

МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ  
Белорусский национальный технический университет

---

Кафедра «Высшая математика»

## МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Учебно-методическое пособие  
для обучающихся по специальностям  
1-37 01 02 «Автомобилестроение (по направлениям)»,  
1-37 01 01 «Двигатели внутреннего сгорания»

*Рекомендовано учебно-методическим объединением по образованию  
в области транспорта и транспортной деятельности*

Минск  
БНТУ  
2022

УДК 519.2(075.8)

ББК 22.172я7

М34

А в т о р ы:

*Т. И. Чепелева, Н. И. Чепелев,*

*С. Н. Чепелев, М. В. Щукин*

Р е ц е н з е н т ы:

кафедра общей математики и информатики

механико-математического факультета БГУ,

(зав. кафедрой, д-р экон. наук, профессор *С. А. Самаль*);

д-р экон. наук, профессор, зав. кафедрой бизнес-администрирования

Института бизнеса БГУ *Г. А. Хацкевич*

**Математическая статистика** : учебно-методическое пособие для  
М34 обучающихся по специальностям 1-37 01 02 «Автомобилестроение  
(по направлениям)» 1-37 01 01 «Двигатели внутреннего сгорания» /  
М. В. Щукин [и др.]. – Минск : БНТУ, 2022. – 84 с.  
ISBN 978-985-583-821-1.

Учебно-методическое пособие предназначено для студентов вузов, изучающих математическую статистику, может быть использовано аспирантами, магистрантами, инженерами для расчетов статистических данных любой структуры.

В пособие включены теоретические вопросы для подготовки к экзамену, задачи для аудиторной и самостоятельной работы, типовые расчеты по математической статистике, а также таблицы значений функций, необходимые для решения задач.

В пособии имеется краткое описание обращения к программным средствам системы STATISTICA, предназначенной для статистической обработки данных в среде Windows.

УДК 519.2(075.8)

ББК 22.172я7

ISBN 978-985-583-821-1

© Белорусский национальный  
технический университет, 2022

## Введение

**Математическая статистика** – это раздел математики, посвященный математическим методам систематизации, обработки и использования статистических данных для научных и практических выводов обоснований, где статистические данные – это сведения о числе объектов в какой-либо совокупности.

**Статистическим методом исследования** называется метод, опирающийся на рассмотрение статистических данных о совокупностях объектов.

**Предмет математической статистики** составляет формальная математическая сторона статистических методов исследования, относящаяся к специфической природе изучаемых объектов.

**Статистика** – это наука, позволяющая увидеть закономерности в хаосе случайных данных, выделить устойчивые связи в них, определить действия человека с тем расчетом, чтобы увеличить долю правильно принятых им решений среди всех принимаемых. «Статистика» (от лат. «status») – состояние, положение вещей.

Статистика родилась из потребностей практической деятельности человека. Конфуций (5 в. до н. э.) рассказывает о переписи населения в Китае, проведенной в 2238 г. до н. э. Книга Моисея является одной из древнейших по исчислению населения, способного носить оружие.

Статистическая наука призвана изучать закономерности в случайных данных, которые будут использоваться в дальнейшем при исследовании тех или иных процессов, явлений, дает возможность вскрывать причинно-следственные связи. Развитие и совершенствование цивилизаций неизбежно ведет к расширению и усилению потребностей статистических методов расчета.

На практике применение статистических методов – чрезвычайно сложное дело, требует больших интеллектуальных усилий и временных затрат.

## Теоретические вопросы для подготовки к экзамену

1. Выборочная совокупность. Вариационный ряд.
2. Полигон и гистограмма.
3. Эмпирическая функция распределения и ее свойства.
4. Выборочная средняя и выборочная дисперсия.
5. Оценки параметров распределения. Точечные оценки и требования, предъявляемые к ним.
6. Точечные оценки для математического ожидания и дисперсии.
7. Интервальные оценки. Доверительный интервал.
8. Распределение Стьюдента.
9. Распределение Пирсона.
10. Построение доверительного интервала для математического ожидания нормально распределенной СВ (случайной величине) при известном среднем квадратическом отклонении.
11. Построение доверительного интервала для математического ожидания нормально распределенной СВ при неизвестном среднем квадратическом отклонении.
12. Построение доверительного интервала для среднего квадратического отклонения нормально распределенной СВ.
13. Понятие о статистических гипотезах и критериях согласия.
14. Критерий согласия Пирсона  $\chi^2$ .
15. Критерий согласия Колмогорова.
16. Выборочный коэффициент корреляции и его свойства.
17. Уравнение регрессии. Линейная регрессия. Определение коэффициентов линейной регрессии методом наименьших квадратов.
18. Нелинейная регрессия. Определение параметров нелинейной регрессии.

# АНАЛИТИЧЕСКАЯ СИСТЕМА TIBCO STATISTICA™

## Общая информация

С внедрением информационных технологий в науку, технику, производство, медицину и во все сферы народного хозяйства бурно развивается рынок статистического программного обеспечения, открывающего путь к новым технологиям статистической обработки данных, максимально сокращающий рутинные процедуры, связанные с обработкой многомерных, сложных данных и зависимостей в данных.

Огромное предпочтение на этом рынке отдается программному продукту Statistica компании TIBCO Software Inc., США (далее – STATISTICA), представляющему собой интегрированную систему статистического анализа и обработки данных.

Система STATISTICA состоит из следующих основных компонентов:

- рабочих книг – являются оптимизированными ActiveX-контейнерами, которые позволяют эффективно обрабатывать большое количество Документов. Все Документы упорядочены в иерархическом виде с использованием отдельных папок или узлов Документов (по умолчанию для каждого нового Анализа создается отдельный узел). Используя дерево Рабочей книги, можно легко управлять структурой, изменяя положение отдельных Документов, папок или целых ветвей дерева;

- таблиц данных (мультимедийных таблиц) – для ввода и задания исходных данных, вывода численных и мультимедийных результатов анализа и др.;

- графиков – графической системы для визуализации данных и результатов статистического анализа;

- набора специализированных статистических модулей;

- специального инструментария для подготовки отчетов;

- встроенного языка программирования STATISTICA VISUAL BASIC.

Специализированными модулями системы являются основные статистические функции, таблицы, непараметрическая статистика, дисперсионный анализ, множественная регрессия, нелинейное оце-

нивание, кластерный анализ, факторный анализ, дискриминантный функциональный анализ, каноническая корреляция и др.

### Этапы статистического анализа:

- ввод исходных данных в электронную таблицу и их предварительное преобразование (выборка, ранжирование и т. д.);
- визуализация данных с помощью одного из типов графиков;
- применение конкретной процедуры статистической обработки;
- вывод результатов анализа в виде электронных таблиц с численной и текстовой информацией и графиков;
- подготовка и печать отчета.

### STATISTICA имеет два интерфейса:

1. Классический (состоит из меню команд).
2. Ленточный.

По умолчанию в STATISTICA 13 активирован ленточный режим (рис. 1). Переключение между режимами осуществляется путем нажатия соответствующей кнопки в строке заголовка при ленточном режиме либо через меню Вид – Лента при классическом режиме.

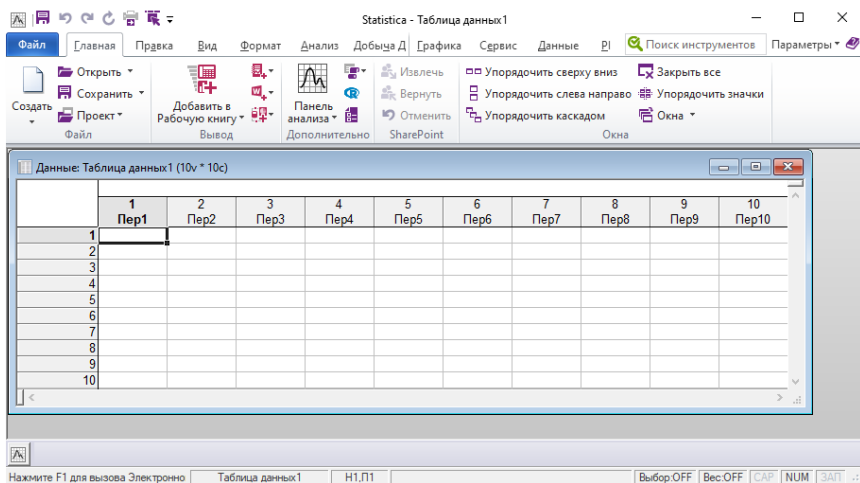


Рис. 1. Интерфейс STATISTICA

Взаимодействия с системой носит интерактивный характер и осуществляется при помощи последовательного выбора необходимых команд из меню (классический интерфейс) либо ленты (ленточный интерфейс).

