либо отклоняется, либо не отклоняется.

При проверке гипотез по одному из критериев возможны два ошибочных решения:

- неправильное отклонение нулевой гипотезы – ошибка первого рода;

- неправильное принятие нулевой гипотезы – ошибка второго рода.

Статистическая гипотеза представляет собой некоторое предположение о законе распределения случайной величины или о параметрах этого закона, формулируемое на основе выборки.

Примерами статистических гипотез являются предположения: генеральная совокупность распределена по экспоненциальному закону; математические ожидания двух экспоненциально распределенных выборок равны друг другу.

В первой из них высказано предположение о виде закона распределения, а во второй – о параметрах двух распределений. Гипотезы, в основе которых нет никаких допущений о конкретном виде закона распределения, называют непараметрическими, в противном случае – параметрическими. Гипотезу, утверждающую, что различие между сравниваемыми характеристиками отсутствует, а наблюдаемые отклонения объясняются лишь случайными колебаниями в выборках, на основании которых производится сравнение, называют нулевой(основной) гипотезой и обозначают – H_0 .

Наряду с основной гипотезой рассматривают и альтернативную (конкурирующую, противоречащую) ей гипотезу H_1 . и если нулевая гипотеза будет отвергнута, то будет иметь место альтернативная гипотеза.

Различают простые и сложные гипотезы. Гипотезу называют простой, если она однозначно характеризует параметр распределения случайной величины. Например, если l является параметром экспоненциального распределения, то гипотеза H_0 о равенстве $l=10$ – простая гипотеза.

Сложной называют гипотезу, которая состоит из конечного или бесконечного множества простых гипотез. Сложная гипотеза H_0 о неравенстве $l > 10$ состоит из бесконечного множества простых гипотез H_0 о равенстве $l = b_i$, где b_i – любое число, большее 10. Гипотеза H_0 о том, что математическое ожидание нормального распределения равно двум при неизвестной дисперсии, тоже является сложной.

Сложной гипотезой будет предположение о распределении случайной величины X по нормальному закону, если не фиксируются конкретные значения математического ожидания и дисперсии.

Проверка гипотезы основывается на вычислении некоторой случайной величины – критерия, точное или приближенное распределение, которого известно. Обозначим эту величину через z , ее значение является функцией от элементов выборки $z = z(x_1, x_2, \dots, x_n)$.

Процедура проверки гипотезы предписывает каждому значению критерия одно из двух решений – принять или отвергнуть гипотезу. Тем самым все выборочное пространство и соответственно множество значений критерия делятся на два непересекающихся подмножества S_0 и S_1 . Если значение критерия z попадает в область S_0 , то гипотеза принимается, а если в область S_1 , то гипотеза отклоняется.

Множество S_0 называется областью принятия гипотезы или областью допустимых значений, а множество S_1 – областью отклонения гипотезы или критической областью. Выбор одной области однозначно определяет и другую область.

Принятие или отклонение гипотезы H_0 по случайной выборке соответствует истине с некоторой вероятностью и, соответственно, возможны два рода ошибок. Ошибка первого рода возникает с вероятностью α тогда, когда отвергается

верная гипотеза H_0 и принимается конкурирующая гипотеза H_1 .

Ошибка второго рода возникает с вероятностью b в том случае, когда принимается неверная гипотеза H_0 , в то время как справедлива конкурирующая гипотеза H_1 .

Доверительная вероятность – это вероятность не совершить ошибку первого рода и принять верную гипотезу H_0 .

Вероятность отвергнуть ложную гипотезу H_0 называется мощностью критерия. Следовательно, при проверке гипотезы возможны четыре варианта исходов, представленных в таблице 1:

Таблица 1 – Возможные варианты исхода по выдвигаемым гипотезам

Гипотеза H_0	Решение	Вероятность	Примечание
Верна	Принимается	$1 - \alpha$	Доверительная вероятность
	Отвергается	α	Вероятность ошибки 1-го рода
Неверна	Принимается	b	Вероятность ошибки 2-го рода
	Отвергается	$1 - b$	Мощность критерия

Например, рассмотрим случай, когда некоторая несмещенная оценка параметра θ вычислена по выборке объема n , и эта оценка имеет плотность распределения $f(\theta)$, рисунок 42.

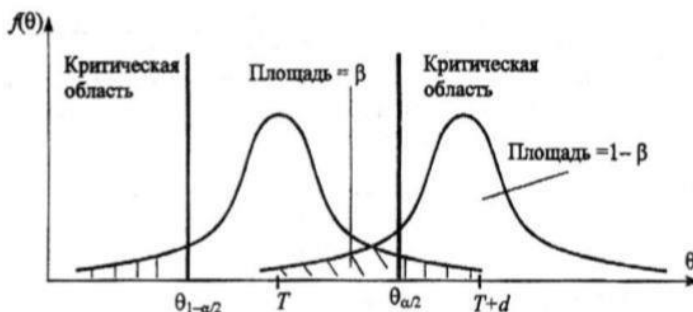


Рисунок 42 – Области принятия и отклонения гипотезы

Предположим, что истинное значение оцениваемого параметра равно T . Если рассматривать гипотезу H_0 о равенстве $\theta=T$, то насколько велико должно быть различие между θ и T , чтобы эту гипотезу отвергнуть. Ответить на данный вопрос можно в статистическом смысле, рассматривая вероятность достижения некоторой заданной разности между q и T на основе выборочного распределения параметра.

Целесообразно полагать одинаковыми значения вероятности выхода параметра θ за нижний и верхний пределы интервала. Такое допущение во многих случаях позволяет минимизировать доверительный интервал, т.е. повысить мощность критерия проверки. Суммарная вероятность выхода параметра θ за пределы интервала с границами $\theta_{1-a/2}$ и $\theta_{a/2}$, составляет величину a . Эту величину следует выбирать такой, чтобы выход за пределы интервала был маловероятен. Если оценка параметра попала в заданный интервал, то нет основания подвергать сомнению проверяемую гипотезу. Гипотезу равенства $\theta=T$ можно принять. Но если после получения выборки окажется, что оценка выходит за установленные пределы, то в этом случае есть серьезные основания отвергнуть гипотезу H_0 . Отсюда следует, что

вероятность допустить ошибку первого рода равна α (равна уровню значимости критерия).

Если предположить, например, что истинное значение параметра в действительности равно $T+d$, то согласно гипотезе H_0 о равенстве $\theta=T$ – вероятность того, что оценка параметра θ попадет в область принятия гипотезы, составит β , рисунок 42.

При заданном объеме выборки вероятность совершения ошибки первого рода можно уменьшить, снижая уровень значимости α . Однако при этом увеличивается вероятность ошибки второго рода β (снижается мощность критерия).

Аналогичные рассуждения можно провести для случая, когда истинное значение параметра равно $T-d$.

Единственный способ уменьшить обе вероятности состоит в увеличении объема выборки (плотность распределения оценки параметра при этом становится более "узкой"). При выборе критической области руководствуются правилом Неймана – Пирсона: следует так выбирать критическую область, чтобы вероятность α была мала, если гипотеза верна, и велика в противном случае. Однако выбор конкретного значения α относительно произволен. Употребительные значения лежат в пределах от 0,001 до 0,2.

В целях упрощения ручных расчетов составлены таблицы интервалов с границами $\theta_{1-\alpha/2}$ и $\theta_{\alpha/2}$ для типовых значений n и различных способов построения критерия. При выборе уровня значимости необходимо учитывать мощность критерия при альтернативной гипотезе. Иногда большая мощность критерия оказывается существеннее малого уровня значимости, и его значение выбирают относительно большим, например, 0,2. Такой выбор оправдан, если последствия ошибок второго рода более существенны, чем ошибок первого рода. Например, если отвергнуто правильное решение "продолжить работу пользователей с текущими паролями", то ошибка первого рода

приведет к некоторой задержке в нормальном функционировании системы, связанной со сменой паролей.

Если же принято решения не менять пароли, несмотря на опасность несанкционированного доступа посторонних лиц к информации, то эта ошибка повлечет более серьезные последствия.

В зависимости от сущности проверяемой гипотезы и используемых мер расхождения оценки характеристики от ее теоретического значения применяют различные критерии. К числу наиболее часто применяемых критериев для проверки гипотез о законах распределения относят критерии хи-квадрат Пирсона, Колмогорова, Мизеса, Вилкоксона, о значениях параметров – критерии Фишера, Стьюдента.

2.2 Типовые распределения

При проверке гипотез широкое применение находит ряд теоретических законов распределения [15]. Наиболее важным из них является нормальное распределение. С ним связаны распределения хи-квадрат, Стьюдента, Фишера, а также интеграл вероятностей. Для указанных законов функции распределения аналитически не представимы. Значения функций определяются по таблицам или с использованием стандартных процедур пакетов прикладных программ. Указанные таблицы обычно построены в целях удобства проверки статистических гипотез в ущерб теории распределений – они содержат не значения функций распределения, а критические значения аргумента $z(\alpha)$. Для односторонней критической области $z(\alpha)=z_{1-\alpha}$, т.е. критическое значение аргумента $z(\alpha)$ соответствует квантили $z_{1-\alpha}$ уровня $1-\alpha$, рисунок 43, так как:

$$\int_z^{\infty} f(z)dz = \alpha = 1 - \int_{-\infty}^{z(\alpha)} f(z)dz \quad (1)$$

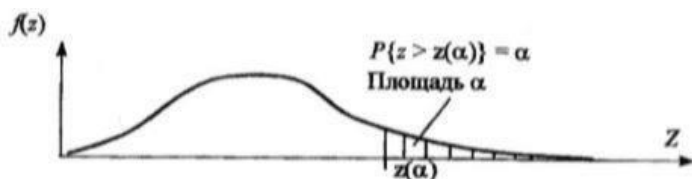


Рисунок 43 – Односторонняя критическая область

Для двусторонней критической области, с уровнем значимости α , размер левой области α_2 , правой α_1 ($\alpha_1 + \alpha_2 = \alpha$) рисунок 44. Значения $z(\alpha_2)$ и $z(\alpha_1)$ связаны с квантилями распределения соотношениями:

$$z(\alpha_1) = z_1 - \alpha_1, \quad z(\alpha_2) = z(\alpha_2) \quad (2)$$

так как:

$$\int_{-\infty}^{z(\alpha_1)} f(z)dz = 1 - \alpha_1 \quad (3)$$

$$\int_{-\infty}^{z(\alpha_2)} f(z)dz = 1 - \alpha_2$$

Для симметричной функции плотности распределения $f(z)$ критическую область выбирают из условия $\alpha_1 = \alpha_2 = \alpha/2$ (обеспечивается наибольшая мощность критерия).

В таком случае левая и правая границы будут равны $|z(\alpha/2)|$.

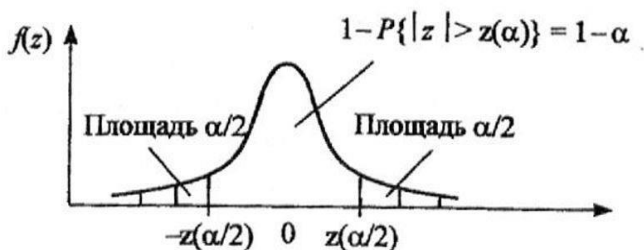


Рисунок 44 – Двусторонняя критическая область

2.2.1 Нормальное распределение

Этот вид распределения является наиболее важным в связи с центральной предельной теоремой теории вероятностей: распределение суммы независимых случайных величин стремится к нормальному с увеличением их количества при произвольном законе распределения отдельных слагаемых, если слагаемые обладают конечной дисперсией. Кроме того, А.М. Ляпунов доказал, что распределение параметра стремится к нормальному, если на параметр оказывает влияние большое количество факторов и ни один из них не является превалирующим. Функция плотности нормального распределения:

$$f(x) = \frac{1}{\sqrt{2\pi m_2}} \exp \left[-\frac{(x-m_1)^2}{2m_2} \right] \quad (4)$$

– унимодальная, симметричная, аргумент x может принимать любые действительные значения, рисунок 43, 45.

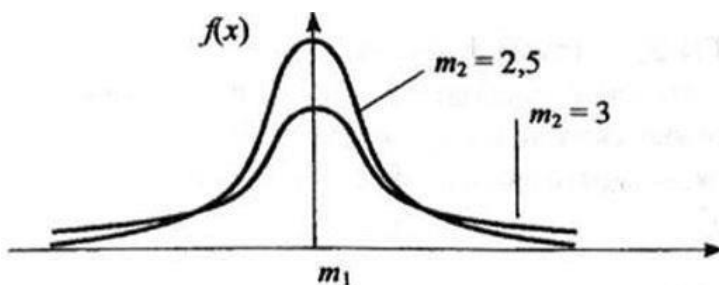


Рисунок 45 - Плотность нормального распределения

Функция плотности нормального распределения стандартизованной величины u имеет вид:

$$f(u) = \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{(u)^2}{2} \right] \quad (5)$$

Вычисление значений функции распределения $\Phi(u)$ для стандартизованного неотрицательного аргумента u ($u \geq 0$) можно произвести с помощью полинома наилучшего приближения:

$$\Phi(u) = 1 - 0,5 (1 + 0,196854u + 0,115194u^2 + 0,000344u^3 + 0,019527u^4) - 4 \quad (6)$$

Такая аппроксимация обеспечивает абсолютную ошибку не более 0,00025. Для вычисления $\Phi(u)$ в области отрицательных значений стандартизованного аргумента u ($u < 0$) следует воспользоваться свойством симметрии нормального распределения

$$\Phi(u) = 1 - \Phi(-u) \quad (7)$$

Интеграл вероятностей связан с функцией нормального распределения стандартизированной величины соотношением

$$\Phi(u) = 0,5 + F(u) \quad (8)$$

2.2.2 Распределение хи-квадрат

Распределению хи-квадрат (χ^2 – распределению) с k степенями свободы соответствует распределение суммы квадратов n стандартизованных случайных величин u_j :

$$x = \sum_{j=1}^n u_j^2 \quad (9)$$

каждая из которых распределена по нормальному закону, причем k из них независимы, $n \geq k$.

Функция плотности распределения хи-квадрат с k степенями свободы:

$$f(x) = \left[2^{k/2} \Gamma\left(\frac{k}{2}\right) \right]^{-1} (x)^{\frac{k}{2}-1} e^{-x/2}, x \geq 0 \quad (10)$$

где $\Gamma(k/2)$ – гамма-функция.

Функция плотности при k , равном одному или двум, – монотонная, а при $k > 2$ – унимодальная, несимметричная, рисунок 46.

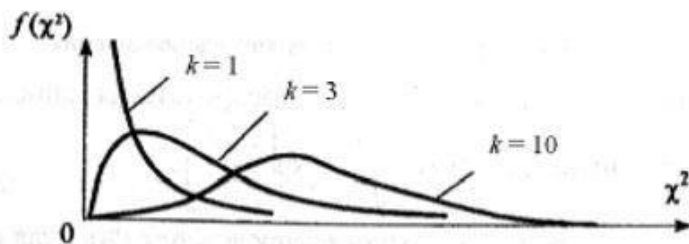


Рисунок 46 – Плотность распределения хи-квадрат

Математическое ожидание и дисперсия величины χ^2 равны соответственно k и $2k$. Распределение хи-квадрат является частным случаем более общего гаммараспределения, а величина, равная корню квадратному из хи-квадрат с двумя степенями свободы, подчиняется распределению Рэля.

С увеличением числа степеней свободы ($k > 30$) распределение хи-квадрат приближается к нормальному распределению с математическим ожиданием k и дисперсией $2k$.

2.2.3 Распределение Стьюдента

Распределение Стьюдента (t-распределение, предложено в 1908 г. английским статистиком В. Госсетом, публиковавшим научные труды под псевдонимом Student) характеризует распределение случайной величины:

$$t = \frac{u_0}{\sqrt{(u_1^2 + u_2^2 + \dots + u_k^2)/k}} \quad (11)$$

где u_0, u_1, \dots взаимно независимые нормально распределенные случайные величины с нулевым средним и

конечной дисперсией. Аргумент t не зависит от дисперсии слагаемых. Функция плотности распределения Стьюдента

$$f(t) = \frac{\Gamma[(k+1)/2]}{\sqrt{\pi k} \Gamma(k/2)} \left[1 + \frac{t^2}{k} \right]^{-\frac{(k+1)}{2}} \quad (12)$$

Величина k характеризует количество степеней свободы. Плотность распределения – унимодальная и симметричная функция, похожая на нормальное распределение, рисунок 47.

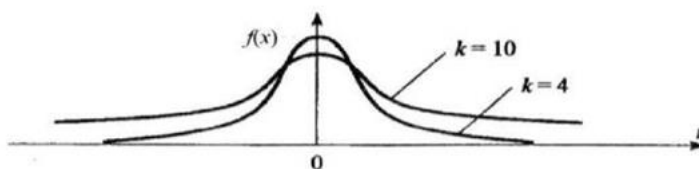


Рисунок 47 – Плотность распределения Стьюдента

Область изменения аргумента t от минус до плюс бесконечности. Математическое ожидание и дисперсия равны 0 и $k/(k-2)$ соответственно, при $k > 2$. По сравнению с нормальным распределение Стьюдента более пологое, оно имеет меньшую дисперсию. Это отличие заметно при небольших значениях k , что следует учитывать при проверке статистических гипотез (критические значения аргумента распределения Стьюдента превышают аналогичные показатели нормального распределения). Таблицы распределения содержат значения для односторонней (пределы интегрирования от $r(k; \alpha)$ до бесконечности:

$$\int_{r(k;a)}^{\infty} f(t)dt = a \quad (13)$$

или двусторонней (пределы интегрирования от $-r(k; \alpha)$ до $r(k; \alpha)$):

$$\int_{-r(k;a)}^{r(k;a)} f(t)dt = a \quad (14)$$

критической области.

Распределение Стьюдента применяется для описания ошибок выборки при $k < 30$. При k , превышающем 100, данное распределение практически соответствует нормальному, для значений k из диапазона от 30 до 100 различия между распределением Стьюдента и нормальным распределением составляют несколько процентов.

Поэтому относительно оценки ошибок малыми считаются выборки объемом не более 30 единиц, большими – объемом более 100 единиц. При аппроксимации распределения Стьюдента нормальным распределением для односторонней критической области вероятность:

$$P\{t > t(k; \alpha)\} = u_{1-\alpha}(0, k/(k-2)), \quad (15)$$

Где $u_{1-\alpha}(0, k/(k-2))$ – квантиль нормального распределения. Аналогичное соотношение можно составить и для двусторонней критической области:

$$f(x) = \left[\frac{k_1}{k_2} \right]^{k_1/2} \frac{\Gamma[(k_1 + k_2)/2]}{\Gamma(k_1/2)\Gamma(k_2/2)} x^{(k_1-2)/2} \left[1 + \frac{k_1}{k_2} x \right]^{-\frac{(k_1+k_2)}{2}} \quad (16)$$

2.2.4 Распределение Фишера

Распределению Р.А. Фишера (F-распределению Фишера –Снедекора) подчиняется случайная величина:

$$X = [(y_1/k_1) / (y_2/k_2)], \quad (17)$$

равная отношению двух случайных величине y_1 и y_2 , имеющих хи-квадрат распределение с k_1 и k_2 степенями свободы. Область изменения аргумента x от 0 до бесконечности. Плотность распределения

В этом выражении k_1 обозначает число степеней свободы величины y_1 с большей дисперсией, k_2 – число степеней свободы величины y_2 с меньшей дисперсией. Плотность распределения – унимодальная, несимметричная, рисунок 48:

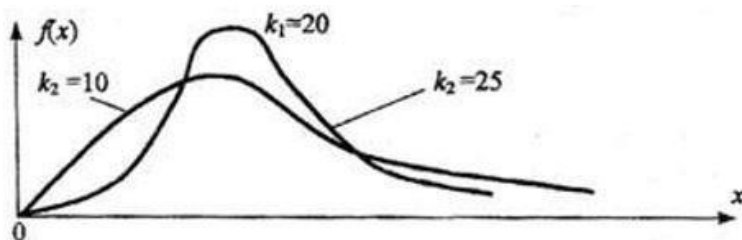


Рисунок 48 – Плотность распределения Фишера

Математическое ожидание случайной величины x

$$m_1 = k_2 / (k_2 - 2) \text{ при } k_2 > 2, \quad (18)$$

дисперсия $m_2 [2 k_2^2 (k_1 + k_2 - 2)] / [k_1 (k_2 - 2)^2 (k_2 - 4)]$ при $k_2 > 4$.

При $k_1 > 30$ и $k_2 > 30$ величина x распределена приблизительно нормально: с центром распределения $(k_1 - k_2)/(2k_1k_2)$ и дисперсией $(k_1 + k_2)/(2k_1k_2)$.

2.3 Проверка гипотез о законе распределения

Обычно сущность проверки гипотезы о законе распределения ЭД заключается в следующем: Имеется выборка ЭД фиксированного объема, выбран или известен вид закона распределения генеральной совокупности. Необходимо оценить по этой выборке параметры закона, определить степень согласованности ЭД и выбранного закона распределения, в котором параметры заменены их оценками. Пока не будем касаться способов нахождения оценок параметров распределения, а рассмотрим только вопрос проверки согласованности распределений с использованием наиболее употребительных критериев.

2.3.1 Критерий хи-квадрат К. Пирсона

Использование этого критерия основано на применении такой меры (статистики) расхождения между теоретическим $F(x)$ и эмпирическим распределением $F_n(x)$, которая приблизительно подчиняется закону распределения χ^2 . Гипотеза H_0 о согласованности распределений проверяется путем анализа распределения этой статистики. Применение критерия требует построения статистического ряда.

Итак, пусть выборка представлена статистическим рядом с количеством разрядов u . Наблюдаемая частота попаданий в i -й разряд n_i . В соответствии с теоретическим законом распределения ожидаемая частота попаданий в i -й разряд составляет F_i . Разность между наблюдаемой и ожидаемой

частотой составит величину $(n_i - F_i)$. Для нахождения общей степени расхождения между $F(x)$ и $F_n(x)$ необходимо подсчитать взвешенную сумму квадратов разностей по всем разрядам статистического ряда:

$$\chi^2 = \sum_{i=1}^n \frac{(n_i - F_i)^2}{F_i} \quad (19)$$

Величина χ^2 при неограниченном увеличении n имеет распределение хи-квадрат (асимптотически распределена как хи-квадрат). Это распределение зависит от числа степеней свободы k , т.е. количества независимых значений слагаемых в выражении. Число степеней свободы равно числу y минус число линейных связей, наложенных на выборку. Одна связь существует в силу того, что любая частота может быть вычислена по совокупности частот в оставшихся $y-1$ разрядах. Кроме того, если параметры распределения неизвестны заранее, то имеется еще одно ограничение, обусловленное подгонкой распределения к выборке. Если по выборке определяются f параметров распределения, то число степеней свободы составит $k = y - f - 1$.

Очевидно, что чем меньше расхождение между теоретическими и эмпирическими частотами, тем меньше величина критерия. Область принятия гипотезы H_0 определяется условием $\chi^2 < \chi^2(k, \alpha)$ – критическая точка распределения хи-квадрат с уровнем значимости α .

Вероятность ошибки первого рода равна α , вероятность ошибки второго рода четко определить нельзя, потому что существует бесконечно большое множество различных способов несовпадения распределений. Мощность критерия зависит от количества разрядов и объема выборки.

Критерий рекомендуется применять при $n > 200$, допускается применение при $n > 40$, именно при таких условиях критерий состоятелен (как правило, отвергает неверную нулевую гипотезу). Теоретическое значение вероятности F_i попадания случайной величины в i -й интервал определяется по формуле:

$$F_i = \frac{1}{\sigma\sqrt{2\pi}} \int_{x_{i-1}}^{x_i} \exp\left(-\frac{(x-\mu_1)^2}{2\sigma^2}\right) dx. \quad (20)$$

Для нормального закона возможные значения случайной величины лежат в диапазоне от минус до плюс бесконечности, поэтому при расчетах оценок вероятностей крайний левый и крайний правый интервалы расширяются до минус и плюс бесконечности соответственно. Вычислить значения функции нормального распределения можно, воспользовавшись стандартными функциями табличного процессора или полиномом наилучшего приближения.

2.3.2 Критерий А.Н. Колмогорова

Для применения критерия А.Н. Колмогорова ЭД требуется представить в виде вариационного ряда (ЭД недопустимо объединять в разряды). В качестве меры расхождения между теоретической $F(x)$ и эмпирической $F_n(x)$ функциями распределения непрерывной случайной величины X используется модуль максимальной разности:

$$d_n = \max |F(x) - F_n(x)| \quad (21)$$

А.Н. Колмогоров доказал, что какова бы ни была функция распределения $F(x)$ величины X при неограниченном

увеличении количества наблюдений n функция распределения случайной величины $d_n\sqrt{n}$ асимптотически приближается к функции распределения

$$K(\lambda) = P(d\sqrt{n} < \lambda) = \sum_{k=-\infty}^{\infty} (-1)^k \exp(-2k^2\lambda^2). \quad (22)$$

Иначе говоря, критерий А.Н. Колмогорова характеризует вероятность того, что величина $d_n\sqrt{n}$ не будет превосходить параметр 1 для любой теоретической функции распределения. Уровень значимости α выбирается из условия:

$$P(d_n\sqrt{n} > \lambda) = \alpha = 1 - K(\lambda), \quad (23)$$

В силу предположения, что почти невозможно получить это равенство, когда существует соответствие между функциями $F(x)$ и $F_n(x)$. Критерий А.Н. Колмогорова позволяет проверить согласованность распределений по малым выборкам, он проще критерия хи-квадрат, поэтому его часто применяют на практике. Но требуется учитывать два обстоятельства:

В соответствии с условиями его применения необходимо пользоваться следующим соотношением:

$$\begin{aligned} d_n &= \max(d_n^+, d_n^-), \\ d_n^+ &= \max_{1 \leq j \leq n} \left| \frac{j}{n} - F(x) \right|; \\ d_n^- &= \max_{1 \leq j \leq n} \left| F(x) - \frac{j-1}{n} \right| \end{aligned} \quad (24)$$

Условия применения критерия предусматривают, что теоретическая функция распределения известна полностью –

известны вид функции и значения ее параметров. На практике параметры обычно неизвестны и оцениваются по ЭД. Но критерий не учитывает уменьшение числа степеней свободы при оценке параметров распределения по исходной выборке. Это приводит к завышению значения вероятности соблюдения нулевой гипотезы, т.е. повышается риск принять в качестве правдоподобной гипотезу, которая плохо согласуется с ЭД (повышается вероятность совершить ошибку второго рода). В качестве меры противодействия такому выводу следует увеличить уровень значимости α , приняв его равным 0,1 – 0,2, что приведет к уменьшению зоны допустимых отклонений.

2.3.3 Критерий Р. Мизеса

В качестве меры различия теоретической функции распределения $F(x)$ и эмпирической $F_n(x)$ по критерию Мизеса (критерию – ω^2) выступает средний квадрат отклонений по всем значениям аргумента x :

$$\omega_n^2 = \int_{-\infty}^{+\infty} [F_n(x) - F(x)]^2 dF(x) \quad (25)$$

Статистика критерия:

$$n\omega_n^2 = \frac{1}{12n} + \sum_{i=1}^n \left[F(x_i) - \frac{i-0,5}{n} \right]^2 \quad (26)$$

При неограниченном увеличении n существует предельное распределение статистики $n\omega_n^2$. Задав значение вероятности α можно определить критические значения $n\omega_n^2(\alpha)$. Проверка гипотезы о законе распределения осуществляется обычным образом: если фактическое значение

$n\omega_n^2$ окажется больше критического или равно ему, то согласно критерию Мизеса с уровнем значимости α гипотеза H_0 о том, что закон распределения генеральной совокупности соответствует $F(x)$, должна быть отвергнута.

Достоинством критерия Мизеса является быстрая сходимость к предельному закону, для этого достаточно не менее 40 наблюдений в области часто используемых на практике больших значений $n\omega_n^2$ (а не несколько сот, как для критерия хи-квадрат).

Сопоставляя возможности различных критериев, необходимо отметить следующие особенности:

Критерий Пирсона устойчив к отдельным случайным ошибкам в ЭД. Однако его применение требует группирования данных по интервалам, выбор которых относительно произволен и подвержен противоречивым рекомендациям.

Критерий Колмогорова слабо чувствителен к виду закона распределения и подвержен влиянию помех в исходной выборке, но прост в применении.

Критерий Мизеса имеет ряд общих свойств с критерием Колмогорова: оба основаны непосредственно на результатах наблюдения и не требуют построения статистического ряда, что повышает объективность выводов; оба не учитывают уменьшение числа степеней свободы при определении параметров распределения по выборке, а это ведет к риску принятия ошибочной гипотезы. Их предпочтительно применять в тех случаях, когда параметры закона распределения известны априори, например, при проверке датчиков случайных чисел.

При проверке гипотез о законе распределения следует помнить, что слишком хорошее совпадение с выбранным законом распределения может быть обусловлено некачественным экспериментом («подчистка» ЭД) или

предвзятой предварительной обработкой результатов (некоторые результаты отбрасываются или округляются).

Выбор критерия проверки гипотезы относительно произволен. Разные критерии могут давать различные выводы о справедливости гипотезы, окончательное заключение в таком случае принимается на основе неформальных соображений. Точно также нет однозначных рекомендаций по выбору уровня значимости.

Рассмотренный подход к проверке гипотез, основанный на применении специальных таблиц критических точек распределения, сложился в эпоху "ручной" обработки ЭД, когда наличие таких таблиц существенно снижало трудоемкость вычислений. В настоящее время математические пакеты включают процедуры вычисления стандартных функций распределений, что позволяет отказаться от использования таблиц, но может потребовать изменения правил проверки.

2.4 Методы оценки параметров распределения

2.4.1 Точечная оценка параметров распределения

Точечная оценка предполагает нахождение единственной числовой величины, которая и принимается за значение параметра. Такую оценку целесообразно определять в тех случаях, когда объем ЭД достаточно велик. Причем не существует единого понятия о достаточном объеме ЭД, его значение зависит от вида оцениваемого параметра (к этому вопросу вернёмся при изучении методов интервальной оценки параметров, а предварительно будем считать достаточной выборку, содержащую не менее чем 10 значений). При малом объеме ЭД точечные оценки могут значительно отличаться от

истинных значений параметров, что делает их непригодными для использования.

Задача точечной оценки параметров в типовом варианте постановки состоит в следующем:

Имеется выборка наблюдений (x_1, x_2, \dots, x_n) за случайной величиной X . Выборка должна быть представительной. Объем выборки n фиксирован. Известен вид закона распределения величины X , например, в форме плотности распределения $f(T, x)$, где T – неизвестный (в общем случае векторный) параметр распределения. Параметр является неслучайной величиной.

Требуется найти оценку θ параметра T закона распределения.

Существует несколько методов решения задачи точечной оценки параметров, самыми распространенными из которых являются методы максимального (наибольшего) правдоподобия, моментов и квантилей.

2.4.2 Метод максимального правдоподобия

Метод был предложен Р. Фишером в 1912 г. Метод основан на исследовании вероятности получения выборки наблюдений (x_1, x_2, \dots, x_n) за некоторым количественным признаком X . Значения выборки случайные, поэтому будем рассматривать их как независимые случайные величины X_1, X_2, \dots, X_n , имеющие одинаковые распределения. Если количественный признак X непрерывный, то случайные величины X_1, X_2, \dots, X_n имеют одинаковые плотности распределения вероятностей, зависящих от x и от вектора параметров распределения $\bar{\theta} = (\theta_1, \theta_2, \dots, \theta_n)$, т.е. плотности $f(x, \bar{\theta})$.

Суть метода максимального правдоподобия заключается в свойстве случайной величины реализовывать в эксперименте

в основном те свои значения (x_1, x_2, \dots, x_n) , вероятность которых максимальна. В этом случае совместная плотность распределения:

- вероятностей $f(x_1, x_2, \dots, x_n, \bar{\theta})$ случайных величин X_1, X_2, \dots, X_n должна быть максимальной для непрерывного признака или совместное распределение вероятностей
- $P(X_1=x_1, X_2=x_2, \dots, X_n=x_n)$ случайных величин X_1, X_2, \dots, X_n должно быть максимальным для дискретного признака.

Ввиду независимости случайных величин X_1, X_2, \dots, X_n их совместная плотность распределения будет равна произведению плотностей распределения, т.е.:

$$f(x_1, x_2, \dots, x_n, \bar{\theta}) = f(x_1, \bar{\theta}) \cdot f(x_2, \bar{\theta}) \cdot \dots \cdot f(x_n, \bar{\theta}), \quad (27)$$

а совместная вероятность равна произведению вероятностей событий:

$$P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n, \bar{\theta}) = p_{x_1}(\bar{\theta}) \cdot \dots \cdot p_{x_n}(\bar{\theta}). \quad (28)$$

Функцией правдоподобия $L(x_1, x_2, \dots, x_n, \bar{\theta})$ называется функция, которая в случае непрерывного количественного признака определяется по формулам: для непрерывного признака:

$$L(x_1, x_2, \dots, x_n, \bar{\theta}) = f(x_1, \bar{\theta}) \cdot f(x_2, \bar{\theta}) \cdot \dots \cdot f(x_n, \bar{\theta}). \quad (29)$$

для дискретного признака

$$L(x_1, x_2, \dots, x_n, \bar{\theta}) = p_{x_1}(\bar{\theta}) \cdot \dots \cdot p_{x_n}(\bar{\theta}). \quad (30)$$

В качестве оценки $\bar{\theta}_1, \dots, \bar{\theta}_n$ неизвестных параметров распределения $\theta_1, \dots, \theta_n$ берут те значения, при которых функция правдоподобия достигает максимума.

Технически задача поиска максимального значения функции правдоподобия облегчается, если рассмотреть не саму функцию, а натуральный логарифм от неё, т.е. функцию:

$$\ln L(x_1, x_2, \dots, x_n, \bar{\theta}) \quad (31)$$

За оценки неизвестного вектора параметров q берется решение $\bar{\theta}_1, \dots, \bar{\theta}_n$ уравнения правдоподобия:

$$\frac{\partial \ln L(x_1, x_2, \dots, x_n, \bar{\theta})}{\partial \theta_k} = 0; k = \overline{1, n} \quad (32)$$

Для проверки того, что точка оптимума соответствует максимуму функции правдоподобия, необходимо найти вторую производную от этой функции. И если вторая производная в точке оптимума отрицательна, то найденные значения параметров максимизируют функцию.

Итак, нахождение оценок максимального правдоподобия включает следующие этапы: построение функции правдоподобия (ее натурального логарифма); дифференцирование функции по искомым параметрам и составление системы уравнений; решение системы уравнений для нахождения оценок; определение второй производной

функции, проверку ее знака в точке оптимума первой производной и формирование выводов.

Метод максимального правдоподобия позволяет получить состоятельные, эффективные (если таковые существуют, то полученное решение даст эффективные оценки), достаточные, асимптотически нормально распределенные оценки. Этот метод может давать как смещенные, так и несмещенные оценки. Смещение удается устранить введением поправок. Метод особенно полезен при малых выборках. Если функция максимального правдоподобия имеет несколько максимумов, то из них выбирают глобальный.

2.4.3 Метод моментов

Метод предложен К. Пирсоном в 1894 г.

Будем считать, что вид функции распределения изучаемого количественного признака известен, но параметры этого распределения неизвестны. Нужно оценить эти параметры по выборке. Сущность метода моментов заключается в том, что по конкретной выборке x_1, x_2, \dots, x_n генеральной совокупности X , распределение которой известно с точностью до параметров $\theta_1, \dots, \theta_m$ выбирается столько эмпирических моментов, сколько требуется оценить неизвестных параметров распределения. Желательно применять моменты младших порядков, так как погрешности вычисления оценок резко возрастают с увеличением порядка момента.

Вычисленные по ЭД оценки моментов приравниваются к теоретическим моментам; параметры распределения определяются через моменты, и составляются уравнения, выражающие зависимость параметров от моментов, в результате получается система уравнений. Решение этой

системы дает оценки параметров распределения генеральной совокупности.

2.4.4 Интервальная оценка параметров распределения

Интервальный метод оценивания параметров распределения случайных величин заключается в определении интервала (а не единичного значения), в котором с заданной степенью достоверности будет заключено значение оцениваемого параметра.

Интервальная оценка характеризуется двумя числами – концами интервала, внутри которого предположительно находится истинное значение параметра. Иначе говоря, вместо отдельной точки для оцениваемого параметра можно установить интервал значений, одна из точек которого является своего рода "лучшей" оценкой. Интервальные оценки являются более полными и надежными по сравнению с точечными, они применяются как для больших, так и для малых выборок. Совокупность методов определения промежутка, в котором лежит значение параметра T , получила название методов интервального оценивания. К их числу принадлежит метод Неймана.

Постановка задачи интервальной оценки параметров заключается в следующем: имеется выборка наблюдений x_1, x_2, \dots, x_n , за случайной величиной X . Объем выборки n фиксирован. Предположим, что статистическая характеристика $\bar{\theta}$, рассчитанная по данным выборки, является выборочной оценкой неизвестного параметра θ генеральной совокупности, причём θ – это постоянное число. Оценка $\bar{\theta}$ характеризует параметр θ тем точнее, чем меньше абсолютная величина разности $|\theta - \bar{\theta}|$.

Если $|\theta - \bar{\theta}| < \delta$, где $\delta > 0$, то число δ называется точностью оценки $\bar{\theta}$.

Нельзя утверждать, что оценка абсолютно $\bar{\theta}$ удовлетворяет неравенству $|\theta - \bar{\theta}| < \delta$. Однако можно задать вероятность γ , с которой это неравенство осуществляется.

Надёжность или доверительная вероятность оценки $\bar{\theta}$ – это вероятность γ , с которой осуществляется неравенство:

$$|\theta - \bar{\theta}| < \delta \quad (33)$$

Принято надёжность оценки задавать перед процессом оценивания параметра генеральной совокупности. В качестве вероятности γ берут число, близкое к 1 (от 0,95 до 0,999).

Пусть $P = (|\theta - \bar{\theta}| < \delta) = \gamma$. Если неравенство $|\theta - \bar{\theta}| < \delta$ заменить равносильным ему двойным неравенством:

$$-\delta < \theta - \bar{\theta} < \delta \text{ или } \bar{\theta} - \delta < \theta < \bar{\theta} + \delta \quad (34)$$

то получим:

$$P(\bar{\theta} - \delta < \bar{\theta} + \delta) = \gamma \quad (35)$$

Вероятностный смысл данного отношения таков: вероятность того, что интервал $(\bar{\theta} - \delta < \bar{\theta} + \delta)$ покрывает неизвестный параметр $\bar{\theta}$, равна γ .

Интервал, который покрывает оцениваемый неизвестный параметр с заданной вероятностью γ называется доверительным интервалом.

Границы $(\bar{\theta} - \delta)$ и $(\bar{\theta} + \delta)$ называются доверительными границами интервала. Они определяются на основе выборочных данных и являются функциями от случайных

величин X_1, X_2, \dots, X_n , и, следовательно, сами являются случайными величинами.

2.4.5 Общий метод построения доверительных интервалов

Доверительный интервал для оценки математического ожидания нормального распределения при известном среднем квадратичном отклонении.

Будем считать, что количественный признак X генеральной совокупности подчиняется нормальному закону распределения с параметрами α и σ . Среднее квадратичное данного распределения σ считается известным. Нужно оценить неизвестное математическое ожидание α генеральной совокупности по выборочной средней \bar{X} , т.е. необходимо определить доверительный интервал, который покрывает параметр α с заданной надёжностью γ .

Имеет место следующее утверждение: если случайная величина X подчиняется нормальному закону распределения, то выборочная средняя \bar{X} , найденная по независимым наблюдениям, также подчиняется нормальному закону распределения с параметрами:

$$M(\bar{X}) = \alpha; \quad \sigma(\bar{X}) = \frac{\sigma}{\sqrt{n}}. \quad (36)$$

Для того, чтобы определить доверительный интервал, необходимо, чтобы выполнялось равенство $P(|\bar{X} - \alpha| < \delta) = \gamma$.

Для нормально распределённой случайной величины X имеет место равенство:

$$P(|\bar{X} - \alpha| < \delta) = 2\Phi\left(\frac{\delta}{\sigma(\bar{X})}\right). \quad (37)$$

\bar{X} имеет нормальное распределение $\sigma(\bar{X}) = \frac{\sigma}{\sqrt{n}}$. Поэтому:

$$P(|\bar{X} - \alpha| < \delta) = 2\Phi\left(\frac{\delta\sqrt{n}}{\sigma}\right) = 2\Phi(t); \quad t = \frac{\delta\sqrt{n}}{\sigma} \quad (38)$$

Из равенства $t = \frac{\delta\sqrt{n}}{\sigma}$ выразим $\delta = \frac{\sigma}{\sqrt{n}}$ и получим:

$$P\left(|\bar{X} - \alpha| < t \frac{\sigma}{\sqrt{n}}\right) = 2\Phi(t). \quad (39)$$

Так как вероятность P задана и равна γ , окончательно получаем:

$$P\left(\bar{x} - t \frac{\sigma}{\sqrt{n}} < \alpha < \bar{x} + t \frac{\sigma}{\sqrt{n}}\right) = 2\Phi(t) = \gamma. \quad (40)$$

Таким образом, с доверительной вероятностью γ можно утверждать, что доверительный интервал $(\bar{x} - t \frac{\sigma}{\sqrt{n}} < \alpha < \bar{x} + t \frac{\sigma}{\sqrt{n}})$ покрывает неизвестный параметр α с точностью $\delta = t \frac{\sigma}{\sqrt{n}}$

\bar{x} – находим по данным выборки, n – это объём выборки, σ – величина, известная заранее.

Число t находится из равенства $2\Phi(t) = \gamma$, или $\Phi(t) = \frac{\gamma}{2}$. По таблице функции Лапласа определяется для заданной доверительной вероятности γ величина t и затем находится величина δ .

2.4.6 Методика Вальда проверки гипотезы о свойствах случайной величины

Рассматривается следующая задача: имеются экспериментальные данные о некоторой случайной величине ξ ; известно аналитическое выражение для плотности распределения этой случайной величины – некоторое

выражение $f(x, \theta)$, где x - аргумент, а θ – некоторый параметр, принимающий одно из двух возможных значений $\theta = \theta_1$ или $\theta = \theta_2$, причем $\theta_1 < \theta_2$. Какое именно из этих двух значений принимает параметр θ – неизвестно. Именно это и надо определить по значениям случайной величины ξ , которые можно разыгрывать неограниченно много.

Примером такой плотности может служить плотность нормального распределения:

$$f(x, \theta) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\theta)^2}{2\sigma^2}}, \quad (41)$$

где, кроме параметра θ , участвует так же еще один параметр σ .

Опишем методику Вальда решения поставленной задачи.

Фиксируем достаточно малые числа $\alpha, \beta \in (0;1)$, и далее фиксируем числа:

$$A = \frac{1-\beta}{\alpha}, \quad B = \frac{\beta}{1-\alpha}; \quad (42)$$

при достаточно малых α, β между числами A, B имеет место неравенство: $A > B$ именно для этого неравенства должны быть достаточно малы числа α, β . Пусть, далее $\xi_1, \xi_2, \dots, \xi_k$ - значения случайной величины ξ , получаемые экспериментально. Построим суммы:

$$\Lambda_t = \sum_{i=1}^t \ln \frac{f(\xi_i, \theta_2)}{f(\xi_i, \theta_1)}. \quad (43)$$

Основная теорема Вальда: если $\Lambda_t > \ln A$, то с вероятностью ошибки β ; если же $\Lambda_t > \ln B$, то $\theta = \theta_I$ с вероятностью ошибки α .

Принято рассматривать в описанной ситуации две величины - n и Δg_t , то минимальное значение t , при котором выполняется одно из неравенств:

$$\Lambda_t > \ln A \text{ или } \Lambda_t > \ln B \quad (44)$$

является случайной величиной, зависящей от той или иной реализации случайной величины $\xi_1, \xi_2, \dots, \xi_k \dots$; именно это минимальное значение t и является случайной величиной n .

Учитывая смысл только что введенного обозначения n , введем величину:

$$\Delta g_t = \frac{1}{n} \sum_{i=1}^n \ln \frac{f(\xi_i, \theta_2)}{f(\xi_i, \theta_1)}; \quad (45)$$

очевидно, и эта величина - также случайная. На практике часто имеют смысл математические ожидания последних двух случайных величин; их находят как обычные средние арифметические значения реализаций этих величин, конструируемых по различным реализациям $\xi_1, \xi_2, \dots, \xi_k \dots$

ГЛАВА 3. ЛАБОРАТОРНЫЙ ПРАКТИКУМ

Лабораторная работа 1

Определение уравнения линейной регрессии по экспериментальным данным однофакторного эксперимента

Теоретическая часть

Регрессионный анализ представляет один из разделов математической статистики, целью которого является функциональная аппроксимация случайных данных по заданной выборке. Регрессионный анализ использует следующий математический аппарат: методы аппроксимации, методы дизайна эксперимента и статистические оценки.

Пусть в результате измерений некоторой физической величины получены её значения Y в функции от некоторого фактора X , определяющие зависимость $Y=f(X)$ (Таблица 1).

Таблица 1 – Значения физической величины Y от некоторого фактора X

X	x_1	x_2	\dots	x_n
Y	y_1	y_2	\dots	y_n

Требуется найти формулу, выражающую данную зависимость аналитически. Один из подходов состоит в построении интерполяционного многочлена, значения которого в узлах интерполяции x_i совпадают со значениями данной функции $f(x_i) = f_i$. Если значения функции $f(x)$ известны с некоторой погрешностью, то требование

совпадения значений в узлах интерполяции не оправдано, поскольку оно не означает совпадение характеров исходной и интерполирующей функции. Поэтому поставим задачу следующим образом – найти функцию вида $y = f(x)$ которая в точках x_1, x_2, \dots, x_n принимает значения, близкие к табличным значениям y_1, y_2, \dots, y_n . Будем считать, что величина x – детерминированная, а получаемые значения y – случайные. В выражении искомой функциональной зависимости должны присутствовать числовые параметры, которые подлежат определению из опытных данных:

$$F(x) = F(x, a_1, a_2, \dots, a_m). \quad (1)$$

Такая функция будет m -параметрической. Один из подходов в построении такой зависимости заключается в построение теоретической кривой таким образом, чтобы вероятность появления экспериментальных данных, была максимальной. Преобразуем, заменив случайные параметры a_1, a_2, \dots, a_m их математическими ожиданиями A_1, A_2, \dots, A_m , представляющими неизвестные детерминированные величины, а все случайные составляющие выделим в отдельное слагаемое:

$$F(x) = F(x, A_1, A_2, \dots, A_m) + Z.$$

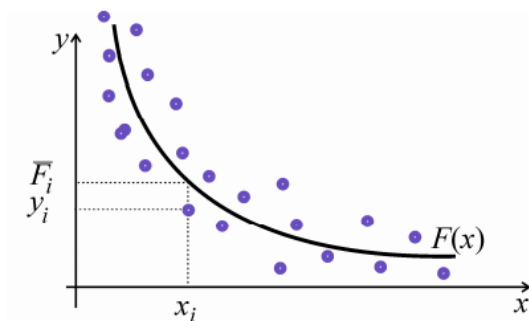


Рисунок 1 – Теоретическая кривая и экспериментальные точки

Предположим, что все опыты независимы, при измерениях нет систематических ошибок и промахов, ошибка измерений распределена по нормальному закону и все опыты равнозначные. В этом случае случайная величина Z имеет нормальное распределение и нулевое математическое ожидание. Пусть D_z – дисперсия величины Z , σ_z – среднеквадратичное отклонение. Тогда для каждого i -го опыта можно записать:

$$z_i = y_i - F(x, a_1, a_2, \dots, a_m).$$

В соответствии с выборочным методом будем считать каждую выборку единственной реализацией случайной величины z_i , которая имеет такое же распределение, как и Z . Плотность распределения каждой из z_i :

$$f(z_i) = \frac{1}{\sigma_z \sqrt{2\pi}} \exp\left(-\frac{z_i^2}{2\sigma_z^2}\right),$$

а совместная плотность распределения независимых z_i :

$$\begin{aligned} f(z_1, z_2, \dots, z_n) &= \frac{1}{\sigma_z^n (2\pi)^{\frac{n}{2}}} \prod_{i=1}^n \exp\left(-\frac{z_i^2}{2\sigma_z^2}\right) \\ &= \frac{1}{\sigma_z^n (2\pi)^{\frac{n}{2}}} \exp\left(-\frac{\sum_{i=1}^n z_i^2}{2\sigma_z^2}\right). \end{aligned}$$

Совместная плотность распределения должна быть максимально возможной. Максимизируя за счет подбора параметров a_1, a_2, \dots, a_m можно заключить, что первый множитель от них не зависит, знаменатель экспоненты также не зависит, учитывая знак перед числителем в экспоненте, можно рассчитать, что максимальная совместная плотность достигается при выполнении условия:

$$\Phi(a_1, a_2, \dots, a_m) = \sum_{i=1}^n z_i^2 = \sum_{i=1}^n (y_i - F(x, a_1, a_2, \dots, a_m))^2 \rightarrow \min. \quad (2)$$

где a_1, a_2, \dots, a_m являются случайными параметрами.

Таким образом, параметры теоретической кривой $F(x)$ выбираются так, чтобы сумма квадратов отклонений экспериментальных ординат от теоретических была наименьшей. Представление случайной величины F в виде (2) означает, что математическое ожидание каждого измерения равно:

$$M(y_i) = F(x_i, A_1, A_2, \dots, A_m), \quad (3)$$

где A_1, A_2, \dots, A_m математические ожидания.

В случае одного входного фактора X провести аппроксимацию зависимости (1) во многих случаях можно произвести методом наименьших квадратов. Этот метод можно применить для двух модельных зависимостей:

- двухпараметрической функции

$$F_1(x, a, b) = a \cdot \exp(b \cdot x); \quad (4)$$

- линейной зависимости

$$F_2(x, a, b) = a \cdot x + b. \quad (5)$$

Простейший способ найти параметры a, b , минимизирующие выражение (2) — вычислить его частные производные по каждому из этих параметров, приравнять их к нулю, после чего решить систему линейных уравнений:

$$\frac{\delta \Phi}{\delta a} = -2 \sum_{i=1}^n (y_i - a - bx_i) = 0;$$

$$\frac{\delta \Phi}{\delta b} = -2 \sum_{i=1}^n (y_i - a - b x_i) x_i = 0.$$

Значения параметров a и b , минимизирующие функцию потерь, удовлетворяют уравнениям

$$a = \frac{\sum y_i}{n} - b \frac{\sum x_i}{n}; \quad (6)$$

$$b = \frac{\frac{\sum x_i y_i}{n} - \frac{\sum x_i \sum y_i}{n^2}}{\frac{\sum x_i^2}{n} - \left(\frac{\sum x_i}{n}\right)^2}.$$

Уравнение (5) называют уравнением линейной регрессии, а параметры a и b – коэффициентами линейной регрессии.

Существует возможность с помощью подходящего преобразования переменных получить линейную зависимость и использовать уравнения (6) для функций, представленных в таблице 2.