## Запуск STATISTICA

Запустив ОС Windows, нажав кнопку «Пуск», выбрать в меню «Программы» папку, которая содержит систему STATISTICA. В этой папке выбрать ярлык программы STATISTICA и щелкнуть по нему мышью. После инсталляции системы папка STATISTICA 13 будет содержать следующие элементы:

- STATISTICA – ярлык для запуска программы;
- электронное руководство STATISTICA – справочная система.

После первого запуска STATISTICA 13 на экране появляется окно с ленточным интерфейсом и пустой таблицей исходных данных.

С целью совместимости с предыдущими версиями системы мы будем использовать классический интерфейс (рис. 2). Поэтому рекомендуем переключиться на него заранее через соответствующий пункт строки заголовка программы.

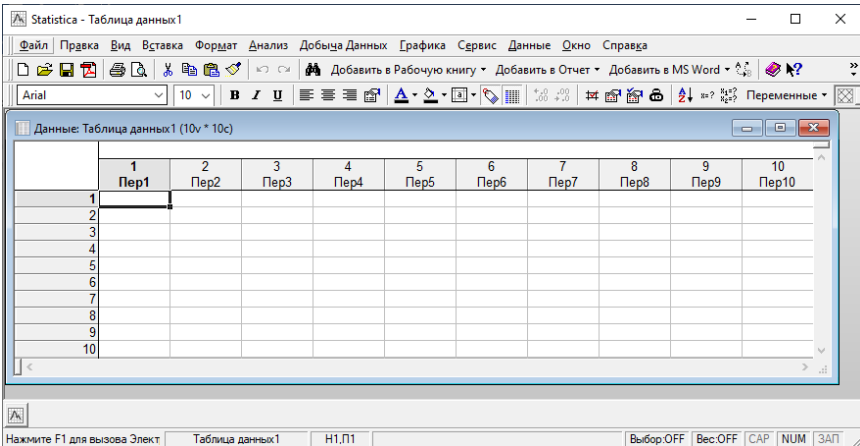


Рис. 2. Классический интерфейс

## Главное окно STATISTICA

«Строка заголовка» содержит название активного статистического модуля – STATISTICA Таблица данных 1. В правой части – слева направо:

- минимизации размеров окна;
- восстановления окна;
- закрытия окна.

Вторая строка – «строка меню», содержит команды, упорядоченные по функциональному признаку.

STATISTICA работает с разными типами документов, при этом каждый документ имеет свое меню и свою панель инструментов. Для электронной таблицы "SpreadSheet" ввода исходных данных меню содержит следующие команды:

- File – Файл;
- Server – Сервер (если активирован);
- Edit – Правка;
- View – Вид;
- Insert – Вставка;
- Format – Формат;
- Statistics – Анализ;
- Data Mining – Добыча данных;
- Graphs – Графика;
- Tools – Сервис;
- Data – Данные;
- Window – Окно;
- Help – Справка.

«Панель инструментов» расположена в третьей строке и в левом первом столбце. Большую часть основного окна STATISTICA занимает рабочая область, на которую выводятся документы, с которыми ведется работа:

- рабочие книги;
- таблицы данных (мультимедийные таблицы);
- отчеты;
- графики;
- макросы (STATISTICA Visual Basic).

Окно документа состоит из строки заголовка, рабочей области для ввода и (или) отображения информации и инструментов по управлению окном. Дополнительно STATISTICA имеет контекстные меню, позволяющие получать быстрый доступ к наиболее часто используемым командам для работы с тем или иным объектом в активном окне.



Контекстное меню электронной таблицы с исходными данными обеспечивает доступ к группам команд:

- построение графиков для выделенного блока значений;
- вычисление основных статистик;
- задание спецификаций переменных, операции с текстовыми значениями;
- преобразование структуры электронной таблицы (добавление, удаление строки и т. д.);
- операции с выделенным блоком значений;
- операции с «Буфером обмена Windows».

### ***Spreadsheet – электронная таблица с исходными данными***

Таблицы данных STATISTICA основаны на технологии мультимедийных таблиц. Система работает как с исходными данными, так и с численными и текстовыми результатами. Таблица данных STATISTICA является двумерной таблицей, которая может содержать практически неограниченное число наблюдений (строк) и переменных (столбцов), при этом каждая ячейка может содержать неограниченное количество символов (рис. 3). В ячейки можно помещать не только числа и текст, но и любые другие документы, использующие технологию ActiveX, например, звук, видео, графику, анимацию, отчеты с вложенными объектами. Поскольку Таблицы данных STATISTICA могут содержать макросы и поддерживают настраиваемый пользовательский интерфейс, то их можно использовать для пользовательских приложений (например, добавлять список опций или ряд кнопок в верхнем левом углу), презентациях, анимациях и т. д.

Для работы с таблицами исходных данных существует большое количество инструментов, которые доступны при помощи выпадающих и контекстных меню и из панели инструментов, включая:

- операции, которые изменяют структуру электронной таблицы (добавление, удаление, копирование, перемещение переменных и случаев);
- операции по заданию спецификаций (имен, форматов и т. д.) для переменных и случаев;
- операции с выделенным блоком значений;

– операции, реализованные при помощи Drag-and-Drop – перетаскивать и опустить, включая операции по копированию, перемещению и автозаполнению блока и др.

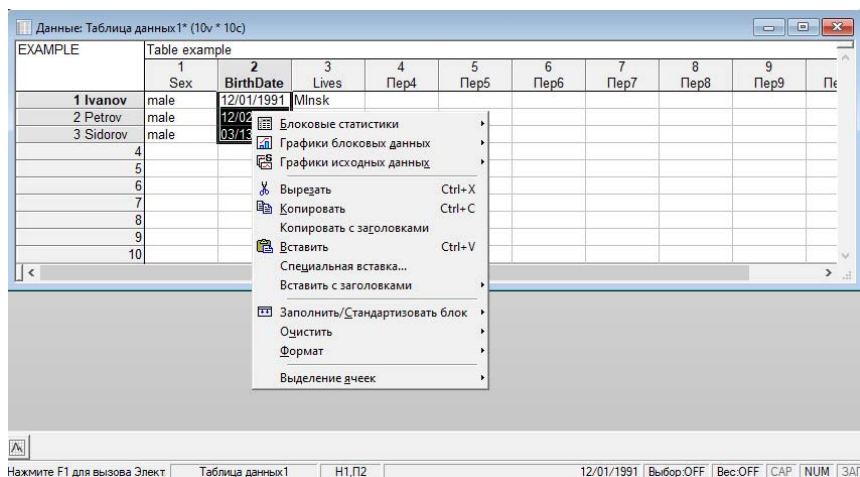


Рис. 3. Окно Spreadsheet

**Variables** – переменная – имеет свое имя, формат и другие атрибуты, которые называются спецификацией и задаются пользователем. Переменная представляет собой наблюдаемую величину. Результаты наблюдений записываются в строках таблицы – случаях Cases.

В Spreadsheet можно задать спецификации переменных, которые включают в себя:

- формат отображения данных (количество десятичных разрядов в представлении чисел, формат представления дат, времени, соответствие между численными и текстовыми значениями и др.);
- код, приписываемый пропущенным данным;
- длинные имена переменных и комментарии к ним;
- формулы, используемые для определения, перекодировки или преобразования значений переменных;
- динамические связи между файлом данных STATISTICA и другим Windows-совместимым файлом и др.

Можно включать в Таблицу любые другие внедренные или связанные объекты (например, мультимедийные объекты, Макросы). Текст в ячейках может иметь практически неограниченную длину (обычно в настройках системы STATISTICA она ограничена 1000 символами для предотвращения случайной вставки большого объема нежелательной информации). Имеются широкие возможности для форматирования текста в ячейках.

Чтобы создать новую Таблицу данных выполните следующие действия:

1. Выберите команду Создать в меню Файл или нажмите кнопку Создать на панели инструментов.
2. В диалоге Создать новый Документ выберите вкладку Таблица.
3. Задайте Число переменных и Число наблюдений.
4. В группе Положение выберите способ создания Таблицы данных: В новой Рабочей книге или В отдельном окне. Выберите опцию В новой Рабочей книге для создания Таблицы данных в пустой Рабочей книге. Выберите опцию В отдельном окне, чтобы создать пустую Таблицу в новом окне.
5. Нажмите кнопку ОК.

Для того чтобы изменить содержимое ячейки, не удаляя предыдущие данные, выполните следующие действия:

1. Дважды нажмите в соответствующем поле ячейки. Вы перейдете в режим редактирования, курсор будет установлен в поле ячейки. Вы можете нажать клавишу F2 и также перейти в режим редактирования выбранной ячейки.
2. Внесите соответствующие изменения в ячейку.
3. Нажмите клавишу ENTER.

Для того чтобы полностью заменить содержимое ячейки, нажмите один раз мышью по соответствующей ячейке и введите данные.