Таблица 2 – Аппроксимирующий функции, нахождение коэффициентов которых сводится к нахождению коэффициентов линейной регрессии

Функция $y=f(x)$	Линеаризованная форма $y = ax+b$	Замена переменных и постоянных
$y = \frac{a}{x} + b$	$y = a \cdot \frac{1}{x} + b$	$X = \frac{1}{x}, Y = y$
$y = \frac{d}{x+c}$	$y = -\frac{1}{c} \cdot xy + \frac{d}{c}$	$X = xy, Y = y, c = -\frac{1}{a}, d = -\frac{b}{a},$
$y = \frac{1}{ax+b}$	$\frac{1}{y} = ax + b$	$X = x, Y = \frac{1}{y}$
$y = \frac{x}{ax+b}$	$\frac{1}{y} = a \frac{1}{x} + b$	$X = \frac{1}{x}, Y = \frac{1}{y}$
$y = a \cdot \ln(x) + b$	$y = a \cdot \ln(x) + b$	$X = \ln(x), Y = y$

$y = c \cdot e^{ax}$	$\ln(y)$ $= ax + \ln(c)$	$X = x, Y = \ln(y), c = e^b$
$y = c \cdot x^a$	$\ln(y)$ $= a \cdot \ln(x)$ $+ \ln(c)$	$X = \ln(x), Y = \ln(y), c = e^b$
$y = \frac{1}{(ax + b)^2}$	$\frac{1}{\sqrt{y}} = ax + b$	$X = x, Y = \frac{1}{\sqrt{y}}$
$y = \frac{c \cdot x}{e^{dx}}$	$\ln\left(\frac{y}{x}\right)$ $= -dx + \ln(c)$	$X = x, Y = \ln\left(\frac{y}{x}\right), c = e^b, d = -a$
$y = \frac{1}{1 + c \cdot e^{ax}}$	$\ln\left(\frac{1}{y} - 1\right)$ $= ax + \ln(c)$	$X = x, Y = \ln\left(\frac{1}{y} - 1\right), c = e^b$

Задание на лабораторную работу

По экспериментальным данным представленным в таблице 4 по варианту, заданному преподавателем, найти уравнение линейной регрессии методом наименьших квадратов и с помощью аппроксимирующей функции lsqnonlin() в Matlab.

Пример 1. Данные измерения температуры остывания чашечки кофе с течением времени представлены в таблице 3.

Таблица 3 – Зависимость температуры чашки кофе от времени при остывании

Т, мин						5	5	8	0
Т t, °C	5	6	5	9	5	0	8	7	4

Для аппроксимации данных зависимостью $F_1(t, a, b) = a \cdot \exp(b \cdot t)$ проведем ее линеаризацию. Для этого

прологарифмируем это выражение и введем замену переменных. Получим: $\ln(F_1(t, a, b)) = \ln(a) + b \cdot t$. Обозначим $Y_1 = \ln(F_1(t, a, b))$, $X = t$, $A = b$, $B = \ln(a)$. Тогда $Y_1 = A \cdot X + B$, где A , B - неизвестные коэффициенты уравнения линейной регрессии.

Листинг программы в Matlab для нахождения коэффициентов A и B :

```
t = [0 2 3 5 8 15 25 28 30] ;
T =[85 76 55 49 45 30 28 27 24];
% Аппроксимация данных зависимостью T-a*exp(b*t)
% замена переменных
Y1=log(T);
X=t;
N=length(X);
p=zeros(2,2);
q=zeros(1,2);
% определение коэффициентов матрицы
p(1,1)=sum(X.^2);
p(1,2)=sum(X);
p(2,1)=p(1,2);
p(2,2)=N;
q(1)=dot(X,Y1);
q(2)=sum(Y1);
% решение СЛАУ
S=p^-1*q';
% значения параметров
A=S(1);
B=S(2);
% обратная замена переменных
a1=exp(B);
```

```

b=A;
% Аппроксимация данных зависимостью T=a*t+b
k=zeros(2,2);
d=zeros(1,2);
% определение коэффициентов матрицы
k(1,1)=sum(t.^2);
k(1,2)=sum(t);
k(2,1)=k(1,2);
k(2,2)=N;
d(1)=dot(t,T);
d(2)=sum(T);
% решение СЛАУ
W=k^-1*d';
% значения параметров
a2=W(1)
b2=W(2)
% Построение графика функции
tp=0:0.01:30;
T1=a1*exp(b1*tp);
T2=a2*tp+b2;
plot(t,T,'rd',tp,T1,'b-',tp,T2,'g-','LineWidth',2)
xlabel('t,мин');
ylabel('T,C')
% Вычисление суммы квадратов отклонений:
T1E=a1*exp(b1*t);
T2E=a2*t+b2;
Sig1=sum((T1E-T).^2)
Sig2=sum((T2E-T).^2)

```

В результате работы программы получаем значение параметров приближающих функций и записываем ответ:

$$F_1 = 67,5813 \cdot \exp(-0,0365 \cdot t), \quad F_2 = -1,6237 \cdot t + 67,4831. \text{ На}$$

рисунке 2 представлены исходные данные и аппроксимирующие функции.

Результат вычисления отклонений демонстрирует, что наиболее подходящей зависимостью является $F_1(t, a, b) = a \cdot \exp(b \cdot t)$. Отклонения для F_1 и F_2 составляют соответственно:

$$\text{Sig1}=683,8837;$$

$$\text{Sig2}=926,4355.$$

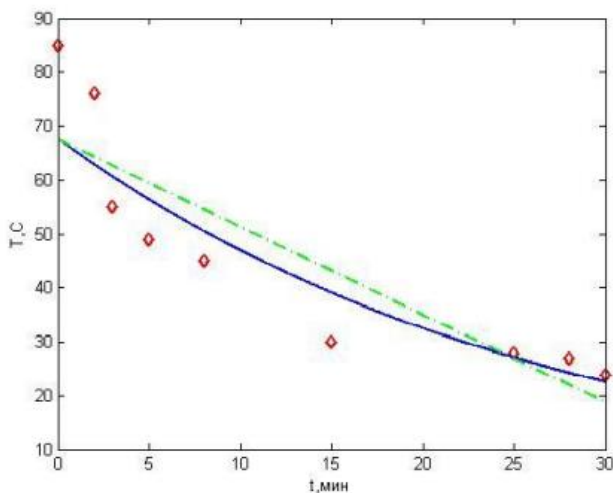


Рисунок 2 – Исходные данные для примера 1 и аппроксимирующие функции

Пример 2 Используя табличные значения примера 1, найти параметры аппроксимирующих функций с помощью встроенной функции Matlab.

Параболическая зависимость приводит к наименьшему значению отклонения экспериментальных данных от теоретической кривой. Данный случай демонстрирует, что выбор эмпирической зависимости не всегда должен быть основан на подобном анализе. При выборе аппроксимирующей функции в первую очередь во внимание следует принимать

физический смысл решаемой задачи (для описания процесса остывания наиболее подходящей монотонно убывающая функциональная зависимость).

Для примера выберем две аппроксимирующие зависимости:

$$F_1(t, a, b) = a \cdot \exp(b \cdot t) \text{ и } F_2(t, a, b, c) = a \cdot t^2 + b \cdot t + c.$$

Создадим файлы FFF1.m и FFF.m, содержащие функции, которая возвращают значения вектор-функций, представляющих собой отклонения теоретических кривых от экспериментальных точек:

```
function z=FFF1(P,x,y)
i=1:length(x);
z(i)=y(i)-(P(1)*exp(P(2)*x(i)));
function z=FFF(P,x,y)
i=1:length(x);
z(i)=y(i)-(P(1)*x(i).^2+P(2)*x(i)+P(3));
```

Вычислим значения аппроксимирующих функций с помощью функции lsqnonlin() и визуализируем исходные данные и аппроксимирующие функции:

```
t=[0 2 3 5 8 15 25 28 30] ;
T=[85 76 55 49 45 30 28 27 24];
% аппроксимация данных зависимостью T=a*exp(b*t)
z=[0 0];
P=lsqnonlin('FFF1',z,[],[],[],t,T)
tp=0:0,01:30;
T1=P(1)*exp(P(2)*tp);
% аппроксимация данных зависимостью T=a*t^2+b*t+c
z=[0 0 0];
P1=lsqnonlin('FFF',z,[],[],[],t,T)
```

```

tp=0:0,01:30;
T2=P1(1)*tp.^2+P1(2)*tp+P1(3);
T1E=P1(1)*exp(P(2)*t);
T2E=P1(1)*t.^2+P1(2)*t+P1(3);
Sig1=sum( (T1E-T) .^2)
Sig2=sum((T2E-T).^2)
% построение графика функции
plot(t,T,'bo',tp,Tl,'r-',tp,T2,'g--','LineWidth',2)
xlabel('t,мин');
ylabel('T,C')

```

Исходные данные и графики аппроксимирующих функций представлены на рисунке 3.

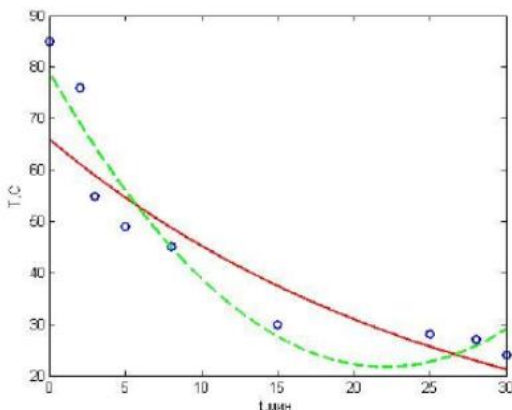


Рисунок 3 – Исходные данные для примера 2 и аппроксимирующие функции

Вид аппроксимирующих функций: $F_1 = 65,9695 \cdot \exp(-0,0365 \cdot t)$,

$$F_2 = 0,1177 \cdot t^2 - 5,1993 \cdot t + 790855.$$

Отклонения для F_1 и F_2 составляют соответственно:

$$\text{Sig1} = 728,127;$$

$$\text{Sig2} = 283,5526.$$

Таблица 4 – Задания для выполнения лабораторной работы 1

Вариант 1	В таблице приведены данные зависимости удельной намагнитченности I ферромагнитных тел в состоянии насыщения от различных значений температуры для электролитического железа Fe					
	$t, ^\circ\text{C}$	-188	105	526	700	772
	I	221	213	179	120	24
Вариант 2	Нормальная термопара Ле-Шателье (Pt-Pt10%Rh) рекомендована международным комитетом как основной прибор при точных измерениях температур. В таблице приведены данные зависимости значений электродвижущей силы E от температуры					
	$t, ^\circ\text{C}$	100	300	700	1200	1600
	$E, \text{мВ}$	1,435	3,240	7,339	13,155	17,922
Вариант 3	Нормальная термопара Ле-Шателье (Pt-Pt10%Rh) рекомендована международным комитетом как основной прибор при точных измерениях температур. В таблице приведены данные зависимости значений электродвижущей силы E от температуры T					
	$T, ^\circ\text{C}$	0	200	400	1100	1500
	$E, \text{мВ}$	0,112	1,605	3,44	10,993	17,007
Вариант 4	В таблице приведены данные зависимости плотности водных растворов серной кислоты при различных значениях концентрации солей (при температуре 40°C)					
	$\rho, \text{г/см}^3$	0,9986	1,0371	1,1202	1,3555	1,4816
	$n, \%$	1	7	19	42	60
Вариант 5	В таблице приведены данные зависимости удельной намагнитченности I ферромагнитных тел в состоянии насыщения от различных значений температуры T для кобальта Co					
	$T, ^\circ\text{C}$	17	470	804	971	1149
	I	166	151	125	99	18
Вариант 6	В таблице приведены данные зависимости коэффициента внутреннего трения (вязкости) для воды, свободной от примесей, от температуры					
	$T, ^\circ\text{C}$	0	15	45	65	75

	$\eta \cdot 100$ г/(см·с)	1,788	1,14	0,595	0,435	0,379
Вариант 7	В таблице приведены данные зависимости коэффициента внутреннего трения (вязкости) для водного раствора этилового спирта (80%) от температуры					
	T, °C	5	26	50	60	75
	I	3,125	1,784	0,968	0,789	0,6
Вариант 8	В таблице приведены данные зависимости плотности водных растворов серной кислоты при различных значениях концентрации солей (при температуре 0°C)					
	ρ , г/см ³	0,9999	1,0660	1,1351	1,2661	1,4618
	n, %	0	9	19	36	55
Вариант 9	В таблице приведены данные зависимости коэффициента внутреннего трения (вязкости) для водного раствора сахара (20%) от температуры					
	T, °C	5	26	50	60	75
	$\eta \cdot 100$ г/(см·с)	3,125	1,784	0,968	0,789	0,6
Вариант 10	В таблице приведены данные зависимости удельной намагниченности I ферромагнитных тел в состоянии насыщения от различных значений температуры T для железной магнитной руды Fe ₃ O ₄					
	T, °C	-253	110	497	557	578
	$\eta \cdot 100$ г/(см·с)	99	88	46	26	0

Содержание отчета:

1. Исходные данные.
2. Листинг программы для нахождения коэффициентов линейной регрессии.
3. Результаты нахождения коэффициентов уравнения регрессии.
4. Графики с исходными данными и аппроксимирующими функциями.

5. Отклонения аппроксимирующих функций от исходных данных и выводы.

Лабораторная работа 2

Определение уравнения линейной и нелинейной регрессии по экспериментальным данным многофакторного эксперимента

Теоретическая часть

Пусть имеются N результатов наблюдения y_i выходной переменной Y в функции от n входных факторов X_i (таблица 5).

Таблица 5 – Экспериментальные данные многофакторного эксперимента

№ п/п	Y	X_1	X_2	...	X_N
1	y_1	x_{11}	x_{21}	...	x_{n1}
2	y_2	x_{12}	x_{22}	...	x_{n2}
...
N	y_N	x_{1N}	x_{2N}	...	x_{nN}

Тогда уравнение множественной линейной регрессии имеет вид

$$Y = b_0 + \sum_{i=1}^n b_i X_i. \quad (7)$$

Задачу нахождения неизвестных коэффициентов b_i удобно решать в матричном виде. При этом коэффициенту b_0 целесообразно поставить в соответствие фиктивный параметр X_0 , который всегда равен единице. Если ввести в рассмотрение матрицу измеренных значений X , вектор наблюдений Y и вектор коэффициентов B :

$$X = \begin{bmatrix} x_{01} & x_{11} & \dots & x_{n1} \\ x_{02} & x_{12} & \dots & x_{n2} \\ \dots & \dots & \dots & \dots \\ x_{0N} & x_{1N} & \dots & x_{nN} \end{bmatrix}; \quad (8)$$

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_N \end{bmatrix}; \quad B = \begin{bmatrix} b_0 \\ b_1 \\ \dots \\ b_N \end{bmatrix},$$

то система уравнений (8) примет следующий матричный вид:

$$X^T X B = X^T Y. \quad (9)$$

Тогда искомые коэффициенты b_i могут быть вычислены как

$$B = (X^T X)^{-1} X^T Y. \quad (10)$$

Решение возможно, если обратная матрица в выражении (10) является невырожденной, т. е. если параметры X_1, X_2, \dots, X_N линейно независимы.

Для нелинейной множественной регрессии связь между входными и выходными параметрами обычно выражают в виде

$$Y = b_0 + \sum_{i=1}^n b_i X_i + \sum_{i=1}^n b_{ii} X_i^2 + \sum_{i,j=1}^{C_n^2} b_{ij} X_i X_j, \quad (11)$$

где $C_n^2 = \frac{n!}{j!(n-j)!}, j = 0, 1, \dots \leq n = 0, 1, \dots$ - число сочетаний из n элементов по 2.

Можно представить члены второго порядка как некоторые новые переменные. Тогда уравнение (11) приобретет вид

$$Y = b_0 + \sum_{i=1}^n b_i X_i + \sum_{k=n+1}^{C_n^2} b_k X_k, \quad (12)$$

Для одномерной системы уравнение нелинейной регрессии примет вид

$$Y = F(X_1) = b_0 + b_1 X_1 + b_2 X_1^2,$$

$$X = \begin{bmatrix} 1 & x_{11} & x_{11}^2 \\ 1 & x_{12} & x_{12}^2 \\ \dots & \dots & \dots \\ 1 & x_{1N} & x_{1N}^2 \end{bmatrix}. \quad (13)$$

Для двумерной системы – $F(X) = b_0 + b_1X_1 + b_2X_2 + b_3X_1X_2 + b_4X_1^2 + b_5X_2^2$,

$$X = \begin{bmatrix} 1 & x_{11} & x_{21} & x_{11}x_{21} & x_{11}^2 & x_{21}^2 \\ 1 & x_{12} & x_{22} & x_{12}x_{22} & x_{12}^2 & x_{22}^2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & x_{1N} & x_{2N} & x_{1N}x_{2N} & x_{1N}^2 & x_{2N}^2 \end{bmatrix}. \quad (14)$$

Задание на лабораторную работу

По экспериментальным данным, представленным в таблице 7 по варианту, заданному преподавателем, найти уравнение линейной и нелинейной множественной регрессии в Matlab.

Пример 3. Дана таблица значений зависимости прибыли Z некоторой компании от объема продаж двух видов товаров X и Y (таблица 6).

Таблица 6 – зависимость прибыли от объема продаж двух видов товаров

X, ед.	2	2	5	5	5	2	10	10	0
Y, ед.	1	2	1	2	3	3	1	2	3
Z, усл.ед.	11,3	16,1	21	26	29	20,9	36	39	13

Выполним линейную аппроксимацию, приняв $Z = b_0 + b_1X + b_2Y$. Процедура определения параметров b_0 , b_1 , b_2 представлена в листинге:

```

x=[2 2 5 5 5 2 10 10 0] ;
y=[1 2 1 2 3 3 1 2 3];
z=[11.3 16.1 21 26 29 20.9 36 39 13];
N=length(x);
A=zeros(3,3);
B=zeros(1,3);
A(1,1)=N;
A(1,2)=sum(x);
A(1,3)=sum(y);
A(2,1)=A(1,2);
A(2,2)=sum(x.^2);
A(2,3)=dot(x,y);
A(3,1)=A(1,3);
A(3,2)=A(2,3);
A(3,3)=sum(y.^2);
B(1)=sum(z);
B(2)=dot(x,z);
B(3)=dot(y,z);
L=A^-1*B'
a0=L(1);
a1=L(2);
a2=L(3) ;
[X/Y]=meshgrid(2:0.2:10,1:0.1:3);
Z=a0+a1*X+a2*Y;
mesh(X,Y,Z)
hold on
plot3(x,y,z,'ro','MarkerSize',8)

```

Исходные данные и результат аппроксимации представлены на рисунке 4. Аппроксимирующая функция будет иметь вид:

$$Z = 1,7422 + 2,9975 \cdot X + 4,0958 \cdot Y.$$

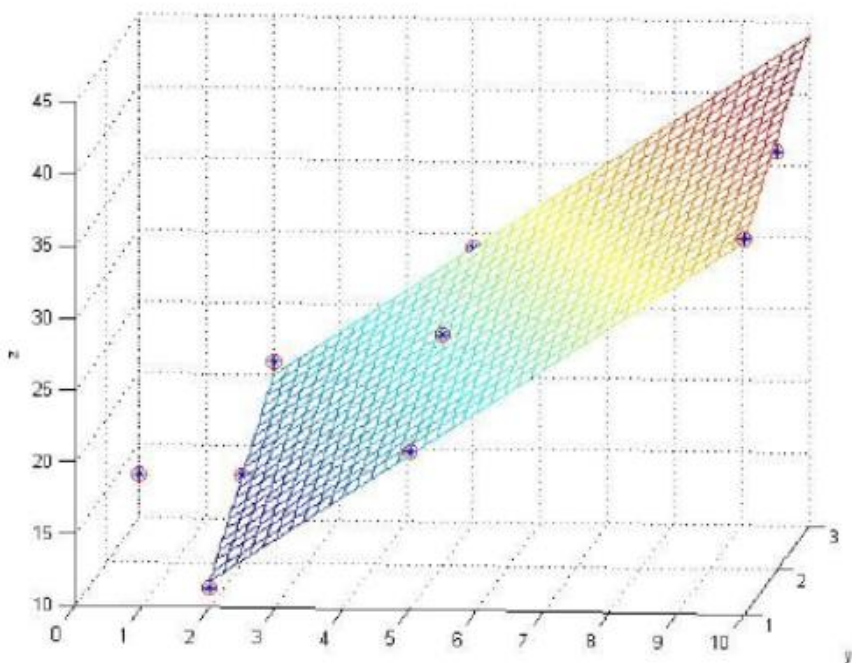


Рисунок 4 – Результат многомерной линейной регрессии

Таблица 7 – Задания для выполнения лабораторной работы 2.

№	Исходные данные									
1	х, ед.	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
	у, ед.	20	80	60	40	10	50	30	80	90
	z, усл. ед.	2,3	1,1	2,8	1,8	3,1	4,5	5,3	4,1	6,2
2	х, ед.	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
	у, ед.	20	80	60	40	10	50	30	80	90
	z,	5,2	2,1	4,8	3,6	5,1	6,5	7,2	6,1	8,2

	усл.е д.									
3	х, ед.	1	2	3	4	5	6	7	8	9
	у, ед.	20	80	50	40	10	40	30	80	100
	z, усл.е д.	12, 3	11, 1	12, 8	11, 8	13, 1	14, 5	15, 3	14, 1	16, 2
4	х, ед.	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
	у, ед.	10	70	60	30	10	60	30	70	80
	z, усл.е д.	1,3	0,1	1,8	0,8	2,1	3,5	4,3	3,1	5,2
5	х, ед.	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
	у, ед.	20	80	60	40	10	50	30	80	90
	z, усл.е д.	2,3	1,1	2,8	1,8	3,1	4,5	5,3	4,1	6,2
6	х, ед.	1	2	3	4	5	6	7	8	10
	у, ед.	20	80	60	40	10	50	30	80	100
	z, усл.е д.	5,3	3,1	4,8	3,8	5,1	6,5	7,3	5,1	9,2
7	х, ед.	10	20	30	40	50	60	70	80	90
	у, ед.	20	80	60	40	10	50	30	80	90
	z, усл.е д.	23	11	28	18	31	45	53	41	62
8	х, ед.	10	20	30	40	50	60	70	80	90
	у, ед.	20	80	60	40	10	50	30	80	90
	z, усл.е	240	111	284	180	315	452	538	412	627

	д.									
9	х, ед.	10	20	30	40	50	60	70	80	90
	у, ед.	20	80	60	40	10	50	30	80	90
	z, усл.е д.	230 0	110 5	280 2	180 1	311 4	453 6	535 6	415 8	620 0
1 0	10	20	30	40	50	60	70	80	90	10
	2	8	6	4	1	6	4	8	10	2
	231	114	285	180	317	455	537	412	620	231

Содержание отчета:

1. Исходные данные.
2. Листинг программы для нахождения коэффициентов линейной и нелинейной многомерной регрессии.
3. Результаты нахождения коэффициентов уравнений регрессии.
4. Графики с исходными данными и аппроксимирующими функциями.
5. Отклонения аппроксимирующих функций от исходных данных и выводы.

Лабораторная работа 3

Аппроксимация экспериментальных данных многофакторного эксперимента с помощью искусственной нейронной сети в Matlab

Теоретическая часть

Для аппроксимация экспериментальных данных многофакторного эксперимента в настоящее время широко используются искусственные нейронные сети (ИНС) прямого распространения. Типовая структура ИНС изображена на рисунке 5, где введены следующие обозначения:

n – количество входов;

m – количество нейронов первого слоя;

p – количество нейронов второго слоя;

w_{ijk} – весовой коэффициент связи j -го нейрона i -го слоя с k -ым входом;

b_{ij} – коэффициент, задающий смещение функции активации j -го нейрона i -го слоя по входу;

N_{ij} – j -й нейрон i -го слоя;

I_k – k -й вход;

O_l – l -й выход.

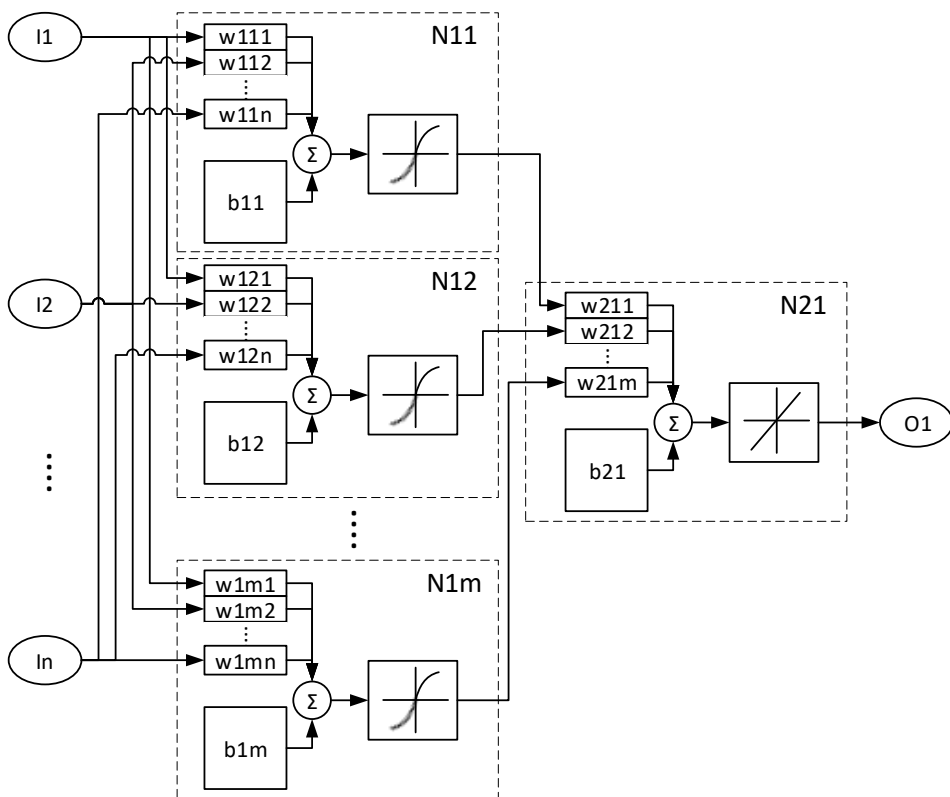


Рисунок 5 – Структура ИНС прямого распространения

Обучение сети будем проводится одним из методов обратного распространения ошибки с оптимизацией значений весовых коэффициентов нейронной сети w_{ijk} , например, Левенберга-Марквардта.