## **Построение статистического графика**

Система STATISTICA включает широкий спектр графических методов для визуального представления результатов исследований. Все графические средства системы STATISTICA обеспечивают возможность выбора встроенного аналитического интерактивного метода анализа и содержат большой набор программ настройки, позволяющих пользователю интерактивно управлять отображением

информации на экране (рис. 4). Гибкие средства управления одновременно несколькими графиками дают возможность пользователю объединять различные графические изображения и осуществлять динамический обмен между приложениями (используя средства OLE).

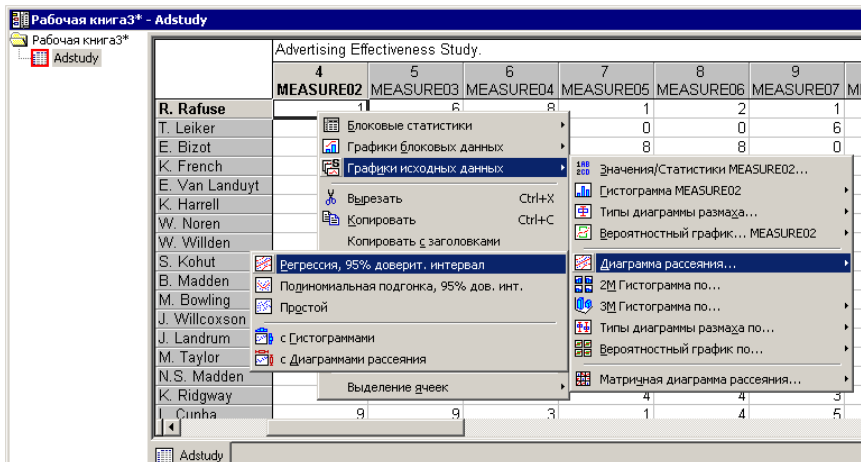


Рис. 4. Меню графиков

Инструменты графического анализа в системе STATISTICA позволяют не только создавать графики, но и обновлять их после внесения каких-либо изменений в массив данных.

Если таблица данных связана с некоторым внешним источником, настройки можно выбрать таким образом, что графики будут автоматически обновляться после получения любой новой информации из этого источника (например, из базы данных или системы сбора данных).

После построения график можно связать с различными переменными или даже с совершенно другой таблицей данных.

Обновление можно задать и другим способом. Изменения свойств точек графика (сервис закрашивание), таких как ярлык, цвет маркера, исключение точек, можно отображать в таблице данных. В том случае, когда несколько графиков связано с одной таблицей данных, это позволяет проводить разведочный анализ данных: любые

подобные изменения на одном из графиков будут моментально перенесены на все остальные.

Встроенный язык программирования STATISTICA Visual Basic полностью поддерживает практически все графические возможности системы STATISTICA. С его помощью можно создавать новые типы графиков и диаграмм, включая составные графические документы, а также редактировать уже существующие графики.

В программном пакете STATISTICA предлагаются разнообразные графические методы, с помощью которых пользователь может запрашивать или самостоятельно организовывать построение графиков. Эти методы, включающие большой набор графиков различных типов, таких как пользовательские, статистические и специализированные, дополняют друг друга, обеспечивая высокий уровень взаимосвязи между числовыми данными (начальными, промежуточными и конечными) и их графическим представлением.

В дополнение к специализированным статистическим графикам, создаваемым в результате статистических процедур, существуют **два основных типа или класса графиков, вызываемых из меню Графика, панели инструментов Графика или контекстных меню:**

- графики исходных данных;
- графики блоковых данных.

Основное различие между этими двумя типами графиков – это данные, используемые при их построении.

**Графики исходных данных** и их расширенная версия в меню Графика являются представлением исходных данных из текущей таблицы данных (обычно, для всех переменных или подмножеств).

Если графики этого типа создаются с помощью контекстного меню таблицы результатов, которая не содержит фактических данных (например, корреляционная матрица), то программа все равно будет использовать исходные данные для построения графика (например, диаграмма рассеяния переменных, заданных выделенными ячейками в корреляционной матрице, из которой вызывается контекстное меню).

**Графики блоковых данных**, с другой стороны, полностью независимы от «исходных данных» или «файла данных». Эти графики являются общим средством визуализации числовых значений выделенного блока любой таблицы (которая может содержать значения

пользовательских подмножеств таблицы результатов или произвольно выбранных подмножеств исходных данных).

Эти две категории графиков предлагают одинаковые настройки и выбор типов графиков. Например, создать специализированный категоризованный тернарный график можно как на основе исходного набора данных, так и на основе какого-то заданного блока значений из таблицы результатов.

## Статистический анализ

В системе STATISTICA для получения совокупности численной, текстовой и графической информации при обработке исходных данных используется меню или лента Анализ (Statistics) в зависимости от интерфейса.

Рассмотрим анализ данных на двух примерах.

### 1. Опция «Основные статистики и таблицы».

Для работы можно открыть учебную таблицу из файла Adstudy.sta (Меню Файл – Открыть папку примеров – Examples\Datasets – Adstudy.sta).

Исходный файл содержит 25 переменных и 50 наблюдений. Эти данные собраны в социологическом опросе, где женщины и мужчины оценивали качество двух реклам. Пол респондента кодировался с помощью переменной 1 (Пол – Gender: 1=мужчина – male, 2=женщина – female). Каждому респонденту случайным образом предлагался на рассмотрение один из двух рекламных роликов (Предпочтение – Advert: 1=Coke, 2=Pepsi). Затем они оценивали привлекательность соответствующего продукта по 23 различным шкалам (с Мера1 – Measure1 до Мера23 – Measure23). В каждой из этих шкал респонденты могли ставить оценки от 0 до 9.

Запустите модуль Основные статистики и таблицы (рис. 5).

*Выберите процедуру* Основные статистики и таблицы из меню Анализ для вызова стартовой панели модуля Основные статистики и таблицы. Этот диалог содержит одну вкладку: Быстрый. Процедура позволяет выполнить большое количество аналитических расчетов.

*ОК.* При нажатии этой кнопки вызывается диалоговое окно анализа, выбранного из списка во вкладке Быстрый.

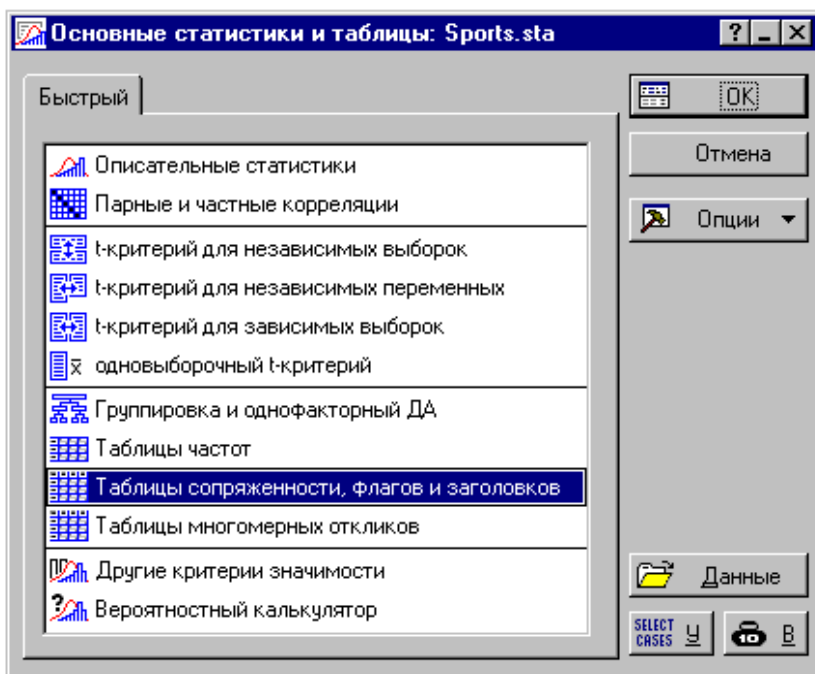


Рис. 5. Основные статистики и таблицы

*Отмена.* При нажатии этой кнопки, закрывается текущее диалоговое окно.

*Опции.* При нажатии этой кнопки, появляется всплывающее меню.

*Диспетчер вывода.* При нажатии этой кнопки, появляется диалоговое окно Диспетчер вывода Анализа/Графика, в котором устанавливаются текущие настройки Вывода в STATISTICA.

*Параметры отображения.* При нажатии этой кнопки, появляется диалоговое окно Параметры отображения Анализа/Графика, в котором устанавливаются текущие настройки отображения в STATISTICA.

*Создать макрос.* При нажатии этой кнопки, появляется диалоговое окно Новый макрос. При работе в STATISTICA, все ваши действия автоматически записываются, и когда вы нажимаете Создать макрос, полная запись ваших действий транслируется в программ-

ный код STATISTICA Visual Basic. Вы можете выполнить полученную программу для автоматического повтора ваших действий.

*Заккрыть Анализ.* При нажатии этой кнопки закрываются все диалоговые окна, ассоциированные с текущим Анализом.

*Данные.* При нажатии этой кнопки, появляется диалоговое окно Выберите Таблицу данных или документ, в котором необходимо выбрать Таблицу данных или документ для выполнения Анализов. Диалоговое окно Выберите Таблицу данных или документ содержит список активных в данный момент Таблиц.

*Выбор наблюдений.* При нажатии этой кнопки появляется диалоговое окно Условия выбора наблюдений Анализа/Графика. Эта функция используется для того, чтобы включить в анализ только подмножество наблюдений. Дополнительную информацию об условиях выбора наблюдений можно получить в соответствующем вводном обзоре, описании синтаксиса и описании диалога.

*В.* При нажатии кнопки В (Вес) появляется диалоговое окно Веса наблюдений Анализа/Графика. Эта функция позволяет «изменять» вклад отдельных наблюдений, пропорциональных значениям выбранной переменной.

Например, проверим, не коррелируют ли между собой оценки в различных шкалах (другими словами, не измеряют ли некоторые шкалы одни и те же свойства рекламы). В стартовой панели Основные статистики и таблицы выберите процедуру Парные и частные корреляции и щелкните ОК (или можете просто дважды щелкнуть на процедуре Парные и частные корреляции).

После выбора процедуры откроется диалоговое окно Корреляции Пирсона.

Вначале выберем переменные. Вы можете выбрать переменные как из одного списка (матрица будет квадратной), так и из двух списков (прямоугольная матрица).

Для нашего примера щелкните на кнопке Квадратная матрица, на экране появится стандартное диалоговое окно Выбор переменных.

Выберите все переменные и затем нажмите кнопку ОК для отображения таблицы результатов.

## **2. Опция «Множественная регрессия».**

Общее назначение множественной регрессии (этот термин был впервые использован в работе Пирсона – Pearson, 1908) состоит в анализе связи между несколькими независимыми переменными



(называемыми также регрессорами или предикторами) и зависимой переменной.

В общественных и естественных науках процедуры множественной регрессии чрезвычайно широко используются в исследованиях. В общем, множественная регрессия позволяет исследователю задать вопрос (и, вероятно, получить ответ) о том, что является лучшим предиктором.

Например, исследователь в области образования мог бы пожелать узнать, какие факторы являются лучшими предикторами успешной учебы в средней школе. А психолога мог бы заинтересовать вопрос, какие индивидуальные качества позволяют лучше предсказать степень социальной адаптации индивида.

Социологи, вероятно, хотели бы найти те социальные индикаторы, которые лучше других предсказывают результат адаптации новой иммигрантской группы и степень ее слияния с обществом.

Заметим, что термин «множественная» указывает на наличие нескольких предикторов или регрессоров, которые используются в модели.

В этом примере будем анализировать данные из файла Poverty.sta (Меню Файл – Открыть папку примеров – Examples\Datasets – Poverty.sta).

Данные основаны на сравнении результатов переписи 1960 и 1970 годов для случайной выборки из 30 округов. Имена округов введены в качестве идентификаторов наблюдений.

Ознакомиться с информацией по каждой переменной можно в электронной таблице Редактор спецификаций переменных (открываемой при выборе команды «Все спецификации...» в меню Данные).

Проанализируем корреляты бедности (т. е. предикторы, «сильно» коррелирующие с процентом семей, живущих за чертой бедности). Таким образом, будем рассматривать переменную 3 (Pt\_Poor) как зависимую или критериальную переменную, а все остальные переменные – в качестве независимых переменных или предикторов.

После выбора команды «Множественная регрессия» с помощью меню Анализ (Statistics), открывается стартовая панель модуля Множественная регрессия (рис. 6).

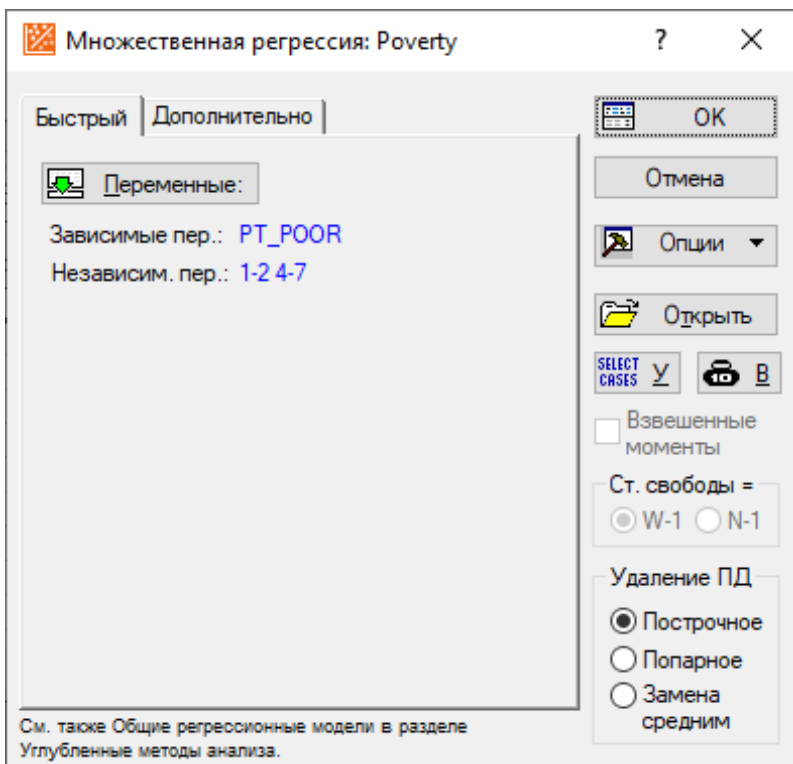


Рис. 6. Множественная регрессия

Можно задать регрессионное уравнение щелчком мыши по кнопке Переменные во вкладке Быстрый стартовой панели модуля Множественная регрессия.

В появившемся окне Выбора переменных выберите Pt\_Poor в качестве зависимой переменной, а все остальные переменные набора данных – в качестве независимых (выделяя их, зажав клавишу Ctrl).

Во вкладке Дополнительно отметьте также опции Показывать описательные статистики, корр. матрицы.

Если все сделано верно, появится окно «Результаты множественной регрессии», из которого доступен весь дальнейший статистический анализ (рис. 7).

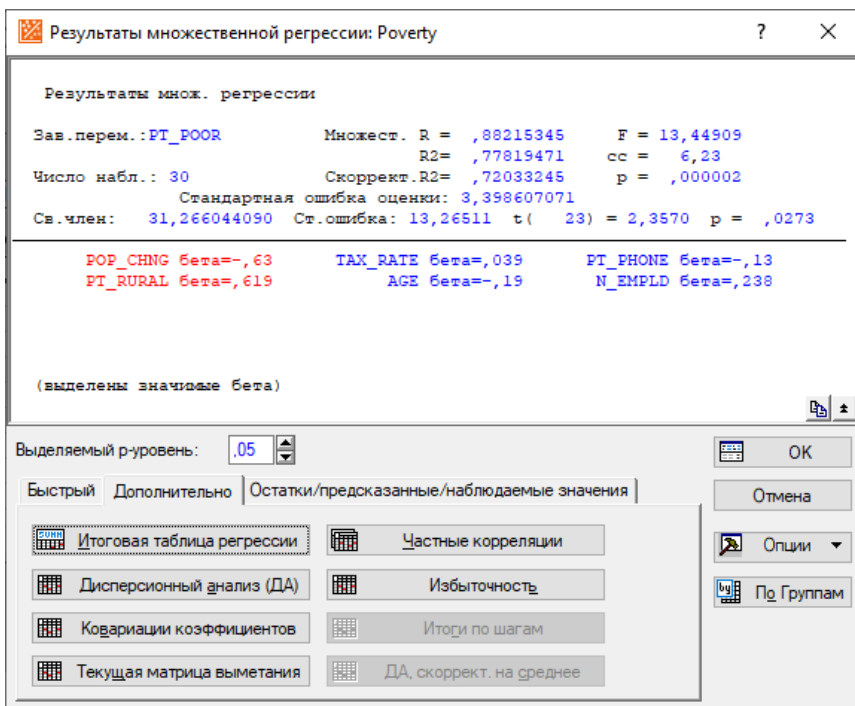


Рис. 7. Результаты множественной регрессии

## Вывод результатов анализа

При выполнении Анализа STATISTICA представляет результат в виде мультимедийных таблиц (Таблиц данных) и Графиков. Существует пять каналов, в которые можно направить все результаты:

1. Рабочие книги.
2. Отдельные окна.
3. Отчеты.
4. Microsoft Word.
5. Ресурсы сети Интернет.