Пример 4. Выбор структуры, настройка параметров и обучение ИНС.

Цель обучения нейронной сети – рациональный выбор скорости комбайна при уборке пшеницы, при которой потери зерна при уборке будут минимальными. Входными

параметрами являются: потери зерна и номер испытываемого комбайна. Для корректной работы нейронной сети необходимо создать массивы данных, предварительно взятых из табличных значений (ссылки на табличные источники в списке литературы). Входной массив данных Inputs для обучения нейронной сети представлен в столбцах 1 и 2 таблицы 8, которая сформирована на основе экспертных данных. Выходной массив данных для обучения нейронной сети Outputs представлен в столбце 3 таблицы 8.

Таблица 8 – Зависимость зернопотерь от подачи зерна для разных комбайнов

Номер испытываемого комбайна	Зернопотери, %	Подача зерна, т/час
1	0,98	7
1	1,26	10
1	1,44	12
1	1,78	14
1	2,1	15
2	0,78	6
2	0,8	7
2	0,98	10
2	1,18	12
2	2,62	14
3	0,8	15
3	0,7	18
3	1,06	4
3	1,38	5
3	2,1	6

Подготовим и настроим ИНС в Matlab с помощью инструмента nnstart. Установим входные параметры (подача зерна и потери), а также выходной параметр – скорость движения комбайна.

На рисунке 6, приведена панель настройки входных параметров.

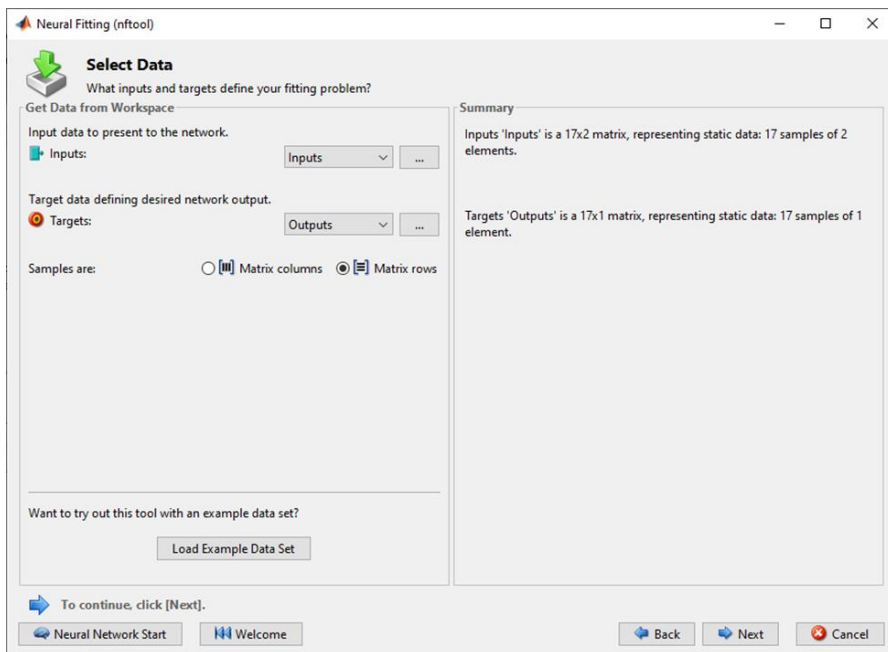


Рисунок 6 – Панель настройки входных параметров

На рисунке 7 изображено разделение данных на обучающую выборку – тренировку(training), на проверку сети(validation) и на тестовый прогон сети(testing). Здесь задаём: 70% - количество слоёв нейронов для обучения, 15% для валидации и оставшиеся 15% для проверки полученного результата.

На рисунке 8 изображена установка количества нейронов в скрытом слое сети. После установки предыдущих показателей следует установка количества нейронов в данном случае 100 (это число получено после нескольких экспериментов). После этого производится обучение нейронной сети. Для этого проведем серию тестов на приборной панели программы, а именно: подберём разное количество нейронов в слое для получения наиболее точного результата, изменим процентное соотношение показателей (обучающей выборки, проверки сети и тестового прогона сети). Этот процесс изображен на рисунке 9.

Проводим серию тестов для нахождения наилучшего результата.

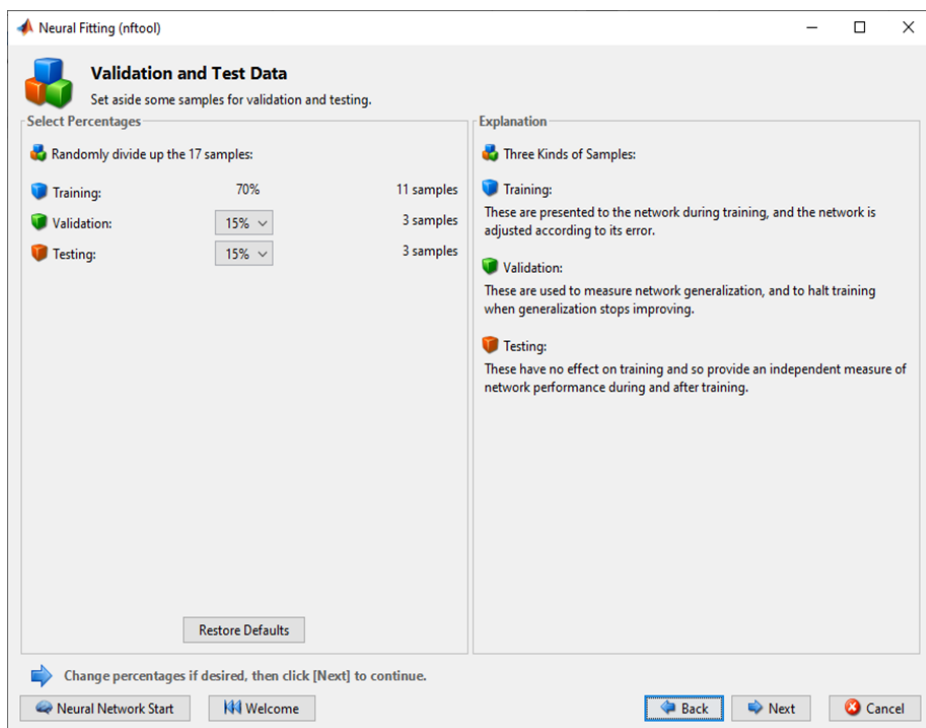


Рисунок 7 – Разделение полученных данных

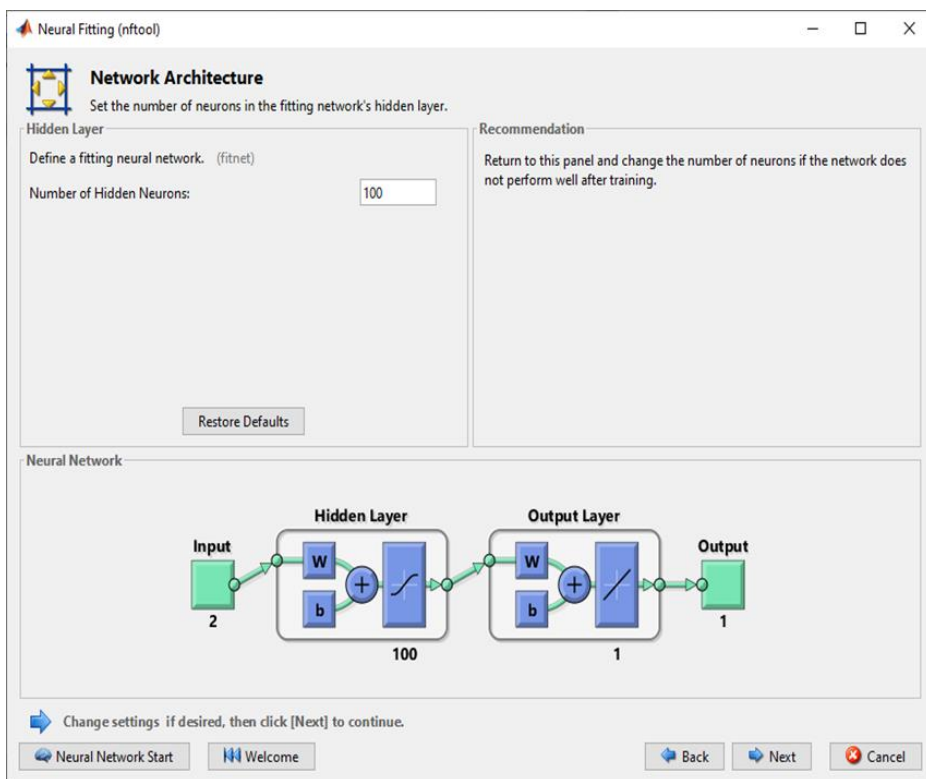


Рисунок 8 – Установка количества нейронов в скрытом слое сети

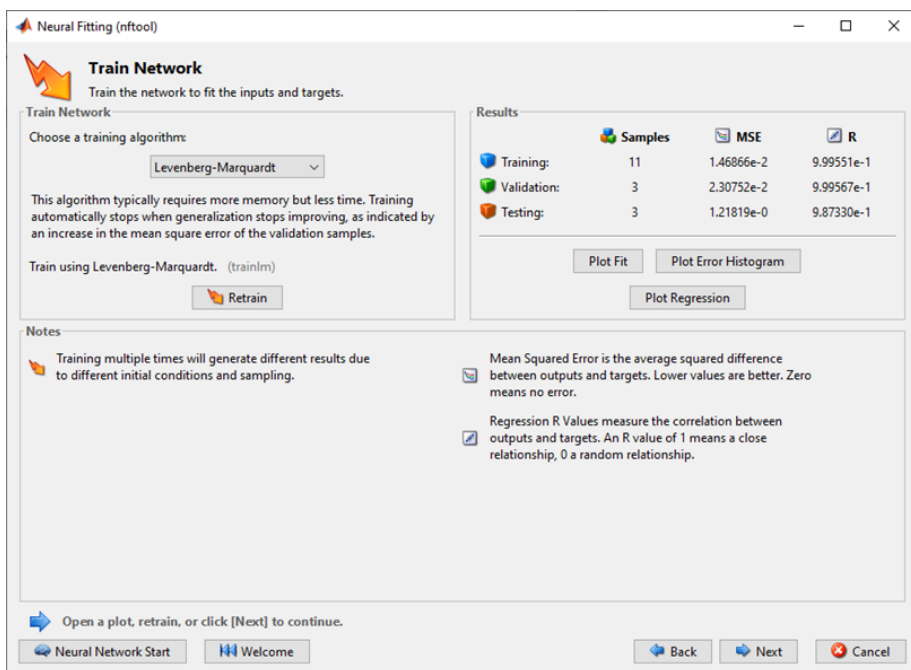


Рисунок 9 - Обучение нейронной сети

Качество обученной сети можно наблюдать на графике 4 (рисунок 10). При условии, что количество проведённых тестов больше 1. Чем ближе точки расположены к линии, а сама линия максимально близко к пунктиру, тем выше точность обученной сети. Параметр R – коэффициент корреляции. Чем ближе его значение к единице тем ближе экспериментальные данные к значениям, предсказываемым ИНС. При неудовлетворительных значениях R переобучить ИНС.

Программный код, использующий обученную ИНС для предсказания подачи зерна приведена в Приложении А (Листинг А.1).

Зависимость подачи от типа комбайна и потерь, построенная на основе обученной нейронной сети

представлена на рисунке 11. Вертикальная ось Y показывает количество поданного зерна, ось X – процент зернопотерь, ось Z – тип комбайна (рассматривать целые значения).

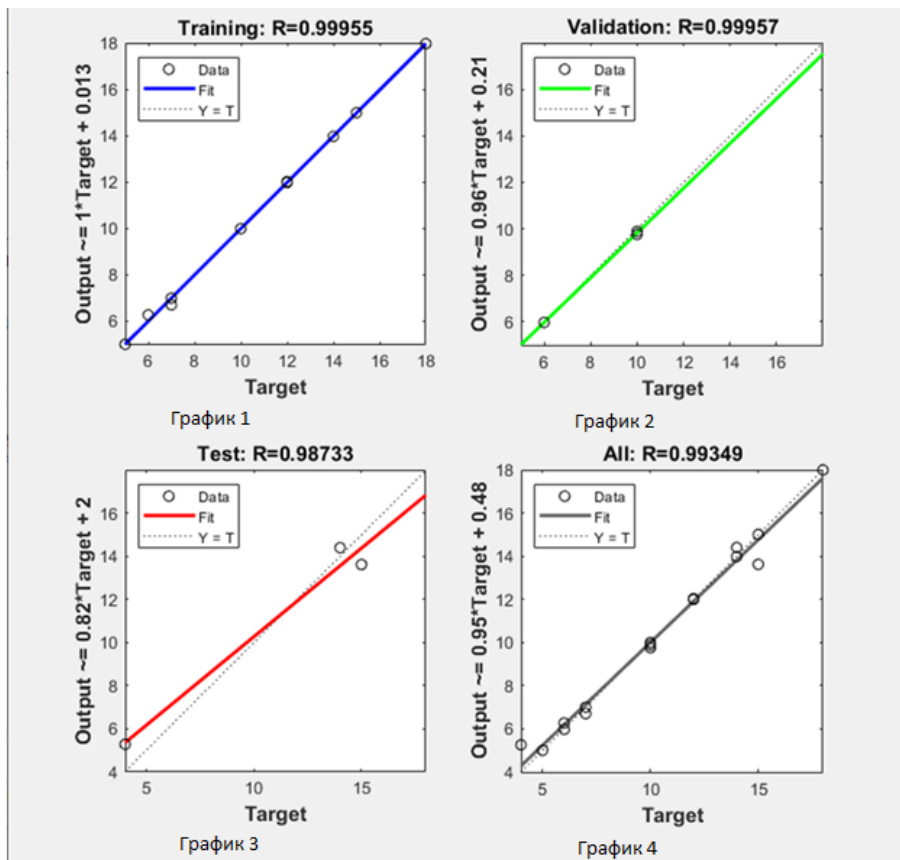


Рисунок 10 – Качество обучения нейронной сети

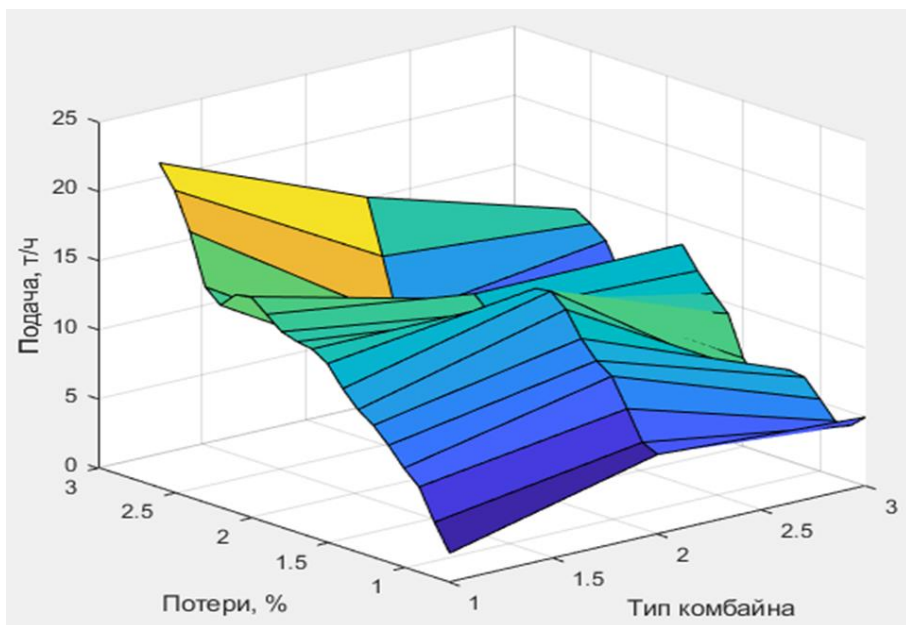


Рисунок 11 - Зависимость подачи от типа комбайна и потерь

Пример 5 Сравнение качества прогнозирования скорости комбайна для системы управления движением, полученных с помощью нейронной сети с другими методами

Сравним результаты, полученные в примере 4 с результатами, которые рассчитаем, используя метод регрессии. Для оценки качества прогноза скорости движения комбайна обоими методами будем использовать известный в дисперсионном анализе коэффициент корреляции R , с помощью которого будем оценивать степень взаимосвязи экспериментальных значений скорости комбайна из таблицы 8 с прогнозируемыми значениями скорости, полученными по модели на основе нейронной сети и модели на основе одного из классических методов аппроксимации. В качестве

классического метода будем использовать метод многомерной нелинейной регрессии, описываемой уравнениями (12) и (14).

Для данных в таблице 4 по модели, описываемой выражениями (8) и (9), составлена программа в среде Matlab, листинг А.2 которой приведен в Приложении А. Результаты расчетов представлены на рисунке 12, на котором изображена поверхность значений скоростей комбайна в функции зернопотерь и номера комбайна, полученная на основе нелинейной регрессии. Полученная модель, как и модель на основе нейронной сети из примера 4 позволяет прогнозировать не только скорости комбайна для значений зернопотерь и номера комбайна из таблицы 8, но и скорость для промежуточных значений, не указанных в таблице. На рисунке 13 графически изображена корреляция между скоростями комбайна, предсказанными по модели на основе нейронной сети и модели, полученной методом регрессии (синяя кривая). Коэффициент корреляции, соответствующий этой кривой, $R=0,9574$. Красная прямая линия на рисунке 13 соответствует идеальному случаю, при котором коэффициент корреляции $R=1$.

Сравнивая значения коэффициентов корреляции для предсказаний скорости движения комбайна, полученных с помощью нейронной сети $R_N=0,9935$ (рисунок 10) и методом регрессии $R_R=0,9574$, можно сделать вывод, что предсказанная с помощью нейронной сети скорость комбайна наиболее приближена к полученной экспериментально (таблица 8). Данный факт свидетельствует о предпочтительности применения модели на основе нейронной сети для решения задачи аппроксимации экспериментальных данных аналитической функцией. Также интересным становится исследование результатов применения аппроксимации одномерных и многомерных массивов

экспериментальных значений выходных факторов в функции от входных в других задачах, где затруднено получение математической модели объекта или процесса в виде аналитических уравнений, полученных из известных уравнений, описывающих базовые физические законы (закон сохранения вещества, энергии, заряда, динамики движения твердых тел и др.)

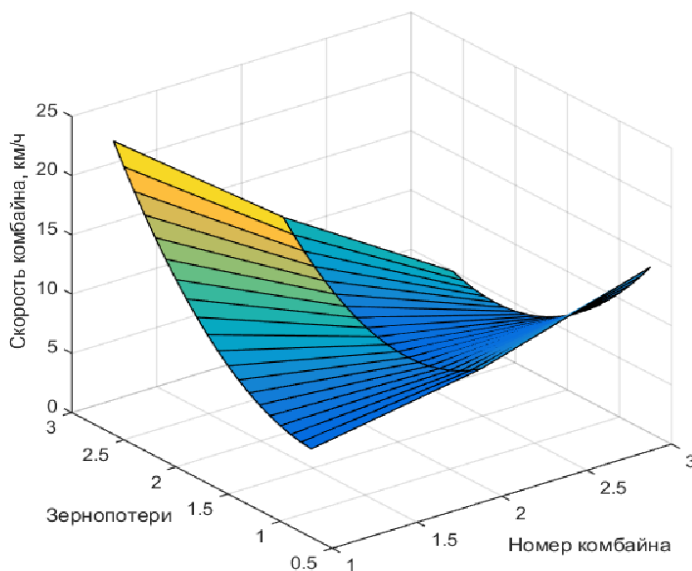


Рисунок 12 – Зависимость подачи от типа комбайна и потерь, полученная методом регрессии

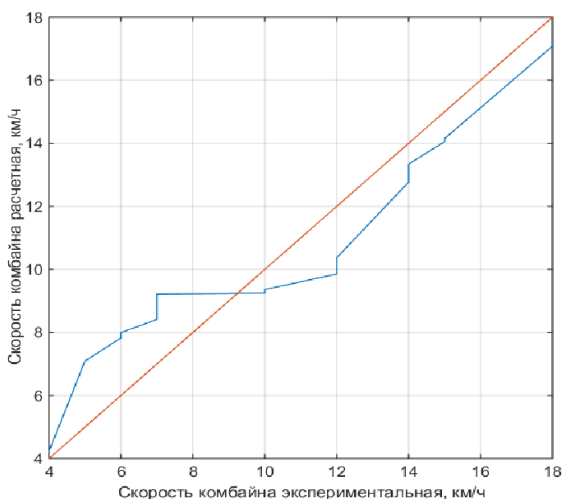


Рисунок 13 – Корреляция предсказанных значений скоростей комбайна, полученных методом аппроксимации многомерной функции с помощью нейронной сети и методом регрессии

Задание на лабораторную работу 3

Для набора экспериментальных данных из лабораторной работы 2 выполнить аппроксимацию с помощью пакета работы с нейронными сетями nnstart. Сравнить результаты аппроксимации с результатами, полученными в лабораторной работе 2 методом нелинейной многомерной регрессии.

Содержание отчета:

1. Исходные данные.
2. Листинг программы, содержащей обученную ИНС.
3. Исходные данные и графики аппроксимирующих функций (ИНС и нелинейной многомерной регрессии).
4. Графики коэффициентов корреляции для разных методов аппроксимации.
5. Выводы.

Лабораторная работа 4

Аппроксимация экспериментальных данных способом избранных точек

Теоретическая часть

В тех случаях, когда задачу аппроксимации значений выходной переменной Y в функции от одного фактора X нельзя решить с помощью линейной регрессии, а при применении нелинейной регрессии получается вырожденная обратная матрица, для нахождения выражения функциональной зависимости $Y=f(X)$ используется способ избранных точек.

Порядок вычислений по этому способу, следующий:

1. Используя типовые сглаженные графики функций для подбора видов уравнений эмпирических формул, выбирают график, который наиболее близко описывает зависимость Y от X . Наиболее известные графики и их уравнения приведены в приложении Б.

2. На кривой сглаженного графика выбирают n точек (по числу подлежащих определению коэффициентов, входящих в формулу, описывающую сглаженный график). Для удобства абсциссы этих точек должны совпадать с делениями шкалы по оси x и отстоять дальше одна от другой (но не на концах кривой, которые ненадежны).

3. Каждую из n пар опытных значений x и y подставляют в избранное уравнение, что дает систему n уравнений с n неизвестными.

4. Эту систему уравнений решают относительно неизвестных, то есть искомых коэффициентов формулы.

5. Если избранный тип уравнения линеен относительно определяемых коэффициентов или его можно привести к линейному виду, применяют определители, что упрощает

вычисления. После определения коэффициентов избранного уравнения устанавливают соответственно уравнения опытной сглаженной кривой, проверяя ряд точек (их количество должно быть не меньше, чем количество опытных точек). Составляют таблицу отклонений результатов вычислений от опытной кривой и определяют абсолютную и относительную ошибки принятой формулы.

6. Пункты 2-5 выполняют до тех пор, пока абсолютная и относительная погрешности не станут превышать допустимых значений. Найденные при этом коэффициенты избранного уравнения будут искомыми.

Пример 6. Найти формулу и определить её коэффициенты для экспериментально полученных значений функциональной зависимости глубины пахоты от затупления лемеха (таблица 9).

Таблица 9 - Опытные и расчетные глубины вспашки в функции степени затупления лемехов

X	Y	lg(Y)	Значения по эмпирической формуле	Отклонения расчетных данных от опытных, см
0	23,5	1,371	23,5	0,0
1	23,1	1,364	23,2	0,1
2	22,4	1,350	22,2	0,2
3	20,8	1,318	20,8	0,0
4	19,2	1,284	19,1	0,1
5	18,6	1,270	18,9	0,3
6	18,5	1,268	18,7	0,2
7	18,4	1,265	18,3	0,1

Найдем эмпирическую формулу для кривой функциональной зависимости глубины пахоты от затупления лемеха. По типичным кривым подбираем подходящую, уравнение которой

$$Y = ab^{-X^c}.$$

Определяем, подходит ли эта формула по экстремальным значениям функции. Предположим, что для нас важно, чтобы формула соответствовала первоначальным степеням затупления. Убеждаемся, что при $X=0$, $Y=a$. Это удовлетворяет наше требование, так как при $X=0$ $Y=a=23,5$. Коэффициенты формулы определяем по способу избранных точек.

Для удобства расчета логарифмируем принятый вид формулы:

$$\lg(Y) = \lg(23,5) - X^c \lg(b).$$

В таблице 9 первый (X) и второй (Y) столбцы представляют степени затупления лемехов и сглаженные опытные значения глубины пахоты. В графе $\lg(Y)$ помещены логарифмы значений глубины пахоты.

Для определения коэффициентов b и c используем координаты двух точек и составляем уравнения:

$$\lg(y_1) = 1,364 = 1,371 - 1^c \lg(b);$$

$$\lg(y_2) = 1,284 = 1,371 - 4^c \lg(b).$$

Из них находим, что $b=1,016$ и $c = 1,82$. Таким образом, окончательно эмпирическая формула будет такой:

$$Y = 23,5 \cdot 1,016^{-X^{1,82}}.$$

Вычисляем для проверки значения y по найденной эмпирической формуле и заносим их в четвертый столбец таблицы. Средняя абсолютная ошибка формулы будет равна:

$$\Delta = \frac{0,0 + 0,1 + 0,2 + 0,0 + 0,1 + 0,3 + 0,2 + 0,1}{8} = 0,125 \text{ см.}$$

Отнесенная к средней глубине пахоты средняя относительная ошибка

$$\Delta_2 = \frac{0,125}{20,5} \cdot 100 = 0,61\%.$$

Максимальная относительная ошибка

$$\Delta_{2\text{макс}} = \frac{0,3}{18,6} \cdot 100 = 1,61\%.$$

Ошибки не превышают допустимого значения (5%), поэтому найденные коэффициенты принимаются как окончательное решение.