Первые четыре из них управляются с помощью Диспетчера вывода (доступ к которому можно получить в меню Файл или в меню Сервис – Параметры). Имеется несколько различных способов размещения результатов в Интернете, различающиеся в зависимости от используемой версии STATISTICA.

Перечисленные каналы вывода можно использовать в различных комбинациях (например, Рабочую книгу и Отчет одновременно). Каждый из каналов вывода можно настроить различными способами. Кроме этого, каждый Объект (Таблицы данных или График), помещенный в Рабочую книгу или Отчет, может содержать другие вложенные и связанные Объекты и Документы.

**Рабочие книги** являются стандартным средством управления выводом. Каждый итоговый Документ (например, Таблица данных или График STATISTICA, а также документ Microsoft Word или Excel) представляется в виде вкладки в Рабочей книге (рис. 8).

Example data file: Performance on a memory test under stressful conditions

	1	2	3	4	5	6
	GROUP	GENDER	TIME	PAID	STRESS_R	CORRECT1
1	EXPERMTL	MALE	BEFORE	NOT_PAID	1,41	12
2	EXPERMTL	MALE	BEFORE	NOT_PAID	1,73	3
3	EXPERMTL	MALE	BEFORE	PAID	0,00	7
4	EXPERMTL	MALE	BEFORE	PAID	1,41	11
5	EXPERMTL	MALE	AFTER_1	NOT_PAID	12,83	8
6	EXPERMTL	MALE	AFTER_1	NOT_PAID	2,24	15
7	EXPERMTL	MALE	AFTER_1	PAID	2,24	15
8	EXPERMTL	MALE	AFTER_1	PAID	13,32	11
9	EXPERMTL	MALE	AFTER_2	NOT_PAID	13,74	19
10	EXPERMTL	MALE	AFTER_2	NOT_PAID	13,32	11
11	EXPERMTL	MALE	AFTER_2	PAID	13,00	9
12	EXPERMTL	MALE	AFTER_2	PAID	13,16	10
13	EXPERMTL	FEMALE	BEFORE	NOT_PAID	14,00	8
14	EXPERMTL	FEMALE	BEFORE	NOT_PAID	12,23	6
15	EXPERMTL	FEMALE	BEFORE	PAID	13,22	8
16	EXPERMTL	FEMALE	AFTER_1	PAID	12,23	7
17	EXPERMTL	FEMALE	AFTER_1	NOT_PAID	13,00	5

Рис. 8. Рабочая книга

Документы можно организовать в виде иерархии папок и элементов (по умолчанию, для каждого Анализа создается отдельный узел) с помощью дерева, в котором каждый отдельный элемент, папка или целая ветвь могут быть легко изменены.

## Занятие 1.

### Статистическое распределение. Эмпирическая функция распределения и ее свойства. Полигон и гистограмма.

#### Числовые характеристики выборки

#### 1.1. Краткие теоретические сведения

**Генеральной совокупностью** называется совокупность элементов, объединенных по некоторому признаку, из которых производится выборка.

**Выборочной совокупностью** или выборкой называется совокупность объектов, случайно выбранных для исследования.

**Объемом выборки** называется количество объектов, входящих в выборку.

Пусть из совокупности извлечена выборка объемом  $n$ .

Выборочная совокупность, расположенная по возрастанию или убыванию значения признака, называется *вариационным рядом*, а ее объекты – *вариантами*.

Если значения вариант совпадают или отличаются незначительно, то их можно сгруппировать, придав частоту каждой variante. В результате получим *сгруппированный вариационный ряд*.

*Частотой или относительной частотой* варианты называется отношение частоты варианты к объему выборки:

$$\omega_i = \frac{m_i}{n}. \quad (1.1)$$

**Статистическим распределением** называется соответствие, по которому каждому возможному значению варианты ставится в соответствие частота (относительная частота) ее появления. Статистическое распределение записывается в виде таблицы, в которой в первой строке перечислены все значения вариант, а во второй – частоты или частоты, которые соответствуют вариантам:

$x_i$	$x_1$	$x_2$	$x_3$	...	$x_k$
$m_i$	$m_1$	$m_2$	$m_3$	...	$m_k$

$$\sum_{i=1}^k m_i = n$$

Для построения интервального статистического ряда множество вариантов разбивают на полуинтервалы  $[a_i; a_{i+1})$ , т. е. производят группировку. Рекомендуется число интервалов  $k$  определять по формуле

$$k = 1 + 1,4 \cdot \ln n. \quad (1.2)$$

Длина интервала равна

$$\Delta = \frac{x_{\max} - x_{\min}}{k}. \quad (1.3)$$

Для наглядности используются графические изображения вариационных рядов в виде полигона и гистограммы.

**Полигоном** частот или частостей называется ломаная линия, соединяющая точки с координатами  $(x_i; m_i)$  или  $(x_i; \omega_i)$ .

**Гистограммой** частот или частостей называют ступенчатую фигуру, составленную из прямоугольников с основанием  $\Delta$  и высотой  $\frac{m_i}{\Delta}$  или  $\frac{\omega_i}{\Delta}$ .

**Эмпирической функцией распределения** называют функцию  $F^*(x)$ , определяющую для каждого значения  $x$  относительную частоту события  $X < x$ :

$$F^*(x) = \omega(X < x) = \frac{m_x}{n}, \quad (1.4)$$

где  $m_x$  – число вариантов (с учетом их кратностей) меньших  $x$ ;

$n$  – объем выборки.

Эмпирическая функция распределения обладает следующими свойствами:

1. Значения эмпирической функции принадлежат отрезку  $[0; 1]$ .
2. Эмпирическая функция является неубывающей функцией.
3. Если  $x_1$  – наименьшее значение варианты, а  $x_k$  – наибольшее значение варианты, то  $F^*(x) = 0$  при  $x \leq x_1$  и  $F^*(x) = 1$  при  $x > x_k$ .

Для описания выборки применяются такие числовые характеристики, как выборочная средняя, выборочная дисперсия, выборочное среднее квадратическое отклонение.

**Выборочной средней** называется среднее значение варианты, вычисленное по данным выборки:

$$\bar{x}_B = \frac{1}{n} \sum_{i=1}^n x_i \quad \text{или} \quad \bar{x}_B = \frac{1}{n} \sum_{i=1}^k m_i x_i,$$

где  $m_i$  – частота варианты  $x_i$ .

**Выборочной дисперсией** называется дисперсия, вычисленная по данным выборки:

$$D_B = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_B)^2 \quad \text{или} \quad D_B = \frac{1}{n} \sum_{i=1}^k m_i (x_i - \bar{x}_B)^2.$$

Выборочная дисперсия равна разности между средним значением квадрата вариант и квадратом выборочного среднего:

$$D_B = \overline{X^2} - (\bar{x}_B)^2, \quad \text{где} \quad \overline{X^2} = \frac{1}{n} \sum_{i=1}^k x_i^2 m_i.$$

**Выборочным средним квадратическим отклонением** называется корень квадратный из выборочной дисперсии:

$$\sigma_B = \sqrt{D_B}.$$

**Задача 1.1.1.** По данному распределению выборки найти эмпирическую функцию распределения и построить полигон частот.

$x_i$	1	3	5	7	9
$m_i$	6	11	23	7	3

## Решение

Определим объем выборки  $m = \sum_{i=1}^k m_i = 6 + 11 + 23 + 7 + 3 = 50$ .

Определим относительные частоты вариант  $\omega_i = \frac{m_i}{n}$ .

$x_i$	1	3	5	7	9
$\omega_i$	0,12	0,22	0,46	0,14	0,06

Запишем эмпирическую функцию распределения:

$$F^*(x) = \omega(X < x) = \begin{cases} 0, & x \leq 1; \\ 0,12, & 1 < x \leq 3; \\ 0,34, & 3 < x \leq 5; \\ 0,80, & 5 < x \leq 7; \\ 0,94, & 7 < x \leq 9; \\ 1, & x > 9. \end{cases}$$

Построим полигон частот (рис. 1.1).