Задание на лабораторную работу 4

Для набора экспериментальных данных из таблицы 10 по варианту, заданному преподавателем, определить подходящий сглаженный график функции и определить неизвестные коэффициенты формулы, описывающей график, способом избранных точек при условии чтобы средняя и максимальная относительная ошибки не превышали 5%.

Содержание отчета:

1. Исходные данные.
2. Листинг программы в Matlab расчета коэффициентов формулы, описывающей график функции методом избранных точек.
3. Проверка соответствия средней и максимальной относительной ошибок допустимым значениям.
4. Графики исходных данных и аппроксимирующей функции.
5. Значения найденных коэффициентов формулы.

Таблица 10 – Экспериментальные данные для выполнения лабораторной работы 4

№	Данные для расчёта												
1	x	-7,465	-7,961	-9,038	-9,631	10,892	11,761	13,618	15,186	17,158	18,936	20,995	-
	y	-7,465	-7,961	-9,038	-9,631	10,892	11,761	13,618	15,186	17,158	18,936	20,995	-
	x	-25,061	-27,102	-28,951	-	-	-	-	-	-	-	-	-
	y	-25,061	-27,102	-28,951	-	-	-	-	-	-	-	-	-
	x	-38,077	-38,515	-36,955	36,748	35,628	34,292	32,244	30,750	28,878	27,115	25,027	23,000
	y	-38,077	-38,515	-36,955	36,748	35,628	34,292	32,244	30,750	28,878	27,115	25,027	23,000
	x	-21,005	-19,024	-17,166	15,382	13,279	11,913	10,584	-9,797	-8,261	-7,814	-7,681	-7,387
	y	-21,005	-19,024	-17,166	15,382	13,279	11,913	10,584	-9,797	-8,261	-7,814	-7,681	-7,387
2	x	6,281	5,916	5,407	4,571	3,844	2,776	1,742	0,347	-1,199	-2,822	-4,341	-6,000
	y	-13,436	-11,852	-10,140	-9,035	-7,190	-6,069	-5,171	-4,193	-3,787	-3,062	-2,247	-2,553
	x	-7,599	-9,323	-10,889	-	-	-	-	-	-	-	-	-

	y	-2,685	-3,282	-3,554	-4,154	-5,020	-6,306	-7,532	-8,556	-	-	-	-
	x	-18,022	-17,659	-17,205	-	-	-	-	-	-	-	-	-
	y	-16,566	-18,148	-19,682	-	-	-	-	-	-	-	-	-
	x	-4,413	-2,814	-1,218	0,081	1,611	2,917	3,617	4,407	5,268	5,802	6,219	6,408
	y	-27,681	-26,817	-26,640	-	-	-	-	-	-	-	-	-
3	x	3,662	3,012	1,951	1,064	0,397	-1,666	-2,962	-4,916	-6,503	-8,674	10,846	13,000
	y	29,137	31,192	33,310	35,170	37,319	38,964	39,735	41,495	42,031	42,863	43,406	43,943
	x	-15,096	-17,318	-19,139	21,036	23,058	24,382	26,343	27,564	28,450	28,600	29,517	29,519
	y	43,277	43,404	42,061	41,464	40,222	38,695	37,165	35,333	33,191	31,174	29,219	27,000
	x	-29,828	-29,082	-28,521	27,249	26,037	24,897	22,791	21,067	19,189	17,242	15,197	13,000
	y	24,907	22,714	20,540	18,665	17,144	15,425	13,941	12,293	12,073	10,719	10,498	10,624
	x	-10,904	-8,756	-6,675	-4,792	-2,860	-1,242	-0,075	1,230	1,796	3,405	3,029	3,106
	y	10,948	11,132	11,773	13,039	13,839	15,526	17,026	18,709	20,781	22,784	24,831	27,000
4	x	50,230	50,255	49,574	48,260	47,382	46,003	44,780	43,298	41,566	39,604	37,909	36,000
	y	15,909	17,720	19,584	21,065	22,537	23,979	25,412	26,670	27,018	27,719	28,537	28,415
	x	34,145	32,239	30,420	28,916	27,146	25,611	24,859	23,937	23,043	22,332	22,033	21,273

5	y	28,665	27,527	27,514	26,536	25,146	24,095	22,644	21,200	19,587	17,599	15,908	14,000
	x	22,267	22,098	23,100	23,323	24,919	25,887	26,991	28,922	30,701	32,373	34,147	36,000
	y	12,152	10,317	8,500	6,750	5,147	3,706	2,743	1,677	0,367	0,117	-0,524	-0,076
	x	37,831	39,814	41,542	43,381	44,668	46,423	46,953	48,480	49,490	49,555	50,606	50,563
	y	-0,299	0,102	0,360	1,974	2,655	3,873	5,544	6,648	8,381	10,315	12,097	14,000
	x	12,794	12,366	11,571	10,801	8,891	7,734	5,675	3,710	1,853	-0,406	-2,682	-5,000
	y	-2,717	-0,486	1,887	3,697	5,851	7,608	9,408	10,402	11,347	12,359	12,887	12,820
	x	-7,359	-9,568	-11,874	13,858	16,044	17,822	18,928	20,514	21,521	22,233	23,014	22,565
	y	12,497	11,855	11,525	10,209	8,760	7,767	5,679	3,871	1,884	-0,429	-2,646	-5,000
	x	-22,467	-21,826	-21,851	20,541	19,105	17,685	15,863	13,974	11,829	-9,664	-7,326	-5,000
6	y	-7,347	-9,660	-11,790	13,660	15,733	17,533	18,939	20,153	21,742	22,126	23,032	22,691
	x	-2,713	-0,495	1,662	3,682	5,778	7,550	9,015	10,643	11,563	12,456	13,077	13,273
	y	-22,915	-22,094	-21,730	20,636	19,024	17,386	16,013	14,125	11,746	-9,651	-7,315	-5,000
	x	22,875	22,014	22,220	21,691	20,500	19,838	18,567	17,177	15,980	14,745	13,429	12,000
	y	26,342	27,802	29,156	30,465	31,752	32,907	33,702	34,337	35,281	35,413	35,130	35,321
	x	10,566	9,278	7,892	6,855	5,723	4,691	3,369	2,738	2,401	1,669	1,741	1,745
	y	35,968	35,320	35,204	34,015	33,530	32,658	31,229	30,407	29,042	27,653	26,395	25,000
	x	1,044	1,606	1,718	2,563	3,445	4,211	5,256	6,406	8,096	9,136	10,589	12,000
	y	23,653	22,307	21,041	19,805	18,765	17,338	16,684	15,700	14,934	14,665	14,356	14,355
	x												

	x	13,400	14,764	16,210	17,214	18,513	19,849	20,548	21,565	22,106	22,418	22,357	22,866
	y	14,764	14,674	14,788	15,410	16,693	17,640	18,447	19,580	20,937	22,307	23,545	25,000
7	x	5,238	4,542	4,205	3,372	2,065	1,377	0,333	-1,039	-2,495	-3,807	-5,404	-7,000
	y	25,504	27,056	28,468	30,058	31,112	32,695	33,700	34,210	34,992	35,683	36,130	36,010
	x	-8,537	-10,104	-11,697	13,167	14,347	15,207	16,563	16,935	17,908	18,656	19,228	19,211
	y	35,872	35,326	34,954	34,491	33,746	32,571	30,951	30,037	28,530	27,064	25,503	24,000
	x	-18,783	-18,743	-17,918	17,714	16,828	15,672	14,177	12,927	11,457	10,154	-8,603	-7,000
	y	22,470	20,986	19,506	18,112	16,727	15,542	14,642	13,783	12,992	12,353	11,755	11,878
	x	-5,439	-3,894	-2,580	-0,850	0,214	1,206	2,758	3,533	4,348	4,287	4,970	5,064
	y	12,301	12,185	13,344	14,075	14,889	15,823	16,863	18,149	19,633	20,910	22,500	24,000
8	x	45,141	44,710	43,133	42,679	40,972	38,858	37,040	35,023	32,829	30,297	27,566	25,000
	y	27,638	30,242	32,568	35,090	37,270	38,904	40,565	42,319	43,902	44,164	44,913	44,724
	x	22,330	19,827	17,231	15,179	12,872	10,564	8,901	7,282	6,379	6,130	4,825	4,566
	y	45,078	44,082	43,483	42,493	41,177	39,467	37,204	34,819	32,519	30,114	27,655	25,000
	x	4,994	5,433	6,298	7,773	9,010	11,090	12,699	15,234	17,432	19,941	22,442	25,000
	y	22,348	19,890	17,182	15,075	13,009	11,034	9,041	7,703	6,659	5,362	5,087	4,950
	x	27,636	30,129	32,826	34,767	37,138	39,058	41,078	42,576	43,188	44,309	44,775	45,146
	y	4,757	5,888	6,285	7,460	9,228	10,810	13,083	15,223	17,335	19,751	22,333	25,000
9	x	13,473	13,137	11,715	11,474	10,037	8,241	6,788	5,173	3,413	1,339	-0,917	-3,000
	y	-2,045	0,085	1,952	4,040	5,817	7,087	8,429	9,915	11,284	11,390	12,045	11,824
	x	-5,187	-7,215	-9,353	-	-	-	-	-	-	-	-	-

					10,971	12,875	14,820	16,163	17,514	18,203	18,296	19,507	19,734
	y	12,210	11,308	10,865	10,089	9,042	7,651	5,752	3,769	1,904	-0,044	-2,028	-4,200
	x	-19,338	-18,994	-18,284	-	-	-	-	-	-	-	-	-
					17,023	16,055	14,293	13,048	10,916	-9,152	-7,101	-5,075	-3,000
	y	-6,369	-8,352	-10,602	-	-	-	-	-	-	-	-	-
					12,275	13,938	15,550	17,224	18,293	19,122	20,264	20,445	20,550
	x	-0,847	1,171	3,410	4,917	6,886	8,442	10,142	11,372	11,770	12,735	13,107	13,446
10	y	-20,774	-19,738	-19,497	-	-	-	-	-	-	-	-	-
					18,536	17,037	15,774	13,865	12,127	10,450	-8,491	-6,384	-4,200
	x	4,736	4,246	3,177	2,596	1,350	0,498	-1,340	-2,886	-4,835	-6,588	-8,621	-
													10,600
	y	-5,903	-4,024	-2,106	-0,388	1,283	2,806	4,252	5,047	6,700	7,405	7,359	7,489
	x	-12,635	-14,650	-16,474	-	-	-	-	-	-	-	-	-
					18,092	19,810	21,782	22,542	24,219	24,863	25,954	25,450	25,943
	y	7,207	7,362	6,207	5,100	4,540	2,911	1,318	-0,248	-2,161	-4,009	-5,832	-7,900
	x	-25,478	-25,921	-24,370	-	-	-	-	-	-	-	-	-
					24,175	23,069	21,750	19,722	18,250	16,401	14,663	12,601	10,600
	y	-9,920	-11,772	-13,692	-	-	-	-	-	-	-	-	-
					15,527	17,470	18,447	19,755	20,950	22,266	22,999	23,315	23,251
	x	-8,631	-6,675	-4,842	-3,082	-1,000	0,346	1,657	2,429	3,954	4,393	4,520	4,813
	y	-23,215	-22,579	-22,354	-	-	-	-	-	-	-	-	-
					20,967	20,266	18,566	17,195	15,663	13,901	11,777	-9,966	-7,900

Лабораторная работа 5

Экспериментальное определение частотных характеристик объекта

Теоретическая часть

Под частотной характеристикой объекта (звена, устройства, системы) принято понимать отношение изображения Фурье $F\{Y(t)\}$ некоторой выходной координаты $Y(t)$ объекта к изображению Фурье $F\{X(t)\}$ некоторой его входной координаты $X(t)$

$$\Phi(j\omega) = \frac{F\{Y(t)\}}{F\{X(t)\}} = |\Phi(j\omega)| \cdot e^{-j \arg \Phi(j\omega)}, \quad (15)$$

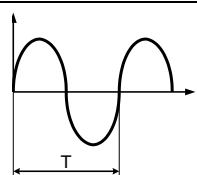
где $|\Phi(j\omega)|$ - амплитудная частотная характеристика,
 $\arg \Phi(j\omega)$ - фазовая частотная характеристика.

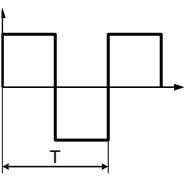
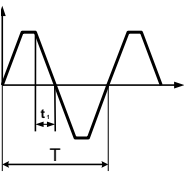
Для экспериментального определения частотной характеристики обычно используют один из трех типов пробных сигналов:

- синусоидальный сигнал;
- прямоугольный сигнал;
- трапецеидальный сигнал.

В таблице 11 представлены характеристики этих сигналов.

Таблица 11 – Характеристики пробных сигналов

№	Тип	Вид	Разложение в ряд Фурье
1	Синусоидальный		$Y(t) = A \cdot \sin \omega t$ $\omega = \frac{2 \cdot \pi}{T}$

2	Прямоугольный		$Y(t) = \frac{4 \cdot A}{\pi} \cdot \sum_{i=0}^{\infty} \frac{\sin(2 \cdot i + 1)\omega t}{2 \cdot i + 1}$ $\omega = \frac{2 \cdot \pi}{T}$
3	Трапецеидальный		$Y(t) = \frac{2 \cdot A \cdot T}{\pi^2 \cdot t_1} \cdot \sum_{i=0}^{\infty} \frac{\sin(2 \cdot i + 1)\omega t_1}{(2 \cdot i + 1)^2} \cdot \sin(2 \cdot i + 1)\omega t$ $\omega = \frac{2 \cdot \pi}{T}$

Рассмотрим использование синусоидального пробного сигнала для экспериментального определения частотных характеристик объекта. Данный пробный сигнал применяют при исследовании линейных и нелинейных объектов.

Для линейного объекта при подаче на его вход синусоидального пробного сигнала

$$X(t) = A \cdot \sin \omega_i t \quad (16)$$

выходная координата будет иметь вид

$$Y(t) = A \cdot |\Phi(j\omega_i)| \cdot \sin(\omega_i t + \arg \Phi(j\omega_i)). \quad (17)$$

Таким образом, для линейного объекта на частоте пробного сигнала ω_i значение амплитудной частотной характеристика $|\Phi(j\omega_i)|$ определяется отношением амплитуды выходного сигнала к амплитуде входного сигнала, а значение фазовой частотной характеристики $\arg \Phi(j\omega_i)$ - фазовым

сдвигом между этими сигналами. Совокупность указанных значений для нескольких фиксированных частот $\omega_i, i = 1, 2, \dots, k$ образует экспериментальную амплитудно-фазовую частотную характеристику (АФЧХ).

Для нелинейного объекта при подаче на его вход пробного сигнала вида (16) выходная координата будет иметь вид

$$Y_H(t) = f_H(X(t), \dot{X}(t)), \quad (18)$$

где f_H - функция нелинейности объекта. Первая гармоника выходного сигнала в этом случае имеет вид

$$Y_{H1}(t) = q(A, \omega_i) \cdot \sin \omega_i t + q'(A, \omega_i) \cdot \cos \omega_i t, \quad (19)$$

где

$$q(A, \omega_i) = \frac{1}{A \cdot \pi} \int_0^{2\pi} F_H(A \cdot \sin \psi, A \cdot \omega_i \cdot \cos \psi) \sin \psi d\psi - \quad (20)$$

- коэффициент гармонической линейризации по синфазной составляющей,

$$q'(A, \omega_i) = \frac{1}{A \cdot \pi} \int_0^{2\pi} F_H(A \cdot \sin \psi, A \cdot \omega_i \cdot \cos \psi) \cos \psi d\psi - \quad (21)$$

- коэффициент гармонической линейризации по квадратурной составляющей.

Первая гармоника (19) может быть представлена в следующем виде

$$Y_{H1}(t) = A_H \cdot \sin(\omega_i t + \phi_\varepsilon), \quad (22)$$

где

$$\begin{aligned} A_H &= R_3(A, \omega_i) \cdot A; \\ R_3(A, \omega_i) &= \sqrt{q^2(A, \omega_i) + q'^2(A, \omega_i)}; \\ \phi_\varepsilon &= \arctg \frac{q'(A, \omega_i)}{q(A, \omega_i)}; \end{aligned} \quad (23)$$

$R_3(A, \omega_i)$ - эквивалентный коэффициент передачи,

$\phi_\varepsilon(A, \omega_i)$ - эквивалентный фазовый сдвиг.

Таким образом, частотная характеристика (15) для нелинейного объекта приобретает вид эквивалентного комплексного коэффициента передачи

$$W_3(j\omega, A) = R_3(A, \omega) \cdot e^{-j\phi_3(A, \omega)}, \quad (24)$$

где $R_3(A, \omega)$ - эквивалентная амплитудная частотная характеристика;

$\phi_3(A, \omega)$ - эквивалентная фазовая частотная характеристика.

Методика экспериментального определения частотной характеристики для нелинейного объекта, следующая:

- для фиксированной частоты выделяется первая гармоника выходного сигнала, по амплитудам и фазовым сдвигам входного сигнала и первой гармоники вычисляются эквивалентные коэффициент передачи и фазовый сдвиг (23);

- выбирается несколько фиксированных частот и для каждой из них проводится эксперимент по первому пункту, в результате чего определяется частотная характеристика (24) для фиксированного значения амплитуды A входного сигнала;

- выбирается несколько значений амплитуд входного сигнала и для каждого из них проводится эксперимент по второму пункту, в результате чего определяется частотная характеристика (24) для разных значений амплитуды A входного сигнала.

В результате эксперимента получается семейство эквивалентных частотных характеристик нелинейного объекта.

В лабораторной работе для определения частотных характеристик объекта используется модель в Matlab Simulink, представленная на рисунке 14.

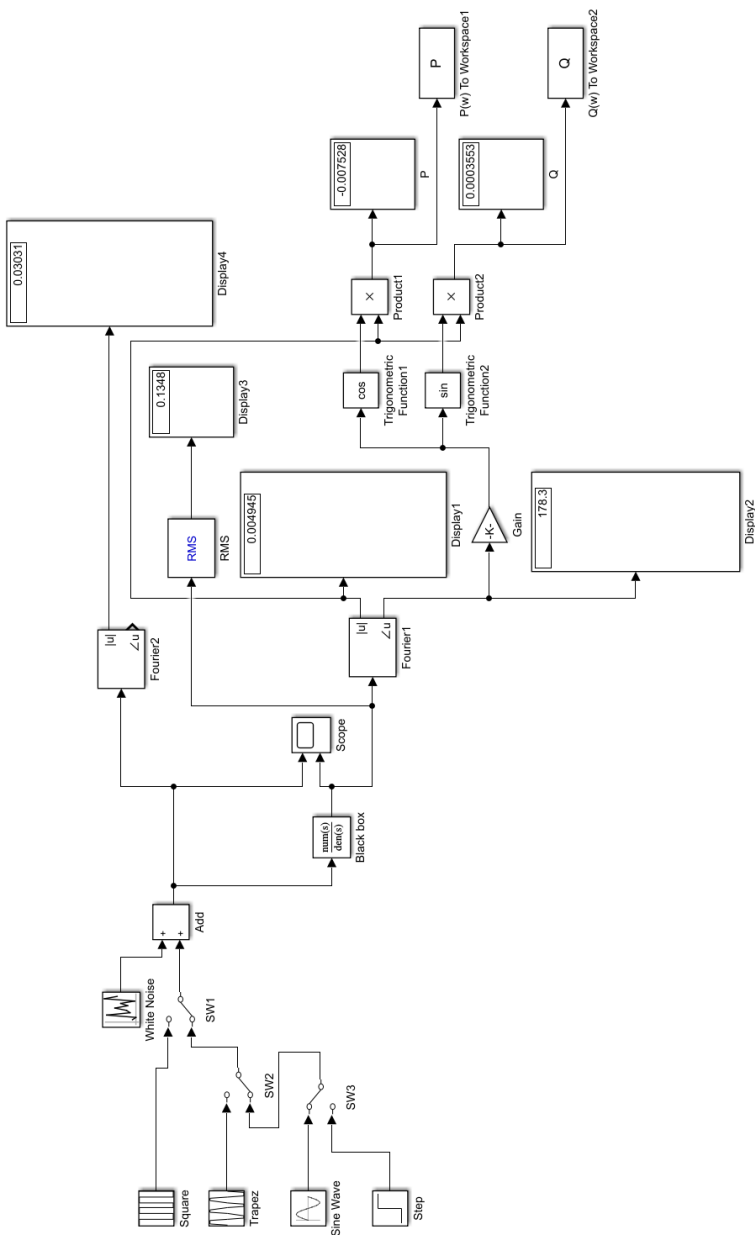


Рисунок 14 – Модель в Simulink для получения экспериментальных частотных характеристик

Блоки Square, Trapez, Sine Wave, Step – источники прямоугольного, трапецеидального, синусоидального и ступенчатого пробных сигналов. SW1, SW2, SW3 – ключи, с помощью которых на вход модели подается один из видов пробных сигналов. На рисунке 14 ключи находятся в положении, в котором на вход модели подается синусоидальный пробный сигнал. Блок White Noise добавляет к пробному сигналу белый шум с помощью блока Add, имитируя помехи, присутствующие в экспериментах с реальным объектом. Исследуемый объект представлен блоком Black Box («черный ящик»), который содержит параметры передаточной функции, заданной преподавателем, для которой необходимо экспериментально определить частотные характеристики. Входной пробный сигнал и отклик объекта на этот сигнал во времени отображаются с помощью блока Scope. Блок Fourier2 рассчитывает амплитуду основной гармоники входного пробного сигнала, которая отображается в блоке Display4. Блок Fourier1 выполняет расчет амплитуды и фазы первой гармоники отклика исследуемого объекта на входной пробный сигнал и отображения их в блоках Display1 и Display2 соответственно. Блок RMS измеряет действующее значение отклика объекта для отображения его в блоке Display1. Блок Gain преобразует значение фазы первой гармоники отклика объекта из градусов в радианы ($K=\pi/180$). С помощью блоков Trigonometric Function1, Product1, Trigonometric Function2, Product2 выполняется расчет действительной P и мнимой Q частей амплитудной частотной характеристики исследуемого объекта, которые отображаются в блоках P и Q соответственно. Блоки $P(w)$ To Workspace1 и $Q(w)$ To Workspace2 предназначены для передачи вещественной частотной характеристики в Workspace пакета Matlab для

последующей обработки (например, построения логарифмических частотных характеристик).

Пример 7. Экспериментальное определение частотных характеристик объекта типа «черный ящик»

Пусть объект задан передаточной функцией вида:

$$W(s) = \frac{0,3 \cdot s^2 + 0,5 \cdot s + 2}{1,2 \cdot s^2 + 0,6 \cdot s + 1}.$$

Проведение измерений откликов объекта на пробное синусоидальное воздействие переменной частоты выполним с помощью модели, представленной на рисунке 14, предварительно введя в настройки блока Black Box коэффициенты полиномов числителя и знаменателя передаточной функции исследуемого объекта управления. Для автоматизации измерений и построения ЛАЧХ используем следующий скрипт в Matlab:

```
a = [1.2 0.6 1]; % знаменатель передаточной функции  
исследуемого объекта
```

```
b = [0.3 0.5 2]; % числитель передаточной функции  
исследуемого объекта
```

```
w = logspace(-1,1); % задание частоты
```

```
h = freqs(b,a,w); % создание передаточной функции  
исследуемого объекта
```

```
mag = abs(h);
```

```
phase = angle(h);
```

```
mag = 20*log10(mag); % расчет значений ЛАЧХ  
исследуемого объекта
```

```
phase = phase*180/pi; % расчет значений ЛФЧХ  
исследуемого объекта
```

```
subplot(2,2,1), semilogx(w,mag);
```

```
subplot(2,2,2), semilogx(w,phase);
```

```
open('Lab_ExFrCh1.slx'); % название модели в Simulink
```

```

Pe = [];
Qe = [];
W1 = [];
W=0;
for i = 1:19
    if i<=10
        W=W+0.1 % задание частоты от 0.1 до 1 с шагом 0.1
    else W=W+1 % задание частоты от 1 до 10 с шагом 1
    end;
    set_param('Lab_ExFrCh1/Sine Wave','Frequency','W'); %
задание частоты источника пробного синусоидального сигнала
    sim('Lab_ExFrCh1'); % запуск моделирования
    Pe = [Pe ans.P.Data]; % запись в WorkSpace значения
действительной части амплитудно-частотной характеристики
    Qe = [Qe ans.Q.Data]; % запись в WorkSpace значения
мнимой части амплитудно-частотной характеристики
    W1 = [W1 W]; % запись в WorkSpace текущего
значения частоты пробного синусоидального сигнала
end;
    A=20*log10(sqrt(Pe.^2+Qe.^2)); % расчет
экспериментальной ЛАЧХ исследуемого объекта
    for i=1:19
        F(i)=180/pi*angle(Qe(i)*sqrt(-1)+Pe(i)); % расчет
экспериментальной ЛФЧХ исследуемого объекта
    end;
    subplot(2,2,3), semilogx(W1,A);
    subplot(2,2,4), semilogx(W1,F);

```

Результаты измерений и расчетов представлены в таблице 12, а ЛАЧХ и ЛФЧХ теоретические и экспериментальные представлены на рисунке 15.