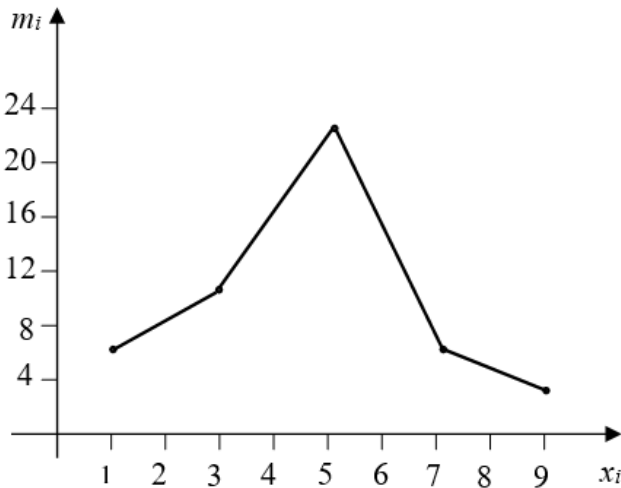


Рис. 1.1



**Задача 1.1.2.** Построить гистограмму частот по данным выборки объема 100 и вычислить числовые характеристики выборки.

$x_i - x_{i+1}$	1–5	5–9	9–13	13–17	17–21
$m_i$	10	20	50	12	8

### Решение

Вычислим относительные частоты по формуле  $\omega_i = \frac{m_i}{n}$  и найдем высоты прямоугольников по формуле  $h_i = \frac{\omega_i}{h}$ , где  $h = 4$ . Вычисления сведем в таблицу.

$x_i - x_{i+1}$	1–5	5–9	9–13	13–17	17–21
$\omega_i$	0,1	0,2	0,5	0,12	0,08
$h_i$	0,025	0,05	0,125	0,03	0,02

Построим гистограмму частот (рис. 1.2).

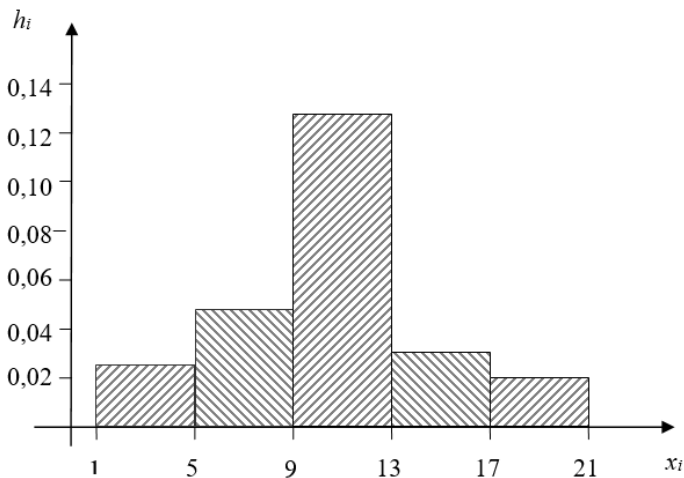


Рис. 1.2

Вычислим числовые характеристики выборки:

$$\begin{aligned}\bar{x}_B &= \frac{1}{n} \sum_{i=1}^k x_i^* m_i = \frac{1}{100} (3 \cdot 10 + 7 \cdot 20 + 11 \cdot 50 + 15 \cdot 12 + 19 \cdot 8) = \\ &= \frac{1}{100} \cdot (30 + 210 + 550 + 180 + 152) = \frac{1122}{100} = 11,22.\end{aligned}$$

$$D_B = \overline{X^2} - (\bar{x}_B)^2.$$

Вычислим  $D_B$  и  $\sigma_B$ :

$$\begin{aligned}\overline{X^2} &= \frac{1}{n} \sum_{i=1}^k (x_i^*)^2 m_i = \frac{1}{100} (9 \cdot 10 + 49 \cdot 20 + 121 \cdot 50 + 225 \cdot 12 + 361 \cdot 8) = \\ &= \frac{1}{100} (90 + 980 + 6050 + 2700 + 2888) = \frac{12\,708}{100} = 127,08.\end{aligned}$$

$$D_B = 127,08 - (11,22)^2 = 1,1916.$$

$$\sigma_B = 1,092.$$

## 1.2. Задачи для аудиторной работы

1.2.1. Даны измерения отклонений от номинала 50 подшипников в мкм:

-1,752;	-0,291;	-0,933;	-0,450;	0,512;
-1,256;	1,701;	0,634;	0,720;	0,490;
-1,531;	-0,433;	1,409;	1,730;	-0,266;
-0,058;	0,248;	-0,095;	-1,488;	-0,361;
0,415;	-1,382;	0,129;	-0,361;	-0,087;
-0,329;	0,086;	0,130;	-0,244;	-0,882;
0,318;	-1,087;	0,899;	1,028;	-1,304;
0,349;	-0,293;	-0,883;	-0,056;	0,757;
-0,059;	-0,539;	-0,078;	0,229;	0,194;
-1,084;	0,318;	0,367;	-0,992;	0,529.

Построить для данной выборки интервальный статистический ряд.

$x_i - x_{i+1}$	-1,75– (-1,25)	-1,25– (-0,75)	-0,75– (-0,25)	-0,25– 0,25	0,25– 0,75	0,75– 1,25	1,25– 1,75
$m_i$	5	8	9	12	9	3	4

1.2.2. Измеряется рост (с точностью до см) 30 наудачу отобран-ных студентов:

178; 160; 154; 183; 155; 153; 167; 186; 163; 155;  
 157; 175; 170; 166; 159; 173; 182; 167; 171; 169;  
 179; 165; 156; 179; 158; 171; 175; 173; 164; 172.

Построить интервальный статистический ряд.

$x_i - x_{i+1}$	150–156	156–162	162–168	168–174	174–180	180–186
$m_i$	4	5	6	7	5	3

1.2.3. По данным выборки объемом 100 найти эмпирическую функцию и построить полигон частот.

$x_i - x_{i+1}$	9–12	12–15	15–18	18–21	21–24	24–27
$m_i$	6	12	33	22	19	8

$$F^*(x) = \begin{cases} 0, & x \leq 9; \\ 0,06, & 9 < x \leq 12; \\ 0,18, & 12 < x \leq 15; \\ 0,51, & 15 < x \leq 18; \\ 0,73, & 18 < x \leq 21; \\ 0,95, & 21 < x \leq 24; \\ 1, & x > 24. \end{cases}$$

1.2.4. Найти эмпирическую функцию распределения и построить полигон частостей по следующим данным.

$x_i$	1	2	3	4	5
$m_i$	4	6	16	26	48

$$F^*(x) = \begin{cases} 0, & x \leq 1; \\ 0,04, & 1 < x \leq 2; \\ 0,1, & 2 < x \leq 3; \\ 0,26, & 3 < x \leq 4; \\ 0,52, & 4 < x \leq 5; \\ 1, & x > 5. \end{cases}$$

1.2.5. Построить гистограмму частот по данным выборки.

$x_i - x_{i+1}$	2-7	7-12	12-17	17-22	22-27
$m_i$	5	10	25	6	4

1.2.6. Построить гистограмму частостей и найти эмпирическую функцию распределения по данным выборки объемом 100.

$x_i - x_{i+1}$	154- 158	158- 162	162- 166	166- 170	170- 174	174- 178	178- 182
$m_i$	10	14	26	28	14	6	2

$$F^*(x) = \begin{cases} 0, & x \leq 154; \\ 0,1, & 154 < x \leq 158; \\ 0,24, & 158 < x \leq 162; \\ 0,5, & 162 < x \leq 166; \\ 0,78, & 166 < x \leq 170; \\ 0,92, & 170 < x \leq 174; \\ 0,98, & 174 < x \leq 178; \\ 1, & x > 178. \end{cases}$$

1.2.7. Найти числовые характеристики по данным выборки:

a)

$x_i$	1	3	6	26
$m_i$	8	40	10	2

$$(\bar{x}_B = 4; \quad D_B = 18,67; \quad \sigma_B = 4,32)$$

б)

$x_i - x_{i+1}$	40,1–40,2	40,2–40,3	40,3–40,4	40,4–40,5	40,5–40,6
$m_i$	7	24	34	26	9

$$(\bar{x}_b = 40,356; \quad D_b = 0,011; \quad \sigma_b = 0,0105)$$

1.2.8. Для проверки оборудования размельчения руды были случайно отобраны и измерены 50 образцов переработанного минерала. Найти выборочное среднее, выборочную дисперсию и выборочное среднее квадратическое отклонение:

0,030; 0,559; 0,407; 2,784; 0,518; 1,185; 1,297;  
 0,614; 0,171; 0,155; 0,081; 30,02; 3,554; 1,155;  
 2,664; 1,889; 0,114; 6,038; 7,815; 0,074; 21,370;  
 0,412; 16,740; 31,820; 0,587; 2,010; 0,558; 0,171;  
 0,894; 4,545; 0,147; 1,642; 0,827; 0,051; 0,486;  
 0,889; 0,340; 0,856; 1,581; 1,474; 2,293; 0,063;  
 1,294; 0,009; 0,114; 1,889; 2,083; 0,138; 2,881;  
 0,114.

$$(\bar{x}_b = 4,67; \quad D_b = 39,25; \quad \sigma_b = 6,26)$$

### 1.3. Задачи для самостоятельной работы

1.3.1. Составить эмпирическую функцию распределения и построить полигон частот по данным выборки.

$x_i$	15	16	17	18	19
$m_i$	1	4	5	6	4

$$F^*(x) = \begin{cases} 0, & x \leq 15; \\ 0,05, & 15 < x \leq 16; \\ 0,25, & 16 < x \leq 17; \\ 0,5, & 17 < x \leq 18; \\ 0,8, & 18 < x \leq 19; \\ 1, & x > 19. \end{cases}$$

1.3.2. Составить эмпирическую функцию распределения и построить гистограмму частот по данным выборки.

$x_i - x_{i+1}$	10–20	20–30	30–40	40–50	50–60	60–70	70–80
$m_i$	1	2	7	18	12	8	2

$$F^*(x) = \begin{cases} 0, & x \leq 10; \\ 0,02, & 10 < x \leq 20; \\ 0,06, & 20 < x \leq 30; \\ 0,20, & 30 < x \leq 40; \\ 0,56, & 40 < x \leq 50; \\ 0,80, & 50 < x \leq 60; \\ 0,96, & 60 < x \leq 70; \\ 1, & x > 70. \end{cases}$$

1.3.3. Интервал движения поездов в метро составляет 2 минуты. Приведены значения случайной величины  $X$  – время ожидания пассажиром поезда. Составить интервальный вариационный ряд и найти среднее время ожидания.

0,000; 0,002; 0,007; 0,025; 0,089; 0,312; 1,068; 1,604; 0,014;  
0,045; 1,747; 1,677; 0,341; 0,952; 0,645; 1,297; 1,981; 0,214;  
1,452; 0,787; 1,654; 0,838; 0,143; 1,317; 0,618; 1,853; 1,555;  
0,653; 1,922; 1,653; 0,617; 0,828; 1,413; 1,030; 1,459; 1,483;  
1,769; 1,265; 1,669; 0,635; 0,787; 1,004; 0,941; 0,612; 1,200;  
1,692; 1,356; 0,908; 1,245; 1,295.

$$(\bar{x}_B = 1,022)$$

1.3.4. Вычислить выборочную дисперсию по данным выборки.

$x_i$	340	360	375	380
$m_i$	20	50	18	12

$$(D_B = 167,29)$$

1.3.5. Вычислить числовые характеристики выборки.

$x_i - x_{i+1}$	10–20	20–30	30–40	40–50	50–60
$m_i$	1	8	10	3	3

$$(\bar{x}_B = 34,6; D_B = 107,84; \sigma_B = 10,35)$$

## Занятие 2.

### Точечные оценки неизвестных параметров распределения

#### 2.1. Краткие теоретические сведения

Пусть изучается СВ  $X$  с законом распределения, зависящим от одного или нескольких параметров. Требуется по выборке, полученной в результате  $n$  испытаний оценить неизвестный параметр  $\theta$ .

Точечной оценкой неизвестного параметра  $\theta$  теоретического распределения называется его приближенное значение, зависящее от данных выборки:

$$\bar{\theta} = \bar{\theta}(x_1, x_2, x_3, \dots, x_n).$$

Точечная оценка должна удовлетворять следующим требованиям:

– оценка должна быть несмещенной, т. е.  $M(\bar{\theta}) = \theta$ ;

– оценка должна быть состоятельной, т. е. она должна сходиться по вероятности к оцениваемому параметру: для  $\forall \varepsilon > 0 \quad \lim_{n \rightarrow \infty} P(|\bar{\theta} - \theta| < \varepsilon) = 1$ ;

– оценка должна быть эффективной: если неизвестный параметр имеет несколько оценок, то в качестве оценки нужно брать оценку с наименьшей дисперсией.

Выборочная средняя  $\bar{x}_v$  является несмещенной и состоятельной оценкой для математического ожидания генеральной совокупности.

Несмещенной и состоятельной оценкой для дисперсии генеральной совокупности является исправленная выборочная дисперсия:

$$D_{\text{и}} = s^2 = \frac{n}{n-1} D_{\text{в}}.$$

**Исправленным** средним квадратическим отклонением называется корень квадратный из исправленной дисперсии:

$$s = \sqrt{D_{\text{и}}}.$$

Для вычисления  $\bar{x}_B$  и  $D_B$  разработано много методов. Одним из наиболее распространенных методов является метод произведений. При вычислении выборочного среднего и выборочной дисперсии поступают следующим образом:

– выбираем «ложный нуль»  $c$ . В качестве «ложного нуля» берется варианта, стоящая посередине вариационного ряда, или варианта, имеющая максимальную частоту;

– переходим к условным вариантам  $U_i$  по формуле  $U_i = \frac{x_i - c}{h}$ ,

где  $h$  – шаг разбиения;

– вычисляем условные моменты 1-го и 2-го порядков:

$$M_1^* = \frac{1}{h} \sum_{i=1}^k U_i m_i; \quad M_2^* = \frac{1}{h} \sum_{i=1}^k U_i^2 m_i;$$

– вычисляем выборочное среднее  $\bar{x}_B$  и выборочную дисперсию  $D_B$ :

$$\bar{x}_B = M_1^* h + c; \quad D_B = \left( M_2^* - (M_1^*)^2 \right) h^2.$$

**Задача 2.1.1.** Методом произведений вычислить выборочную среднюю и выборочную дисперсию по данным выборки.

$x_i$	65	70	75	80	85
$m_i$	2	5	25	15	3

### Решение

В качестве «ложного нуля» возьмем варианту 75,  $c = 75$ . Перейдем к условным вариантам по формуле  $U_i = \frac{x_i - c}{h}$ . Результаты вычислений сведем в таблицу.



$x_i$	$m_i$	$U_i$	$U_i n_i$	$U_i^2 n_i$	$(U_i + 1)^2 m_i$
65	2	-2	-4	8	2
70	5	-1	-5	5	0
75	25	0	0	0	25
80	15	1	15	15	60
85	3	2	6	12	27
$\Sigma$	50	0	12	40	114

Результаты вычислений можно проверить равенством:

$$\sum_{i=1}^k (U_i + 1) m_i = \sum_{i=1}^k m_i + 2 \sum_{i=1}^k m_i U_i + \sum_{i=1}^k U_i^2 m_i;$$

$$114 = 50 + 2 \cdot 12 + 40.$$

Равенство выполняется, следовательно, таблица заполнена верно. Вычислим условные моменты

$$M_1^* = \frac{1}{50} \cdot 12 = 0,24; \quad M_2^* = \frac{1}{50} \cdot 40 = 0,8.$$

Вычислим выборочную среднюю и выборочную дисперсию

$$\bar{x}_B = M_1^* h + C = 0,24 \cdot 5 + 75 = 76,2;$$

$$D_B = \left( M_2^* - (M_1^*)^2 \right) h^2 = (0,8 - 0,0576) \cdot 25 = 18,56.$$

## 2.2. Задачи для аудиторной работы

2.2.1. По данным выборки найти несмещенные оценки для математического ожидания и дисперсии:

$x_i - x_{i+1}$	7,8-8,0	8,0-8,2	8,2-8,4	8,4-8,6	8,6-8,8	8,8-9,0
$m_i$	5	20	80	95	40	10

$$(\bar{x}_B = 8,44; \quad D_B = 0,042)$$

2.2.2. Найти несмещенную оценку для дисперсии по данным выборки:

$x_i$	102	104	108
$m_i$	2	3	5

$$(D_{\text{и}} = 6,93)$$

2.2.3. По данным выборки найти несмещенные оценки для математического ожидания и дисперсии генеральной совокупности:

а) положительные отклонения от номинального размера в партии деталей (в мкм):

17; 21; 8; 20; 23; 18; 22; 20; 17; 12;  
 20; 11; 9; 19; 20; 9; 19; 17; 21; 13;  
 17; 22; 22; 10; 20; 20; 15; 19; 20; 20;  
 13; 21; 21; 9; 14; 11; 19; 18; 23; 19;

$$(\bar{x}_{\text{в}} = 17,2; D_{\text{и}} = 19,7)$$

б) время реакции (в секундах):

8,5; 7,1; 6,7; 6,2; 2,9; 4,4; 6,0; 5,8; 5,4;  
 8,2; 6,9; 6,5; 6,1; 3,8; 6,0; 6,0; 5,6; 5,3;  
 7,7; 6,8; 6,5; 6,1; 4,2; 4,7; 5,6; 5,4; 5,3;  
 7,4; 6,7; 6,4; 6,1; 4,5; 6,0; 5,8; 5,6; 5,1.

$$(\bar{x}_{\text{в}} = 5,9; D_{\text{и}} = 1,3)$$

2.2.4. Даны результаты наблюдений за сроком службы 100 одно-типных станков до выхода за пределы норм точности.