Таблица 12 – Результаты измерений и расчётов

w, рад/с	0,100	0,200	0,300	0,400	0,500	0,600	0,700	0,800	0,900	1,000
lg(w)	-1,000	-0,699	-0,523	-0,398	-0,301	-0,222	-0,155	-0,097	-0,046	0,000
A, о.е.	2,019	2,069	2,179	2,315	2,559	2,838	3,217	3,461	3,366	2,791
Ф, град	6,102	6,317	6,764	7,290	8,160	9,059	10,149	10,784	10,542	8,914
L, дБ	-2,115	-4,228	-7,177	-10,701	-15,689	-23,363	-34,883	-51,709	-72,679	-92,037
w, рад/с	2,000	3,000	4,000	5,000	6,000	7,000	8,000	9,000	10,000	-
lg(w)	0,301	0,477	0,602	0,699	0,778	0,845	0,903	0,954	1,000	-
A, о.е.	0,329	0,171	0,189	0,205	0,217	0,228	0,232	0,238	0,239	-
Ф, град	-9,651	-15,327	-14,483	-13,744	-13,258	-12,857	-12,677	-12,461	-12,417	-
L, дБ	-111,114	-54,115	-27,488	-18,085	-14,313	-11,712	-9,713	-8,280	-7,172	-

Выводы о структуре и параметрах объекта управления на основе полученных частотных характеристик:

- объект имеет статический коэффициент усиления

$$k = 10^{\frac{6,1}{20}} = 2,0184;$$

- порядок числителя объекта равен порядку знаменателя, и равен 2, т.е. исследуемый объект – 2 порядка;

- частота среза объекта $w_{cp} \approx 1,2$.

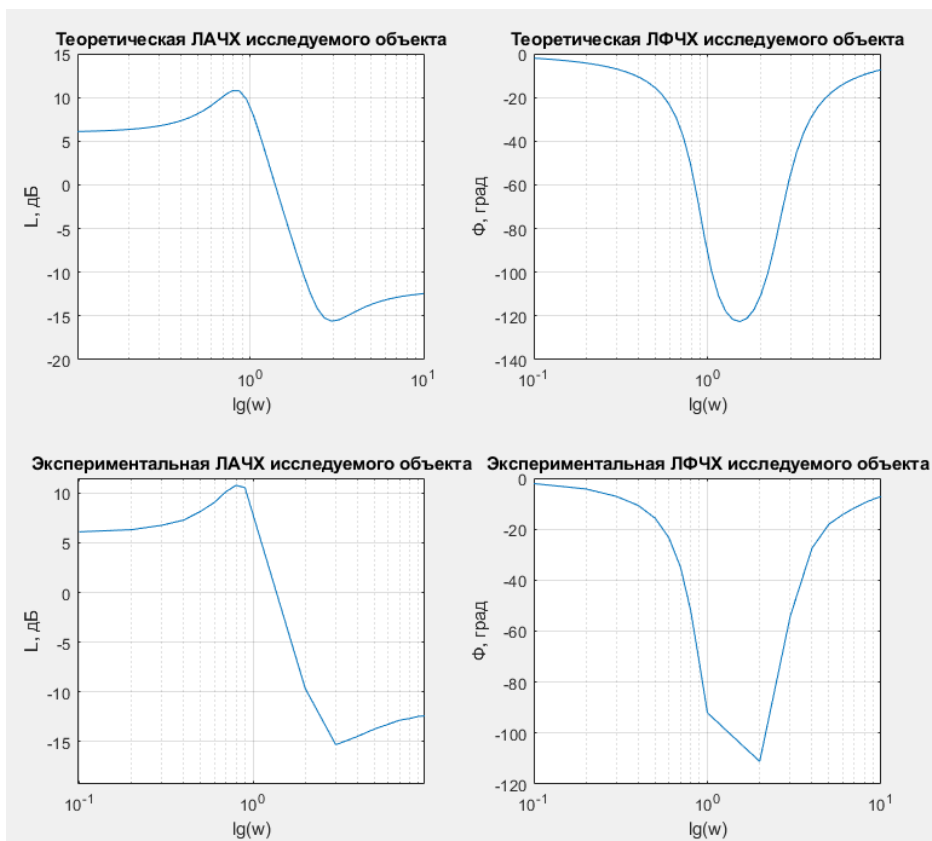


Рисунок 15 – ЛАЧХ и ЛФЧХ теоретические и экспериментальные

Задание на лабораторную работу 5

Для объекта с «неизвестной передаточной» функцией (желательно, чтобы параметры передаточной функции объекта в блок Black Box вводились преподавателем или ассистентом без участия обучающихся чтобы имитировать неопределённость в реальных условиях эксперимента) построить экспериментальные логарифмические частотные характеристики: логарифмическую амплитудно-частотную (ЛАЧХ) и логарифмическую фазо-частотную (ЛФЧХ). По

полученным характеристикам сделать выводы о статических и динамических характеристиках исследуемого объекта и выдвинуть предположения о структуре его передаточной функции. Передаточная функция исследуемого объекта задается преподавателем по номеру варианта из таблицы 14.

При выполнении работы рекомендуется амплитуду входного пробного синусоидального сигнала для всех частот сделать равной 1. Тогда блок Display1 будет сразу отображать значение амплитудной характеристики объекта для заданной частоты. Кроме того, рекомендуется снять характеристики не менее, чем для 20 значений частоты входного сигнала. Для этого в блоке Sine Wave необходимо менять параметр Frequency (rad/sec) последовательно от 0.1 до 1 с шагом 0.1 и от 1 до 10 с шагом 1. Можно также менять частоту и в других диапазонах, учитывая логарифмический масштаб оси частот, например, в диапазоне от 10 до 100 с шагом 10 или диапазоне от 100 до 1000 с шагом 100. Для каждого пробного синусоидального сигнала с заданной частотой необходимо получить и зафиксировать (в файл или в бумажную таблицу) значения амплитуды и фазы отклика объекта (Display1 и Display2 соответственно). По полученным значениям построить логарифмические частотные характеристики объекта.

Содержание отчета:

1. Номер передаточной функции исследуемого объекта.
2. Таблица 13 значений амплитуды A и фазы Φ отклика объекта для каждой частоты ω пробного синусоидального сигнала, а также значения $\lg(\omega)$ и $L=20 \cdot \lg(A)$ для построения ЛАЧХ и ФЧХ.

Таблица 13 – Результаты измерений и расчётов

№ п/п	ω , рад/с	$\lg(\omega)$	A , о.е.	Φ , град	L , дБ
-------	------------------	---------------	------------	---------------	----------

--	--	--	--	--	--

3. Полученные графики ЛАЧХ и ФЧХ исследуемого объекта.

4. Выводы о статических и динамических характеристиках исследуемого объекта и предположения о структуре его передаточной функции.

Таблица 14 – Варианты заданий для выполнения лабораторной работы 5

№	Передаточная функция	№	Передаточная функция
1	$W(s) = \frac{0,2 \cdot s^2 + 0,3 \cdot s + 1}{s^2 + 0,4 \cdot s + 1}$	6	$W(s) = \frac{0,9 \cdot s^2 + 0,7 \cdot s + 2,4}{1,2 \cdot s^2 + 0,8 \cdot s + 1}$
2	$W(s) = \frac{0,3 \cdot s^2 + 0,1 \cdot s + 1,4}{2 \cdot s^2 + 0,5 \cdot s + 1}$	7	$W(s) = \frac{s^2 + 0,6 \cdot s + 1,2}{2 \cdot s^2 + 0,6 \cdot s + 1}$
3	$W(s) = \frac{0,1 \cdot s^2 + 0,4 \cdot s + 2}{1,5 \cdot s^2 + 0,3 \cdot s + 1}$	8	$W(s) = \frac{0,7 \cdot s^2 + 0,5 \cdot s + 2,8}{0,8 \cdot s^2 + 0,4 \cdot s + 1}$
4	$W(s) = \frac{0,4 \cdot s^2 + 0,2 \cdot s + 0,5}{0,6 \cdot s^2 + 0,2 \cdot s + 1}$	9	$W(s) = \frac{0,6 \cdot s^2 + 0,8 \cdot s + 0,9}{0,4 \cdot s^2 + 0,7 \cdot s + 1}$
5	$W(s) = \frac{0,5 \cdot s^2 + 0,3 \cdot s + 1,6}{1,8 \cdot s^2 + 0,1 \cdot s + 1}$	10	$W(s) = \frac{0,8 \cdot s^2 + 0,9 \cdot s + 2,4}{0,9 \cdot s^2 + 0,5 \cdot s + 1}$

Лабораторная работа 6

Определение параметров передаточной функции объекта по экспериментальным частотным характеристикам аналитическим методом

Теоретическая часть

Аналитический метод аппроксимации. Пусть имеется N значений экспериментальной АФЧХ $A_3(j\omega)$. Можно предположить, что передаточная функция и соответствующая ей АФЧХ исследуемого объекта имеют вид

$$A(p) = \frac{\sum_{i=0}^m b_i \cdot p^i}{\sum_{j=0}^n a_j \cdot p^j}$$

или

$$A(j\omega) = \frac{P_1(\omega) + j \cdot Q_1(\omega)}{P_2(\omega) + j \cdot Q_2(\omega)}, \quad (24)$$

где m и n – порядок полиномов числителя и знаменателя соответственно (для физически реализуемых объектов $m \leq n$);

b_i и a_j коэффициенты полиномов числителя и знаменателя при элементах i -го и j -го порядка соответственно;

p – оператор Лапласа;

ω – частота;

$P_1(\omega)$, $P_2(\omega)$ – действительные части полиномов числителя и знаменателя АФЧХ соответственно;

$Q_1(\omega)$, $Q_2(\omega)$ – мнимые части полиномов числителя и знаменателя АФЧХ соответственно.

Экспериментальную частотную характеристику также можно представить в виде суммы действительной и мнимой частей

$$A_3(j\omega) = P_3(\omega) + j \cdot Q_3(\omega), \quad (25)$$

где

$$P_3(\omega) = |A_3(j\omega)| \cdot \cos(\arg A_3(j\omega)),$$

$$Q_3(\omega) = |A_3(j\omega)| \cdot \sin(\arg A_3(j\omega)). \quad (26)$$

Приравняв отдельно действительные и мнимые части выражений (24) и (25) для N известных точек, можно получить систему из $2 \cdot N$ уравнений, содержащую искомые коэффициенты a_i и b_i :

$$\begin{aligned} P_2(\omega_k) \cdot P_3(\omega_k) - Q_2(\omega_k) \cdot Q_3(\omega_k) - P_1(\omega_k) &= 0, \\ Q_2(\omega_k) \cdot P_3(\omega_k) + P_2(\omega_k) \cdot Q_3(\omega_k) - Q_1(\omega_k) &= 0, k = 1, 2, \dots, N \end{aligned} \quad (27)$$

Для решения системы уравнений (27) необходимо задать значения n и m . Для этого можно воспользоваться следующими соображениями:

1. если крайняя высокочастотная асимптота экспериментальной характеристики имеет наклон $-20 \cdot r$ дБ/дек, то $n - m = r$;

2. если объект соответствует замкнутой системе или разомкнутой статической системе, то $\frac{b_0}{a_0} = \Phi_3(0)$;

3. Если объект соответствует разомкнутой астатической системе и крайняя низкочастотная асимптота экспериментальной характеристики имеет наклон $-20 \cdot v$ дБ/дек, то $a_0 = a_1 = \dots = a_v = 0, b_0 = 1$.

Таким образом, значения n и m можно определить как:

1. Для замкнутой системы или разомкнутой статической системы

$$\begin{aligned} n &= \frac{2 \cdot N + r}{2}, m = \frac{2 \cdot N - r}{2}, \text{ если } r - \text{ четное и } n = \frac{2 \cdot N - 1 + r}{2}, \\ m &= \frac{2 \cdot N - 1 - r}{2}, \text{ если } r - \text{ нечетное;} \end{aligned}$$

2. Для разомкнутой астатической системы

$$\begin{aligned} n &= \frac{2 \cdot N + r + v}{2}, m = \frac{2 \cdot N - r + v}{2}, \text{ если } (v - r) - \text{ четное и } \\ n &= \frac{2 \cdot N - 1 + r + v}{2}, m = \frac{2 \cdot N - 1 - r + v}{2}, \text{ если } (v - r) - \text{ нечетное.} \end{aligned}$$

Определить коэффициенты передаточной функции объекта по экспериментальной АФЧХ можно двумя методами:

1. Графическим методом аппроксимации. Суть метода в отображении экспериментальных данных в виде логарифмической АФЧХ и замены ее (аппроксимации) асимптотами, соответствующими элементарным звеньям (интегратору, апериодическому и колебательному звеньям, дифференцирующим звеньям 1-го и 2-го порядка), включая неминимально-фазовые звенья. Очевидно, что с инженерных позиций это наиболее удобный метод, но он весьма груб.

2. Методом наименьших квадратов. Суть метода состоит в уточнении коэффициентов передаточной функции объекта, предварительно полученной с помощью одного из рассмотренных выше методов.

Пусть аппроксимированная передаточная функция объекта имеет вид (24), а поправки к ее первоначальным коэффициентам обозначаются, как Δa_i и Δb_i . Согласно методу наименьших квадратов аппроксимируемая АФЧХ $A(j\omega)$ должна удовлетворять следующему условию:

$$E = \sum_{k=1}^N |A(j\omega_k) - A_{\text{э}}(j\omega_k)|^2 \rightarrow \min,$$

$$2 \cdot N > n + m + 1. \quad (28)$$

Условие (28) эквивалентно следующему:

$$E = \sum_{k=1}^N \{(P(\omega_k) - P_{\text{э}}(\omega_k))^2 + (Q(\omega_k) - Q_{\text{э}}(\omega_k))^2\} \rightarrow \min. \quad (29)$$

Тогда, разложив в выражении (29) действительную $P(\omega_k)$ и мнимую $Q(\omega_k)$ части в кратный ряд Тейлора относительно всех коэффициентов a_i и b_i , и оставив в разложении только члены, содержащие первые разности (т.е. искомые поправки Δa_i и Δb_i), можно получить следующее выражение для E :

$$\begin{aligned}
E \approx \sum_{k=1}^N \left\{ \left(\sum_{i=0}^m \left(\frac{\partial P(\omega_k)}{\partial b_i} \right)_0 \cdot \Delta b_i \right. \right. \\
+ \sum_{i=0}^n \left(\frac{\partial P(\omega_k)}{\partial a_i} \right)_0 \cdot \Delta a_i - \Delta P(\omega_k) \Big)^2 + \\
+ \left(\sum_{i=0}^m \left(\frac{\partial Q(\omega_k)}{\partial b_i} \right)_0 \cdot \Delta b_i \right. \\
+ \sum_{i=0}^n \left(\frac{\partial Q(\omega_k)}{\partial a_i} \right)_0 \cdot \Delta a_i - \Delta Q(\omega_k) \Big)^2 \Big\}, \quad (30)
\end{aligned}$$

где $\Delta P(\omega_k) = P_3(\omega_k) - P_0(\omega_k)$ и $\Delta Q(\omega_k) = Q_3(\omega_k) - Q_0(\omega_k)$.

Нулевые индексы при действительной $P(\omega_k)$ части и мнимой $Q(\omega_k)$ части в выражении (30) соответствуют их вычислению со значениями коэффициентов a_i и b_i , полученными в результате предварительной аппроксимации.

Очевидно, что необходимым и достаточным условием достижения минимума (28) является

$$\begin{aligned}
\frac{\partial E}{\partial a_0} = \frac{\partial E}{\partial a_1} = \dots = \frac{\partial E}{\partial a_k} = \frac{\partial E}{\partial b_0} = \frac{\partial E}{\partial b_1} = \dots = \frac{\partial E}{\partial b_k} \\
= 0. \quad (31)
\end{aligned}$$

Тогда из (30) и (31) можно получить расчетные уравнения для вычисления поправок к коэффициентам

$$\begin{aligned}
\sum_{k=1}^N \left\{ \left(\sum_{i=0}^m \left(\frac{\partial P(\omega_k)}{\partial b_i} \right)_0 \cdot \Delta b_i + \sum_{i=0}^n \left(\frac{\partial P(\omega_k)}{\partial a_i} \right)_0 \cdot \Delta a_i - \Delta P(\omega_k) \right) \right. \\
\cdot \left(\frac{\partial P(\omega_k)}{\partial b_g} \right)_0 +
\end{aligned}$$

$$+ \left(\sum_{i=0}^m \left(\frac{\partial Q(\omega_k)}{\partial b_i} \right)_0 \cdot \Delta b_i + \sum_{i=0}^n \left(\frac{\partial Q(\omega_k)}{\partial a_i} \right)_0 \cdot \Delta a_i - \Delta Q(\omega_k) \right) \cdot \left(\frac{\partial Q(\omega_k)}{\partial b_g} \right)_0 \Bigg\} = 0,$$

$$g = 0, 1, \dots, m; \quad (32)$$

$$\sum_{k=1}^N \left\{ \left(\sum_{i=0}^m \left(\frac{\partial P(\omega_k)}{\partial b_i} \right)_0 \cdot \Delta b_i + \sum_{i=0}^n \left(\frac{\partial P(\omega_k)}{\partial a_i} \right)_0 \cdot \Delta a_i - \Delta P(\omega_k) \right) \cdot \left(\frac{\partial P(\omega_k)}{\partial a_h} \right)_0 + \right. \\ \left. + \left(\sum_{i=0}^m \left(\frac{\partial Q(\omega_k)}{\partial b_i} \right)_0 \cdot \Delta b_i + \sum_{i=0}^n \left(\frac{\partial Q(\omega_k)}{\partial a_i} \right)_0 \cdot \Delta a_i - \Delta Q(\omega_k) \right) \cdot \left(\frac{\partial Q(\omega_k)}{\partial a_h} \right)_0 \right\} = 0,$$

$$h = 0, 1, \dots, n. \quad (33)$$

Окончательно, необходимо оценить величины поправок Δa_i и Δb_i относительно исходных коэффициентов a_i и b_i . Если эти отношения значительны, то необходимо произвести новую предварительную аппроксимацию АФЧХ $A(j\omega)$ аналитическим или графическим методами.

Пример 8. Определение структуры и коэффициентов передаточной функции объекта управления по экспериментальным значениям действительной и мнимой частей амплитудно-частотной характеристики исследуемого объекта.

Пусть имеются экспериментально полученные значения действительной и мнимой частей амплитудно-частотной характеристики исследуемого объекта (таблица 15).

Таблица 15 - Экспериментальные значения действительной и мнимой частей амплитудно-частотной характеристики исследуемого объекта

w, рад/с	0,100	0,200	0,300	0,400	0,500	0,600	0,700	0,800	0,900	1,000
P(w)	2,018	2,064	2,162	2,274	2,463	2,605	2,639	2,145	1,002	- 0,099
Q(w)	- 0,074	- 0,153	- 0,272	- 0,430	- 0,692	- 1,125	- 1,840	- 2,717	-3,213	- 2,789
w, рад/с	2,000	3,000	4,000	5,000	6,000	7,000	8,000	9,000	10,000	-
P(w)	- 0,119	0,100	0,167	0,195	0,211	0,223	0,229	0,236	0,238	-
Q(w)	- 0,307	- 0,139	- 0,087	- 0,064	- 0,054	- 0,046	- 0,039	- 0,034	-0,030	-

Строим логарифмические частотные характеристики объекта в Matlab, используя следующий скрипт:

```
w = [0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1 2 3 4 5 6 7 8 9 10]
P = [2.018 2.064 2.162 2.274 2.463 2.605 2.639 2.145 1.002 -0.099
-0.119 0.100 0.167 0.195 0.211 0.223 0.229 0.236 0.238]
Q = [-0.074 -0.153 -0.272 -0.430 -0.692 -1.125 -1.840 -2.717 -
3.213 -2.789 -0.307 -0.139 -0.087 -0.064 -0.054 -0.046 -0.039 -
0.034 -0.030]
phase = atan2(Q,P);
mag = 20*log10(sqrt(P.^2+Q.^2));
phase = phase*180/pi;
subplot(1,2,1), semilogx(w,mag);
subplot(1,2,2), semilogx(w,phase);
```

ЛАЧХ и ЛФЧХ исследуемого объекта изображены на рисунке 16. Проведём анализ свойств исследуемого объекта для определения его структуры и некоторых параметров передаточной функции:

- объект имеет статический коэффициент усиления

$$k = 10^{\frac{6,1}{20}} = 2,0184;$$

- порядок числителя объекта равен порядку знаменателя, и равен 2, т.е. исследуемый объект – 2 порядка.

Таким образом объект имеет передаточную функцию вида:

$$W(s) = \frac{b_2 \cdot s^2 + b_1 \cdot s + b_0}{a_2 \cdot s^2 + a_1 \cdot s + a_0}.$$

Будем искать коэффициенты передаточной функции в канонической форме, тогда

$b_0=k=2,0184$, $a_0=1$. Таким образом необходимо определить коэффициенты b_2 , b_1 , a_2 , a_1 .

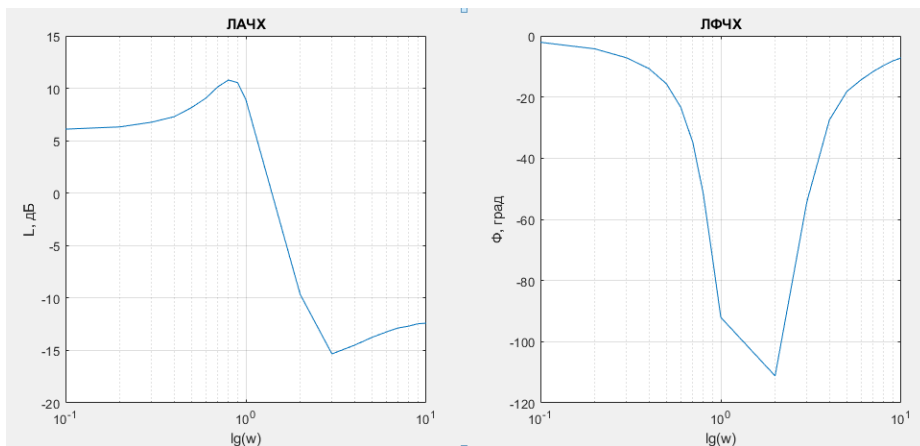


Рисунок 16 – ЛАЧХ и ФЧХ исследуемого объекта

Находим первые приближения для коэффициентов передаточной функции исследуемого объекта с помощью следующего скрипта:

```

for i=1:19
PQ(2*i-1,1)=Pe(i); PQ(2*i-1,2)=-Pe(i)*W1(i)^2; PQ(2*i-1,3)=-W1(i)*Qe(i);
PQ(2*i-1,4)=-1; PQ(2*i-1,5)=W1(i)^2; PQ(2*i-1,6)=0;
PQ(2*i,1)=Qe(i); PQ(2*i,2)=-Qe(i)*W1(i)^2; PQ(2*i,3)=W1(i)*Pe(i);
PQ(2*i,4)=0; PQ(2*i,5)=0; PQ(2*i,6)=-W1(i);
end;
B=-1*PQ(:,1);
PQ1=PQ(:,2:6);
X=pinv(PQ1)*B;

```

Далее составляем и решаем систему уравнений (32) и (33) для нахождения поправок Δa_i и Δb_i , используя следующий скрипт:

```

syms P(a0,a1,a2,b0,b1,b2,w)
P(a0,a1,a2,b0,b1,b2,w)=(b1*a1*w^2+(b0-b2*w^2)*(a0-
a2*w^2))/(a1^2*w^2+(a0-a2*w^2)^2); % аналитическое выражение для
действительной части передаточной функции
Pa0 = diff(P,a0); % частные производные действительной части
Pa1 = diff(P,a1); % передаточной функции по коэффициентам
Pa2 = diff(P,a2);
Pb0 = diff(P,b0);
Pb1 = diff(P,b1);
Pb2 = diff(P,b2);
syms Q(a0,a1,a2,b0,b1,b2,w)
Q(a0,a1,a2,b0,b1,b2,w)=(b1*a0*w-b1*a2*w^3-
a1*b0*w+a1*b2*w^3)/(a1^2*w^2+(a0-a2*w^2)^2); % аналитическое
выражение для мнимой части передаточной функции
Qa0 = diff(Q,a0); % частные производные мнимой части
Qa1 = diff(Q,a1); % передаточной функции по коэффициентам
Qa2 = diff(Q,a2);
Qb0 = diff(Q,b0);
Qb1 = diff(Q,b1);
Qb2 = diff(Q,b2);
PQ2(1:6,1:6)=0;
B1(1:6,1)=0;
for i=1:19 % построение системы уравнений (32) и (33)

```


$PQ2(1,1)=PQ2(1,1)+Pb0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pb0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qb0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qb0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));$
 $PQ2(1,2)=PQ2(1,2)+Pb1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pb0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qb1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qb0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));$
 $PQ2(1,3)=PQ2(1,3)+Pb2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pb0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qb2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qb0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));$
 $PQ2(1,4)=PQ2(1,4)+Pa0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pb0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qa0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qb0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));$
 $PQ2(1,5)=PQ2(1,5)+Pa1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pb0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qa1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qb0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));$
 $PQ2(1,6)=PQ2(1,6)+Pa2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pb0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qa2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qb0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));$
 $B1(1,1)=B1(1,1)+(Pe(i)-P(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i)))*Pb0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+(Qe(i)-Q(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i)))*Qb0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));$
 end;

for i=1:19

$PQ2(2,1)=PQ2(2,1)+Pb0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pb1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qb0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qb1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));$
 $PQ2(2,2)=PQ2(2,2)+Pb1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pb1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qb1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qb1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));$
 $PQ2(2,3)=PQ2(2,3)+Pb2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pb1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qb2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qb1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));$
 $PQ2(2,4)=PQ2(2,4)+Pa0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pb1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qa0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qb1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));$

$PQ2(2,5)=PQ2(2,5)+Pa1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pb1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qa1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qb1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));$
 $PQ2(2,6)=PQ2(2,6)+Pa2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pb1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qa2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qb1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));$
 $B1(2,1)=B1(2,1)+(Pe(i)-P(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i)))*Pb1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+(Qe(i)-Q(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i)))*Qb1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));$
 end;

for i=1:19

$PQ2(3,1)=PQ2(3,1)+Pb0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pb2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qb0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qb2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));$
 $PQ2(3,2)=PQ2(3,2)+Pb1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pb2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qb1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qb2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));$
 $PQ2(3,3)=PQ2(3,3)+Pb2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pb2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qb2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qb2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));$
 $PQ2(3,4)=PQ2(3,4)+Pa0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pb2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qa0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qb2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));$
 $PQ2(3,5)=PQ2(3,5)+Pa1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pb2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qa1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qb2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));$
 $PQ2(3,6)=PQ2(3,6)+Pa2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pb2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qa2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qb2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));$
 $B1(3,1)=B1(3,1)+(Pe(i)-P(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i)))*Pb2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+(Qe(i)-Q(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i)))*Qb2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));$
 end;

for i=1:19

PQ2(4,1)=PQ2(4,1)+Pb0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pa0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qb0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qa0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));
PQ2(4,2)=PQ2(4,2)+Pb1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pa0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qb1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qa0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));
PQ2(4,3)=PQ2(4,3)+Pb2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pa0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qb2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qa0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));
PQ2(4,4)=PQ2(4,4)+Pa0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pa0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qa0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qa0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));
PQ2(4,5)=PQ2(4,5)+Pa1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pa0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qa1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qa0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));
PQ2(4,6)=PQ2(4,6)+Pa2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pa0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qa2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qa0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));
B1(4,1)=B1(4,1)+(Pe(i)-P(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i)))*Pa0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+(Qe(i)-Q(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i)))*Qa0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));

end;

for i=1:19

PQ2(5,1)=PQ2(5,1)+Pb0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pa1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qb0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qa1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));
PQ2(5,2)=PQ2(5,2)+Pb1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pa1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qb1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qa1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));
PQ2(5,3)=PQ2(5,3)+Pb2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pa1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qb2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qa1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));
PQ2(5,4)=PQ2(5,4)+Pa0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pa1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qa0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qa1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));

$PQ2(5,5)=PQ2(5,5)+Pa1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pa1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qa1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qa1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));$
 $PQ2(5,6)=PQ2(5,6)+Pa2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pa1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qa2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qa1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));$
 $B1(5,1)=B1(5,1)+(Pe(i)-P(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i)))*Pa1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+(Qe(i)-Q(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i)))*Qa1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));$
 end;

for i=1:19

$PQ2(6,1)=PQ2(6,1)+Pb0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pa2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qb0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qa2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));$
 $PQ2(6,2)=PQ2(6,2)+Pb1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pa2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qb1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qa2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));$
 $PQ2(6,3)=PQ2(6,3)+Pb2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pa2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qb2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qa2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));$
 $PQ2(6,4)=PQ2(6,4)+Pa0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pa2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qa0(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qa2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));$
 $PQ2(6,5)=PQ2(6,5)+Pa1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pa2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qa1(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qa2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));$
 $PQ2(6,6)=PQ2(6,6)+Pa2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Pa2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+Qa2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))*Qa2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));$
 $B1(6,1)=B1(6,1)+(Pe(i)-P(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i)))*Pa2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i))+(Qe(i)-Q(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i)))*Qa2(1,X(2,1),X(1,1),X(3,1),X(5,1),X(4,1),W1(i));$
 end;

$X1 = \text{pinv}(PQ2) * B1$; % определение поправок

В результате выполнения скрипта получены:

- значения коэффициентов

$$a_0 = 1,000,$$

$$a_2 = 1,193,$$

$$a_1 = 0,604,$$

$$b_0 = 2,012,$$

$$b_2 = 0,299,$$

$$b_1 = 0,503;$$

- значения поправок

$$\Delta a_0 = -0,00842,$$

$$\Delta a_2 = 0,00218,$$

$$\Delta a_1 = 0,00155,$$

$$\Delta b_0 = 0,00237,$$

$$\Delta b_2 = -0,00182,$$

$$\Delta b_1 = 0,0118.$$

Так как поправки не превышают 5%, то передаточная функция объекта определена. С помощью следующего скрипта строим логарифмические частотные характеристики исследуемого объекта (рисунок 17) по найденной передаточной функции и сравниваем их с экспериментальными:

$$a = [1.193 \ 0.604 \ 1];$$

$$b = [0.299 \ 0.503 \ 2.012];$$

$$w = \text{logspace}(-1,1);$$

$$h = \text{freqs}(b,a,w);$$

$$\text{mag} = \text{abs}(h);$$

$$\text{phase} = \text{angle}(h);$$

$$\text{mag} = 20 * \log_{10}(\text{mag});$$

```

phase = phase*180/pi;
subplot(1,2,1), semilogx(w,mag);
subplot(1,2,2), semilogx(w,phase);

```

Частотные характеристики на рисунке 17, полученные для найденной передаточной функции исследуемого объекта, соответствуют экспериментальным, приведенным на рисунке 16. Таким образом, задача определения структуры передаточной функции и её параметров по экспериментально полученным вещественной и мнимой частотным характеристикам исследуемого объекта решена.

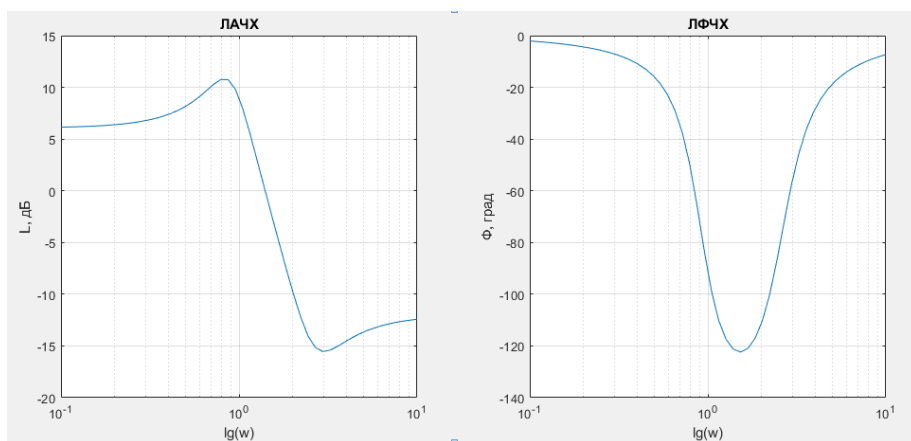


Рисунок 17 – Частотные характеристики исследуемого объекта по найденной передаточной функции

Задание на лабораторную работу

По экспериментально полученным значениям вещественной и мнимой частотных характеристик исследуемого объекта, выбираемых из таблицы 16 по варианту, заданному преподавателем, определить структуру и коэффициенты полиномов числителя и знаменателя

передаточной функции аналитическим методом аппроксимации по наименьшим квадратам.

Содержание отчёта:

1. Таблица экспериментально полученных значений вещественной и мнимой частотных характеристик исследуемого объекта.

2. Скрипты для построения логарифмических частотных характеристик и для определения коэффициентов полинома числителя и знаменателя передаточной функции и поправок к ним.

3. Найденные значения коэффициентов и поправок.

4. Результаты сравнения экспериментальных частотных характеристик с рассчитанными аналитическим методом аппроксимации и выводы.

Таблица 16 – Варианты заданий на лабораторную работу 6

№	Исходные данные										
1	w, рад/ с										
		0,100	0,200	0,300	0,400	0,500	0,600	0,700	0,800	0,900	1,000
	P(w)	1,008	1,029	1,076	1,133	1,238	1,358	1,539	1,678	1,552	0,748
	Q(w)	-	-	-	-	-	-	-	-	-	-
		0,01	0,02	0,04	0,07	0,12	0,22	0,43	0,82	-1,517	1,99
		2	2	5	4	8	8	3	6		3
2	w, рад/ с										-
		2,000	3,000	4,000	5,000	6,000	7,000	8,000	9,000	10,000	
	P(w)	-									-
		0,01									
		3	0,118	0,154	0,169	0,177	0,185	0,189	0,193	0,194	
	Q(w)	-	-	-	-	-	-	-	-		-
		0,20	0,09	0,06	0,04	0,04	0,03	0,03	0,02	-0,022	
		9	7	2	6	0	5	0	6		
	w, рад/ с										
		0,100	0,200	0,300	0,400	0,500	0,600	0,700	0,800	0,900	1,000
	P(w)								-		-
		1,423	1,489	1,629	1,833	2,173	2,243	0,412	1,28	-1,151	0,84
									9		0

	Q(w)	- 0,06 4	- 0,13 8	- 0,26 4	- 0,47 8	- 0,97 7	- 2,20 4	- 3,55 9	- 2,11 7	- -0,986	- 0,51 6
	w, рад/ с	2,000	3,000	4,000	5,000	6,000	7,000	8,000	9,000	10,000	-
	P(w)	- 0,02 6	0,079	0,111	0,124	0,132	0,138	0,141	0,143	0,144	-
	Q(w)	- 0,03 4	- 0,01 3	- 0,00 6	- 0,00 4	- 0,00 4	- 0,00 3	- 0,00 1	- 0,00 1	0,000	-
3	w, рад/ с	0,100	0,200	0,300	0,400	0,500	0,600	0,700	0,800	0,900	1,000
	P(w)	2,030	2,114	2,294	2,562	3,069	3,878	5,032	2,613	-2,657	- 2,42 9
	Q(w)	- 0,02 3	- 0,04 8	- 0,10 5	- 0,19 5	- 0,40 4	- 0,99 5	- 2,94 3	- 7,62 3	-5,008	- 2,25 8
	w, рад/ с	2,000	3,000	4,000	5,000	6,000	7,000	8,000	9,000	10,000	-
	P(w)	- 0,30 0	- 0,08 0	- 0,01 2	0,017	0,032	0,041	0,047	0,052	0,054	-
	Q(w)	- 0,19 5	- 0,10 5	- 0,07 1	- 0,05 3	- 0,04 4	- 0,03 7	- 0,03 3	- 0,02 9	-0,026	-
4	w, рад/ с	0,100	0,200	0,300	0,400	0,500	0,600	0,700	0,800	0,900	1,000
	P(w)	0,500	0,495	0,495	0,484	0,480	0,465	0,453	0,435	0,413	0,400
	Q(w)	0,010	0,021	0,031	0,045	0,062	0,082	0,108	0,147	0,206	0,298
	w, рад/ с	2,000	3,000	4,000	5,000	6,000	7,000	8,000	9,000	10,000	-
	P(w)	0,810	0,713	0,683	0,671	0,670	0,677	0,675	0,678	0,671	-
	Q(w)	- 0,05 2	- 0,03 0	- 0,02 2	- 0,02 1	- 0,02 6	- 0,02 3	- 0,01 5	- 0,01 0	-0,007	-
5	w, рад/ с	0,100	0,200	0,300	0,400	0,500	0,600	0,700	0,800	0,900	1,000
	P(w)	1,623	1,698	1,849	2,111	2,728	3,901	9,778	- 5,96	-2,515	- 1,19

									6		0
	Q(w)						- 0,38 0	- 3,29 0	- 4,13 0		- 0,45 5
		0,009	0,029	0,044	0,052	0,114				-1,265	
	w, рад/ с	2,000	3,000	4,000	5,000	6,000	7,000	8,000	9,000	10,000	-
	P(w)	0,069	0,192	0,228	0,243	0,253	0,262	0,266	0,271	0,272	-
6	Q(w)	- 0,08 2	- 0,05 1	- 0,03 6	- 0,02 9	- 0,02 8	- 0,02 5	- 0,02 1	- 0,01 8		- -0,015
	w, рад/ с	0,100	0,200	0,300	0,400	0,500	0,600	0,700	0,800	0,900	1,000
	P(w)	2,411	2,433	2,488	2,518	2,570	2,485	2,247	1,681	0,966	0,383
	Q(w)	- 0,12 7	- 0,25 9	- 0,43 8	- 0,65 2	- 0,96 1	- 1,36 6	- 1,86 0	- 2,22 9		- 1,96 2
	w, рад/ с	2,000	3,000	4,000	5,000	6,000	7,000	8,000	9,000	10,000	-
7	P(w)	0,403	0,604	0,663	0,689	0,707	0,726	0,731	0,741	0,738	-
	Q(w)	- 0,20 3	- 0,06 4	- 0,03 2	- 0,02 2	- 0,02 5	- 0,02 1	- 0,01 3	- 0,00 9		- -0,006
	w, рад/ с	0,100	0,200	0,300	0,400	0,500	0,600	0,700	0,800	0,900	1,000
	P(w)	1,214	1,253	1,340	1,462	1,670	1,746	1,084	0,235	0,075	0,117
	Q(w)	- 0,01 4	- 0,03 2	- 0,07 7	- 0,16 4	- 0,39 5	- 0,97 0	- 1,64 1	- 1,30 3		- 0,52 7
8	w, рад/ с	2,000	3,000	4,000	5,000	6,000	7,000	8,000	9,000	10,000	-
	P(w)	0,421	0,468	0,478	0,482	0,487	0,496	0,497	0,501	0,498	-
	Q(w)	- 0,10 0	- 0,05 3	- 0,03 6	- 0,03 0	- 0,03 1	- 0,02 6	- 0,02 0	- 0,01 6		- -0,012
	w, рад/ с	0,100	0,200	0,300	0,400	0,500	0,600	0,700	0,800	0,900	1,000
	P(w)	2,814	2,847	2,928	3,008	3,176	3,325	3,569	3,731	3,750	3,093
	Q(w)	- 0,06 7	- 0,12 8	- 0,22 3	- 0,32 5	- 0,47 5	- 0,70 2	- 1,06 7	- 1,62 7		- 3,68 7
	w,	2,000	3,000	4,000	5,000	6,000	7,000	8,000	9,000	10,000	-

	рад/ с										
	P(w)	0,147	0,599	0,721	0,770	0,800	0,828	0,839	0,854	0,854	-
	Q(w)	- 0,41 4	- 0,12 7	- 0,06 5	- 0,04 4	- 0,04 4	- 0,03 8	- 0,02 9	- 0,02 3	- -0,018	-
9	w, рад/ с	0,100	0,200	0,300	0,400	0,500	0,600	0,700	0,800	0,900	1,000
	P(w)	0,899	0,893	0,893	0,878	0,878	0,862	0,860	0,856	0,858	0,870
	Q(w)	0,016	0,036	0,054	0,078	0,105	0,137	0,171	0,217	0,266	0,314
	w, рад/ с	2,000	3,000	4,000	5,000	6,000	7,000	8,000	9,000	10,000	-
	P(w)	1,363	1,503	1,501	1,495	1,500	1,519	1,513	1,519	1,503	-
	Q(w)	0,501	0,302	0,200	0,143	0,098	0,085	0,085	0,081	0,076	-
10	w, рад/ с	0,100	0,200	0,300	0,400	0,500	0,600	0,700	0,800	0,900	1,000
	P(w)	2,413	2,444	2,521	2,597	2,753	2,893	3,100	3,192	3,049	2,340
	Q(w)	- 0,03 4	- 0,06 4	- 0,12 2	- 0,19 0	- 0,30 1	- 0,48 8	- 0,81 1	- 1,31 7	- -2,070	- 2,72 2
	w, рад/ с	2,000	3,000	4,000	5,000	6,000	7,000	8,000	9,000	10,000	-
	P(w)	0,505	0,733	0,795	0,820	0,839	0,861	0,868	0,880	0,877	-
	Q(w)	- 0,50 8	- 0,22 2	- 0,14 0	- 0,10 7	- 0,09 8	- 0,08 5	- 0,06 9	- 0,05 9	- -0,049	-

ЗАКЛЮЧЕНИЕ

Как можно было заметить из вышеперечисленных методов планирования эксперимента и обработки статистических данных, большинство из них сосредоточены на концепции однофакторного эксперимента [16].

Однофакторный, или классический, эксперимент базируется на допущении о том, что исследователь имеет возможность варьировать факторы, участвующие в исследовательской ситуации, по одному. Из этого следует, что экспериментатор способен выделить изучаемую зависимость в чистом виде, может чётко вычленять воздействующие на зависимые переменные факторы (может, скажем, как-то упорядочить их во времени и пространстве, «включать» и «выключать» их по своему усмотрению и т.п.). Однако на самом деле исследовательские ситуации часто оказываются гораздо более сложными.

Выход к более утончённой методологии, имеющей дело с комплексным, принципиально неразделимым действием факторов, был осуществлён прежде всего под влиянием работ английского учёного Рональда Фишера (1890-1962), посвящённых агробиологическим экспериментам в 1925г. В сложных системах факторы, воздействующие на изучаемый объект, действуют не изолированно и не независимо друг от друга, как это предполагала концепция классического эксперимента, а довольно сложным, взаимосвязанным способом. Они зачастую сцеплены между собой таким образом, что попытка варьировать одну независимую переменную автоматически приводит к некоему замысловатому изменению и других факторов. Это означает, что исследователю, приходится иметь дело с особой комплексной организацией этих факторов. Кроме того,

исследователя может интересовать действие не изолированных факторов, которое в реальности не встречается, а именно влияние различных возможных комбинаций факторов.

Идея многофакторного эксперимента (иногда используют упрощённое название факторный эксперимент) состоит в следующем. Исследователь может варьировать независимые переменные как комплекс, т.е. одновременно сразу несколько; после серии экспериментов полученные результаты должны быть подвергнуты специальному статистическому анализу, где каждый участвующий фактор будет оценён по результатам всех опытов данной серии. Используя соответствующие схемы и обрабатывая данные по особым статистическим методикам, позволяющим изучать эффективность совместного полифакторного воздействия (методики дисперсионного анализа), исследователь получает картину, отражающую вклад каждого фактора в изменяющихся условиях. В итоге экспериментатор имеет возможность изучать самые сложные комбинации факторов. Причём это осуществляется достаточно экономичным способом, т.к. информативность экспериментов зависит в данном случае не от их количества в серии, а от концептуальной организации исследований.

Многофакторный эксперимент - мощное средство современной науки. К его достоинствам относятся: эффективность использования времени и средств (ведь проведение ряда экспериментов с отдельными пофакторными модификациями требует значительных затрат), что выражается прежде всего в сокращении числа опытов, необходимых для решения исследовательской задачи; значительная информативность эксперимента (т.к. получаемый результат показывает удельный вес каждого фактора в их совокупном действии); высокая степень достоверности данных (в то время как при попытке использовать методологию классического

эксперимента результаты могут оказаться неудовлетворительными из-за воздействий неподконтрольных факторов).

Как показывает практика, часто исследователи в своих работах рассматривают многофакторный эксперимент как совокупность однофакторных [17]. Обработка данных в большинстве случаев затруднительна, поэтому фактически исследуемая система представляется посредством «черного ящика». Обработка данных осуществляется при помощи современных вычислительных средств и программных продуктов, в полной мере реализующих методы математического моделирования, описанные в данном пособии.

При составлении лабораторного практикума авторами использовались некоторые данные из источников [18-29].

СПИСОК ЛИТЕРАТУРЫ

1. Официальный русскоязычный сайт компании SPSS – <http://www.spss.com>, <http://www.predictivesolutions.ru/software/>
2. Тюрин Ю.Н., Макаров А.А. Анализ данных на компьютере / Под ред. В.Э. Фигурнова. – 3-е изд., перераб. и доп. – М.: ИНФРА-М, 2003. – 544
3. Боровиков, В.П. Прогнозирование в программе STATISTICA в среде Windows: Основы теории и интенсивная практика на компьютере: учеб. пособие / В.П. Боровиков, Г.И. Ивченко. – М.: Финансы и статистика, 2006. – 368 с.
4. Вуколов, Э.А. Основы статистического анализа. Практикум по статистическим методам и исследованию операций с использованием пакетов Statistica и Excel: учеб. пособие / Э.А. Вуколов. – М.: ИНФРА-М, 2004. – 462 с.
5. Кацко, И.А. Практикум по анализу данных на компьютере / И.А. Кацко, Н.Б. Паклин; под ред. Г.В. Гореловой. – М.: КолосС, 2009. – 278 с.
6. <http://www.statsoft.ru/> Электронный ресурс: Российское представительство компании StatSoft. Дата обращения: 03.10.2022.
7. <http://info.statgraphics.com/> Электронный ресурс: Сайт статистического пакета STATGRAPHICSPLUS. Дата обращения: 03.10.2022.
8. Орлова И.В. Экономико-математическое моделирование: практическое пособие по решению задач. – М.: Вузовский учебник, 2007. – 144 с.
9. Орлова, И.В. Экономико-математические методы и модели: компьютерное моделирование: учеб. пособие. – М.: Вузовский учебник, 2009. – 365 с.

10. <http://www.xlstat.com/> Электронный ресурс: Макросдополнение XLSTAT-Pro для MSExcel. Дата обращения: 03.10.2022.
11. <http://www.exponenta.ru/> Электронный ресурс: Образовательный математический сайт. Дата обращения: 03.10.2022.
12. <https://studfile.net/preview/3507786/> Электронный ресурс: Дата обращения: 03.10.2022.
13. <https://www.malavida.com/ru/soft/statgraphics/> Электронный ресурс: Statgraphics, программа для статистики. Дата обращения: 03.10.2022.
14. Реброва И.А. Планирование эксперимента: учебное пособие. – Омск: СиБАДИ, 2010. – 105 с.
15. Рогов В.А., Поздняк Г.Г. Методика и практика технических экспериментов: учеб. пособие для студ. высш. учеб. заведений /В.А. Рогов, Г.Г. Поздняк. – М.: Издательский центр «Академия», 2005. – 288 с.
16. <https://vuzlit.ru/817990> Электронный ресурс. Дата обращения: 03.10.2022.
17. В.Н. Шкляр. Планирование эксперимента и обработка результатов. Конспект лекций для магистров по направлению 220200 «Автоматизация и управление в технических (мехатронных) системах». Издательство Томского политехнического университета, 2010.
18. Лунев В. А. Математическое моделирование и планирование эксперимента: Учеб. пособие. СПб. – 153 с.
19. Лапач С.Н. Регрессионный анализ в условиях неоднородности факторного пространства / С.Н. Лапач, С.Г. Радченко // Математические машины и системы. – 2016. – № 3. – С. 55 – 63.
20. Кононюк А.Е. Основы научных исследований (общая теория эксперимента) - В 4-х кн. - К.1. - К.: 2011.- 508 с.

21. Кононюк спА.Е. Основы научных исследований (общая теория эксперимента) - В 4-х кн. - К.2. - К.: 2011.- 452 с.
22. Галямина И.Г. Управление процессами: учебник для бакалавр. и спец./ И. Г. Галямина.- 2-е изд. - СПб.: Питер, 2013. – 304 с.
23. Базилевский М.П. Критерии нелинейности многофакторных квазилинейных регрессий // Сборник материалов II Всероссийской национальной научной конференции студентов, аспирантов и молодых ученых. 2019. С. 210-213.
24. Масловская А.Г. Методы и средства обработки экспериментальных данных. Практикум. – Благовещенск: Амурский гос. ун-т, 2014. – 96 с.
25. Овсянников С.В. Экспериментальные исследования в мехатронных системах: учеб. пособие / С.В. Овсянников, А.А. Бошляков, А.О. Кузьмина. – М. : Изд-во МГТУ им. Н.Э. Баумана, 2010. – 49 с.
26. Базилевский М.П. Критерии нелинейности квазилинейных регрессионных моделей // Моделирование, оптимизация и информационные технологии. – 2018. – Т. 6. – № 4 (23). – С. 185–195.
27. Борисова М.В., Титов А.Ю., Новиков В.В., Коновалов В.В. Регрессионная модель производительности опорожнения тихоходного смесителя // Вестник Башкирского государственного аграрного университета. 2019. – № 2 (50). – С. 103-108.
28. Наумов В.А., Коржавина Ю.Н., Шибeko А.Г., Сингаев В.И., Алышевский Д.Л. Регрессионная модель плотности имитационного шпика // Известия КГТУ. 2018. – № 49. – С. 145-153.
29. Хайкин Саймон. Нейронные сети: полный курс. 2-е изд.: Пер. с англ. — М.: ООО «И.Д. Вильямс», 2016. — 1104 с.

Приложение А

Листинг А.1 – Код обученной ИНС в Matlab

```
function [Y,Xf,Af] = Podacha1(X,~,~)
%MYNEURALNETWORKFUNCTION      neural      network
simulation function.
%
% Auto-generated by MATLAB, 20-Apr-2021 16:45:26.
%
% [Y] = myNeuralNetworkFunction(X,~,~) takes these arguments:
%
% X = 1xTS cell, 1 inputs over TS timesteps
% Each X{1,ts} = Qx2 matrix, input #1 at timesteps.
%
% and returns:
%
% Y = 1xTS cell of 1 outputs over TS timesteps.
% Each Y{1,ts} = Qx1 matrix, output #1 at timesteps.
%
% where Q is number of samples (or series) and TS is the number
of timesteps.

%#ok<*RPMT0>

% ===== NEURAL NETWORK CONSTANTS =====

% Input 1
x1_step1.xoffset = [1;0.7];
x1_step1.gain = [1;1.04166666666667];
x1_step1.ymin = -1;

% Layer 1
```

b1 = [8.3201061283617416819;-
 7.7823341438912230217;7.2861053330657581384;6.8191728989
 336155564;-6.2913601642241960832;-5.8352460794211289397;-
 5.3896070963327247227;4.7115939949318237367;4.4353083919
 491176346;-3.9453496858097296673;-
 3.4160070938916522465;2.8883918913999222511;2.4184222521
 310569931;-1.9456318810104966666;1.4726933455438524501;-
 0.96617803183259720612;-
 0.07664772820500062589;0.15287950015369303203;-
 0.56674777927528607879;-1.1863822206132328407;-
 1.4128788818133199978;1.627061293268337927;-
 2.4397288749270109598;2.921863706006312178;-
 3.5174989996297303385;4.1297428842065153631;-
 4.2761106211785886444;4.8203992417390058023;5.5018750185
 743359182;5.8624225937176435153;6.3335990744389540907;-
 6.8140273815228491472;7.1469955755602745384;-
 7.9064038263046967359;-8.3353453957587699819];
 IW1_1 = [-0.33369870018065234119
 8.2509248622093966929;8.2954822719909042661 -
 0.021233099762698333379;-8.28042222015443663 -
 0.62988514733770639697;-0.89981489241557655046
 8.2128302304250393462;8.1011943914689457102
 1.8371476106824000674;3.0756760767409514479
 7.6994468278837340236;0.46689050430526690683
 8.2546304090022246669;-5.469528329356703189
 6.3238722885920939376;-1.4014026656409137583

 8.1212718563257073612;3.003306864723240821 -
 7.6992171769419508465;4.6991700306925494601 -
 6.8169389225418779432;-
 2.1702403533576464234 7.9982302939519129836;-
 2.2287881485306062501 -

7.9800722461759328041;8.2807387495784929854 -
 0.17264027358295219727;-6.7562658166017328654 -
 4.782571210470336176;5.1761471318723915047
 6.4583218833169517836;7.3627953233321168369
 3.8209779821677440914;8.2825011908209589251
 0.0090032362398916251922;-
 8.2610986739338336804 -0.54353476986217863498;-
 4.2175042145500558632 -7.0323538738159312445;-
 5.584825127612077722
 6.1193951683503540195;3.3834945083281140477 -
 7.6246633295174213174;-5.3581703318879005948 -
 6.3153788922573550479;8.243129691122829783
 0.80772224375853785006;-5.6950015752071276864 -
 5.885940769341147849;3.6003517449916344617
 7.2690034680079484986;-0.1439644696224760545
 8.3327848752925515186;6.5609527109162337766
 5.1058235565593985683;4.1830063587520776025 -
 7.0450132385336692309;7.6863717854423283526
 3.0513084857529975125;8.2732441588771941099
 0.39266502335954378555;-2.6901701336648464746
 7.8481732266282211441;6.7877901337869399612
 5.0116040328921513947;-5.7521117251008071491
 5.8260120515301796473;-8.1898560661693196039 -
 0.7929999301182942828];

% Layer 2

b2 = 0.63934967699677891506;

LW2_1 = [-0.30856221884826423585 0.44733310701578399282 -
 0.6255224239771308703 0.038011364766081912059 -
 0.80494604877401887411 -0.32615676403132015659
 0.70355400272756651869 0.52647416247697531588 -
 0.25096213211449436287 -0.48004985787546516907 -

0.16353807517026763829	0.18260328855719226704	
0.48615548914117900914	0.072113766406598639458	-
0.35490990525892457397	0.29021995714193388194	-
1.1525794592936142724	0.93805404446110185468	-
0.68060535361524276965	-0.14646127483487389465	
0.19428473814221774996	-0.30657460924043988815	
0.019719392094996707199	-0.19604767424847252544	
0.66627641446057039865	0.16832265738557664059	-
0.92850260209714063908	0.30743860659729477236	-
0.072708339169210725594	-0.099577218410849432928	
0.22287698597032926107	0.86101326045287740474	-
0.39151416627558016659	0.57333532190089020464	-
0.5518391537285276538];		

% Output 1

y1_step1.ymin = -1;

y1_step1.gain = 0.142857142857143;

y1_step1.xoffset = 4;

% ===== SIMULATION =====

% Format Input Arguments

isCellX = iscell(X);

if ~isCellX

 X = {X};

end

% Dimensions

TS = size(X,2); % timesteps

if ~isempty(X)

 Q = size(X{1},1); % samples/series

else

```

    Q = 0;
end

% Allocate Outputs
Y = cell(1,TS);

% Time loop
for ts=1:TS

% Input 1
X{1,ts} = X{1,ts}';
    Xp1 = mapminmax_apply(X{1,ts},x1_step1);

% Layer 1
    a1 = tansig_apply(repmat(b1,1,Q) + IW1_1*Xp1);

% Layer 2
    a2 = repmat(b2,1,Q) + LW2_1*a1;
% Output 1
Y{1,ts} = mapminmax_reverse(a2,y1_step1);
Y{1,ts} = Y{1,ts}';
end

% Final Delay States
Xf = cell(1,0);
Af = cell(2,0);

% Format Output Arguments
if ~isCellX
    Y = cell2mat(Y);

end

```

```

end
% ===== MODULE FUNCTIONS =====

% Map Minimum and Maximum Input Processing Function
function y = mapminmax_apply(x,settings)
y = bsxfun(@minus,x,settings.xoffset);
y = bsxfun(@times,y,settings.gain);
y = bsxfun(@plus,y,settings.ymin);
end
% Sigmoid Symmetric Transfer Function
function a = tansig_apply(n,~)
a = 2 ./ (1 + exp(-2*n)) - 1;
end

% Map Minimum and Maximum Output Reverse-Processing
Function
function x = mapminmax_reverse(y,settings)

x = bsxfun(@minus,y,settings.ymin);
x = bsxfun(@rdivide,x,settings.gain);
x = bsxfun(@plus,x,settings.xoffset);
end

```

Программа расчета подачи в функции от типа комбайна и потерь

```

P=Podacha1(Inputs)
fori=1:3
Komb(i)=i
for j=1:20
Poter(j)=0.7+0.1*(j-1)
Podach(i,j)=Podacha1([Komb(i) Poter(j)])
end

```

```

end
[X,Y] = meshgrid(1:3,0.7:0.1:2.6)
surf(X,Y,Podach')

```

Листинг А.2 – Аппроксимация скорости комбайна в функции типа комбайна и зернопотерь методом наименьших квадратов (многофакторной регрессии)

```

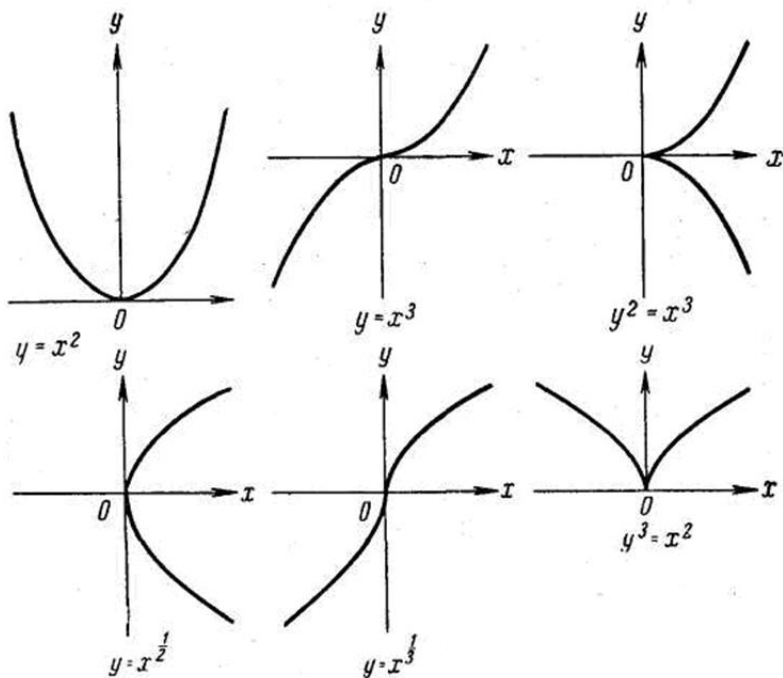
for i=1:15
    X(i,1)=1;
    X(i,2)=Inputs(i,1);
    X(i,3)=Inputs(i,2);
    X(i,4)=Inputs(i,1)*Inputs(i,2);
    X(i,5)=Inputs(i,1)^2;
    X(i,6)=Inputs(i,2)^2;
    Y(i)=Inputs(i,3);
end;
A=(X'*X)^-1*(X'*Y);
for i=1:3
    Komb(i)=i
    for j=1:20
        Poter(j)=0.7+0.1*(j-1)
        R(i,j)=A(1)+A(2)*Komb(i)+A(3)*Poter(j)+A(4)*Komb(i)*Poter(j)
        +A(5)*Komb(i)^2+A(6)*Poter(j)^2;
    end;
end;
[X1,Y1] = meshgrid(1:3,0.7:0.1:2.6);
subplot(1,2,1);
surf(X1,Y1,R');
for i=1:15
    Y2(i)=A(1)+A(2)*Inputs(i,1)+A(3)*Inputs(i,2)+A(4)*Inputs(i,1)*I
nputs(i,2)+A(5)*Inputs(i,1)^2+A(6)*Inputs(i,2)^2;

```

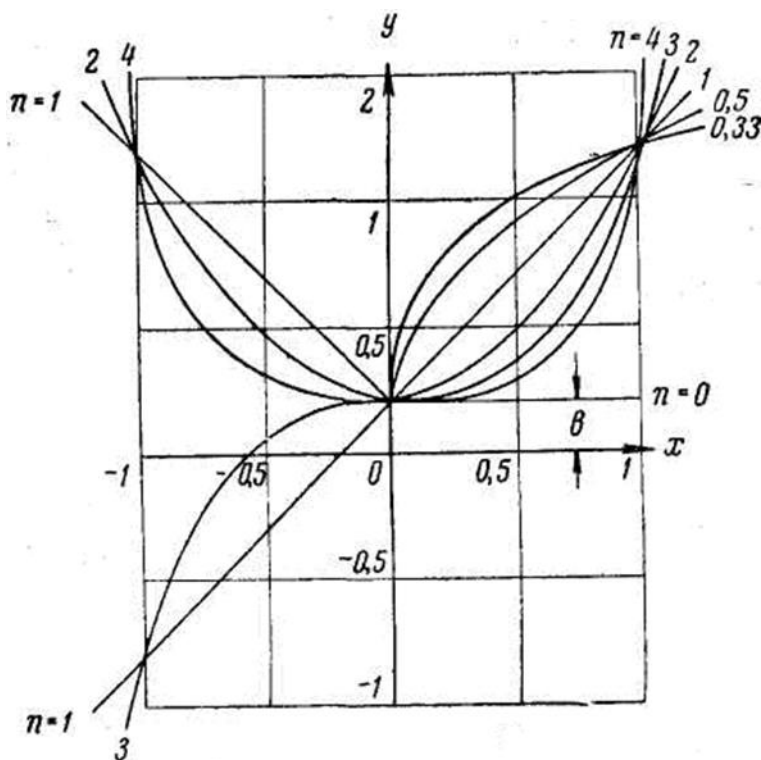
```
end;  
Y=sort(Y);  
Y2=sort(Y2);  
S=corrcoef(Y,Y2);  
subplot(1,2,2);  
plot(Y,Y2,Y,Y)
```


Приложение Б

Графики функций для подбора типов уравнений эмпирических формул



Кривые типа $y = ax^n + b$ при $a=1$



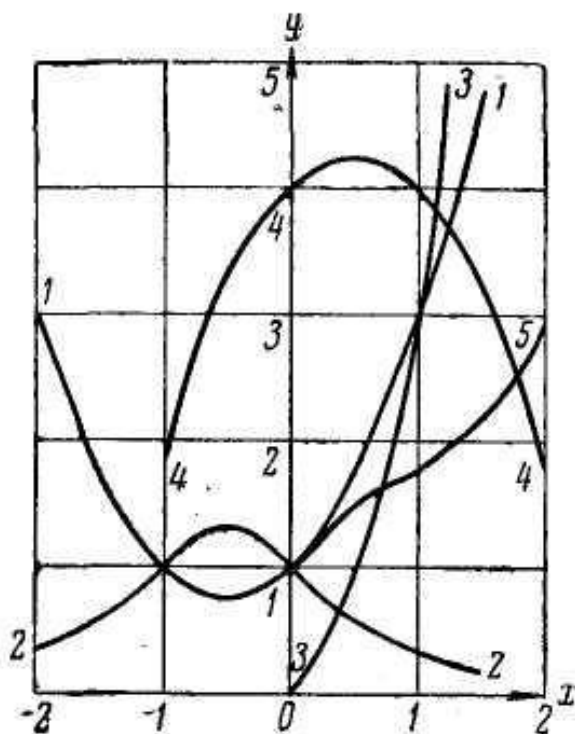
1—1 — кривая типа $y=a+bx+cx^2$ при $a=b=c=1$;

2—2 — кривая типа $y = \frac{1}{a+bx+cx^2}$ при $a=b=c=1$;

3—3 — кривая типа $y=ax+cx^2+cx^3$ при $a=b=c=1$; и $x>0$;

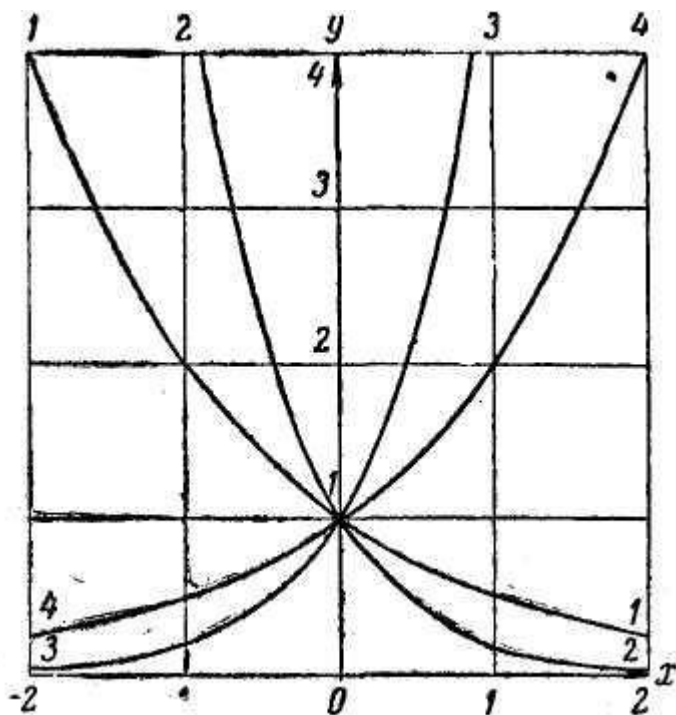
4—4 — кривая типа $y=a+bx+cx$ при $a=4$; $b=1$ и $c=-1$;

5—5 — кривая типа $y=ax+bx^2+cx^3$ при $a<1,0$, $b<1,0$ и $c<1,0$.

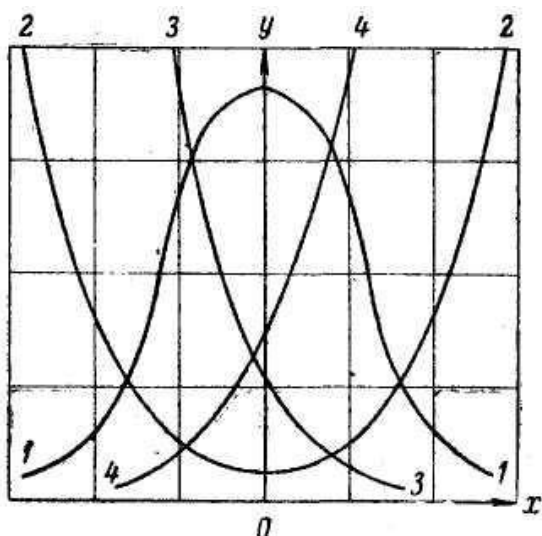


Кривые типа $y=a^x$:

1—1 — при $a=0,5$; 2—2 — при $a=0,2$; 3—3 — при $a=5$; 4—4 — при $a=2$.

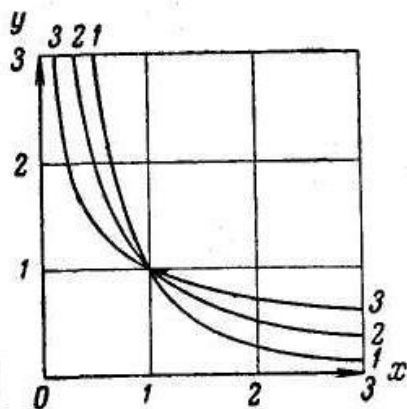


- 1—1 — кривая типа $y = ab^{-xc}$ при $b > 0$ и четном значении c ;
 2—2 — кривая типа $y = ab^{xc}$ при $b > 0$ и четном значении c ;
 3—3 — кривая типа $y = ab^{-xc}$ при $b > 0$ и нечетном значении c ;
 4—4 — кривая типа $y = ab^{xc}$ при $b > 0$ и нечетном значении c

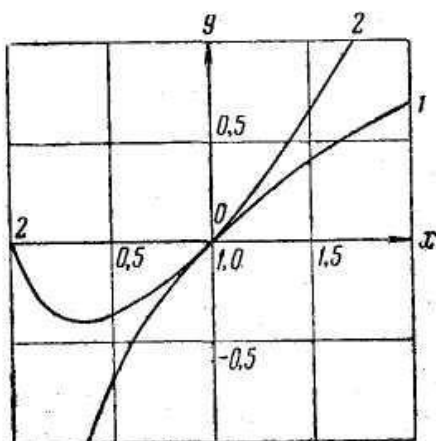


Кривые типа $y=x^m$:

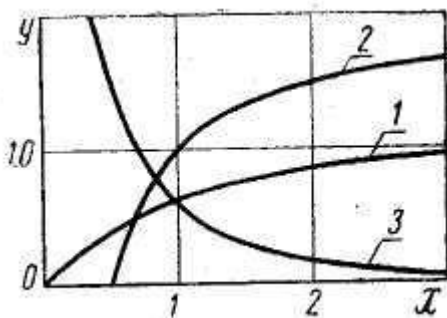
1—1 — при $m=2$; 2—2 — при $m=1$; 3—3 — при $m=0,5$



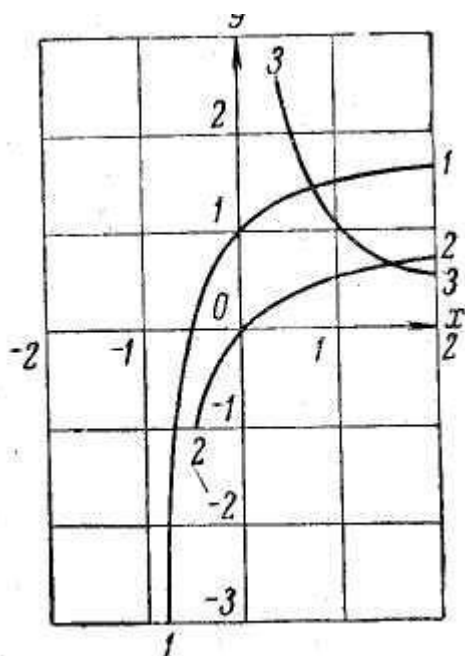
1—1 — кривая типа $y = \ln x$; 2—2 — кривая типа $y = x \ln x$



1 — кривая типа $y = 1 - e^{-x}$; 2 — кривая типа $y = a - (b/x)$ при $a=2$ и $b=1$; 3 — кривая типа $y = \frac{1}{e^x - 1}$.



1—1 — кривая типа $y = \frac{x}{a+bx} + c$ при $a=b=c=1$;
 2—2 — кривая типа $y = \frac{x}{a+bx}$ при $a=b=1$;
 3—3 — кривая типа $y = a/x$ при $a=1$.



Кривые типа $x^m y^n = c$ при $c=1$:

1—1 — при $m=2, n=1$; 2—2 — при $m=1, n=2$; 3—3 — при $m=n=2$; 4—4 — при $m=3, n=2$; 5—5 — при $m=2, n=3$; 6—6 — при $m>n$, m и n нечетные числа; 7—7 — при $m<n$, m и n — нечетные числа.

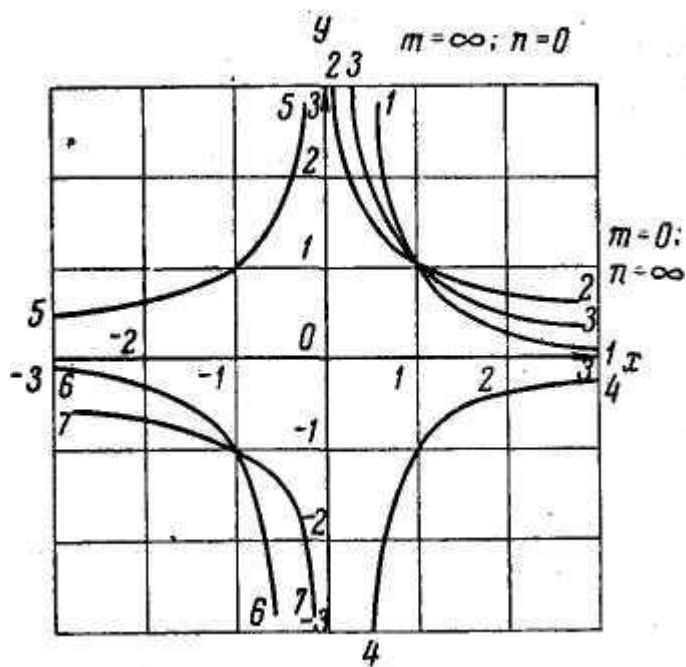
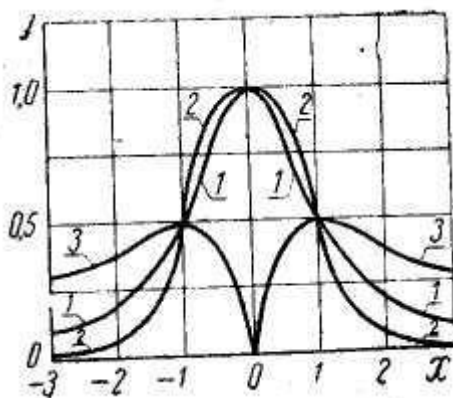
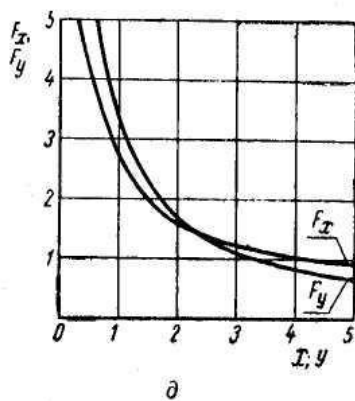
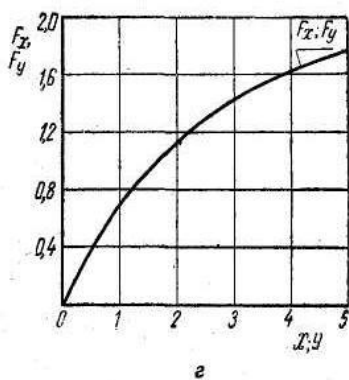
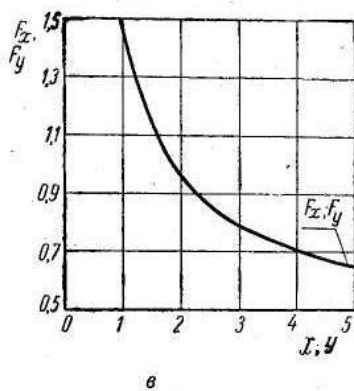
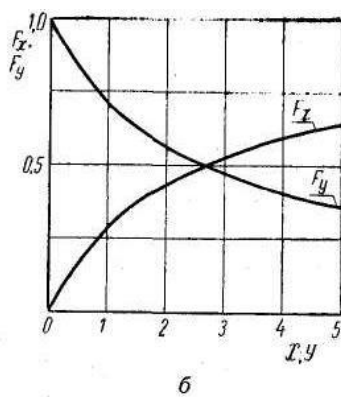
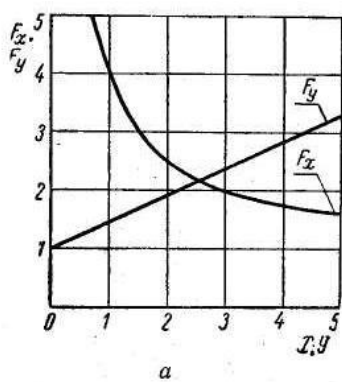
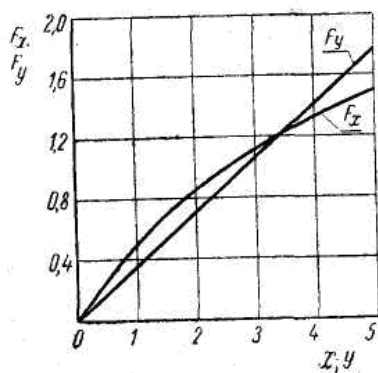


График функций $y = \frac{a}{1+x^n}$ при n — положительном:
 1—1 — при $n=2, a=1$; 2—2 — при $n=4, a=1$; 3—3 — при $y = \frac{|x|}{1+x^2}$.

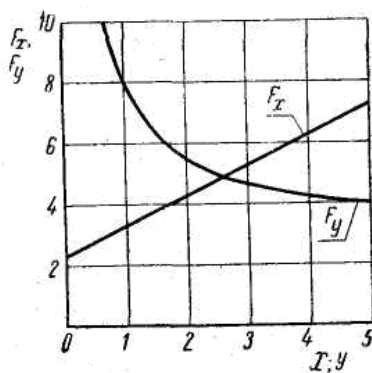


a — функционал	типа	$F_{xy} = \frac{x+y}{x};$
b — функционал	типа	$F_{xy} = \frac{x}{x+y};$
b — функционал	типа	$F_{xy} = \frac{x+y}{xy};$
c — функционал	типа	$F_{xy} = \frac{xy}{x+y};$
d — функционал	типа	$F_{xy} = \frac{x+d}{xy};$
e — функционал	типа	$F_{xy} = \frac{xy}{x+d};$
$ж$ — функционал	типа	$F_{xy} = x + \frac{d}{y};$
$з$ — функционал	типа	$F_{xy} = x + y + \frac{d}{xy};$
$и$ — функционал	типа	$F_{xy} = \frac{x}{y} + \frac{y}{x};$

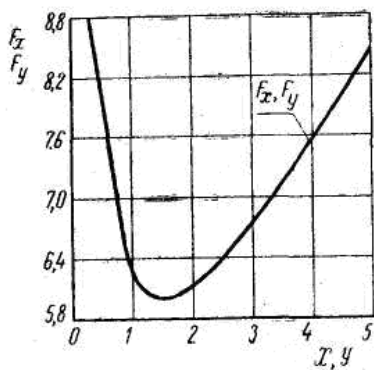




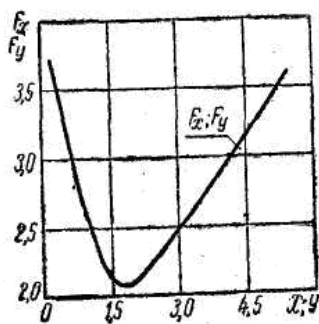
e



ж



з



и