$x_i - x_{i+1}$	20–25	25–30	30–35	35–40	40–45
$m_i$	9	24	35	22	10

Найти несмещенную оценку для дисперсии срока службы.

$$(D_{\text{и}} = 30,14)$$

2.2.5. Приведены результаты измерения диаметра втулок, обрабатываемых автоматом.

$x_i - x_{i+1}$	20,00– 20,04	20,04– 20,08	20,08– 20,12	20,12– 20,16	20,16– 20,20
$m_i$	8	18	45	20	9

Найти оценки для математического ожидания и дисперсии.

$$(\bar{x}_B = 20,1016; D_H = 0,002)$$

### 2.3. Задачи для самостоятельной работы

2.3.1. Даны отклонения напряжений от номинала (мВ):

$x_i - x_{i+1}$	0,00– 0,02	0,02– 0,04	0,04– 0,06	0,06– 0,08	0,08– 0,10	0,10– 0,12	0,12– 0,14	0,14– 0,16
$m_i$	9	15	29	35	32	19	8	3

Найти оценки для математического ожидания и дисперсии.

$$(\bar{x}_B = 0,05; D_H = 0,02)$$

2.3.2. Дана урожайность ржи на различных участках поля.

Урожайность, ц/га	9–12	12–15	15–18	18–21	21–24	24–27
Количество участков	6	12	33	22	19	8

Найти оценку для средней урожайности всего поля.

$$(\bar{x}_B = 18,3)$$

2.3.3. По данной выборке найти оценки для математического ожидания и дисперсии генеральной совокупности.

$x_i$	12	14	16	18	20	22
$m_i$	5	15	50	16	10	4

$$(\bar{x}_B = 16,46; D_H = 4,92)$$

## Занятие 3. Интервальные оценки

### 3.1. Краткие теоретические сведения

Пусть  $\theta^* = \theta^*(x_1, \dots, x_n)$  – функция выборки. Это есть случайная величина, называемая **статистикой**.

**Интервальной** называют оценку, которая определяется случайным интервалом  $(\theta_1^*, \theta_2^*)$ ,  $\theta_1^* < \theta_2^*$ . В качестве интервальной оценки используются доверительные интервалы.

**Доверительным интервалом** для неизвестного параметра  $\theta$ , называется случайный интервал  $(\theta_1^*, \theta_2^*)$ , который с заданной вероятностью  $\gamma$  (надежностью) накрывает неизвестный параметр  $\theta$ .

Если исследуемая СВ распределена по нормальному закону с известным средним квадратическим отклонением  $\sigma$ , то доверительный интервал для математического ожидания определяется неравенством:

$$\bar{x}_B - t_\gamma \frac{\sigma}{\sqrt{n}} < a < \bar{x}_B + t_\gamma \frac{\sigma}{\sqrt{n}}, \quad (3.1)$$

где  $\delta = t_\gamma \frac{\sigma}{n}$  – точность оценки;

$n$  – объем выборки;

$t_\gamma$  – значение аргумента функции Лапласа, при котором

$$\Phi(t_\gamma) = \frac{\gamma}{2}.$$

Если среднее квадратическое отклонение неизвестно, то доверительный интервал для математического ожидания исследуемой СВ определяется неравенством

$$\bar{x}_B - t_{\gamma,n} \frac{s}{\sqrt{n}} < a < \bar{x}_B + t_{\gamma,n} \frac{s}{\sqrt{n}}, \quad \text{где } s = \sqrt{D_n}. \quad (3.2)$$

Значения  $t_{\gamma,n}$  находят по таблице прил. 5 по заданным  $n$  и  $\gamma$ .

Число  $\delta = t_{\gamma,n} \frac{s}{\sqrt{n}}$  называют точностью оценки математического ожидания.

Доверительный интервал для среднего квадратического отклонения исследуемой СВ определяется неравенством

$$s q_1 < \sigma < s q_2. \quad (3.3)$$

Значения  $q_1$  и  $q_2$  находятся по таблице прил. 6 по заданным  $\gamma$  и  $n$ .

**Задача 3.1.1.** Найти доверительный интервал для оценки с надежностью  $\gamma = 0,99$  неизвестного математического ожидания нормально распределенного признака  $X$ , если известно  $\sigma = 4$ , а по данным выборки объемом 100 вычислено  $\bar{x}_b = 12,4$ .

### Решение

Так как известно среднее квадратическое отклонение СВ, то для определения доверительного интервала для математического ожидания воспользуемся неравенством (3.1). Определим значение  $t_\gamma$ :  $\Phi(t_\gamma) = \frac{\gamma}{2} = \frac{0,99}{2} = 0,495 \Rightarrow t_\gamma = 2,58$ . Подставим в неравенство (3.1):

$$12,4 - 2,58 \frac{4}{10} < a < 12,4 + 2,58 \frac{4}{10}; \quad 11,08 < a < 13,432.$$

**Задача 3.1.2.** Для исследования нормально распределенной СВ извлечена выборка объемом 25.

$x_i - x_{i+1}$	10–20	20–30	30–40	40–50	50–60
$m_i$	1	8	10	3	3

Найти с надежностью  $\gamma = 0,95$  доверительные интервалы для математического ожидания и среднего квадратического отклонения исследуемой СВ.

## Решение

По данным выборки методом произведений определим  $\bar{x}_B$  и  $s$ .

$x_i^*$	$m_i$	$U_i$	$U_i m_i$	$U_i^2 m_i$	$(U_i + 1)^2 m_i$
15	1	-2	-2	4	1
25	8	-1	-8	8	0
35	10	0	0	0	10
45	3	1	3	3	12
55	3	2	6	12	27
$\Sigma$	25	0	-1	27	50

Проверка:

$$50 = 25 + 2(-1) + 27; \quad 50 = 50.$$

$$M_1^* = \frac{-1}{25} = -0,04; \quad M_2^* = \frac{27}{25} = 1,08.$$

$$\bar{x}_B = 35 + (-0,04) \cdot 10 = 34,6; \quad D_B = (1,08 - (0,04)^2) \cdot 100 = 107,84;$$

$$D_{\text{и}} = \frac{n}{n-1} D_B = \frac{25}{24} \cdot 107,84 = 112,33; \quad s = \sqrt{D_{\text{и}}} = 10,6.$$

Для определения доверительного интервала для математического ожидания воспользуемся неравенством (3.2):

$$t_{\gamma, n} = t(0,95; 24) = 2,064;$$

$$34,6 - 2,064 \frac{10,6}{5} < a < 34,6 + 2,064 \frac{10,6}{5};$$

$$34,6 - 4,38 < a < 34,6 + 4,38;$$

$$30,22 < a < 38,98.$$

Для определения доверительного интервала для среднего квадратического отклонения воспользуемся неравенством (3.3):

$$q_1 = 0,781; \quad q_2 = 1,391;$$

$$10,6 \cdot 0,781 < \sigma < 10,6 \cdot 1,391;$$

$$8,28 < \sigma < 14,74.$$

### 3.2. Задачи для аудиторной работы

3.2.1. Для определения привеса рыбы за год в одном из рыбхозов проводились выборочные исследования. Разводимые в пруду карпы взвешивались и отпускались обратно.

Результаты 100 таких измерений показали, что годовой привес рыбы в среднем составил 200 г, а дисперсия – 320 г. Найти с надежностью 0,95 доверительный интервал для годового привеса рыбы  $\Delta P$ .

$$(196,49 < \Delta P < 203,51)$$

3.2.2. Одним и тем же прибором со средним квадратическим отклонением случайных ошибок измерения  $\sigma = 40$  м проведено пять различных измерений расстояния. Найти с надежностью 0,95 доверительный интервал для оценки истинного расстояния, если среднее всех проведенных измерений  $\bar{x}_B = 2000$  м.

$$(1964,94 < a < 2035,06)$$

3.2.3. Выборка из большой партии электроламп содержит 100 ламп. Средняя продолжительность горения ламп в выборке равна 1000 ч. Найти с надежностью 0,95 доверительный интервал для продолжительности горения ламп всей партии, если известно, что среднее квадратическое отклонение продолжительности горения лампы  $\sigma = 40$  ч.

$$(992,16 < a < 1007,84)$$

3.2.4. Найти минимальный объем выборки, при котором с надежностью 0,925 точность оценки математического ожидания нормально распределенной СВ равна 0,2, если среднее квадратическое отклонение  $\sigma = 1,5$ .

$$(n = 179)$$

3.2.5. По данным выборки найти доверительный интервал, с надежностью 0,99 накрывающий среднее квадратическое отклонение.

$x_i$	0,1	0,2	0,3	0,4	0,5
$m_i$	2	4	7	6	1

$$(0,077 < \sigma < 0,583)$$

3.2.6. Из генеральной совокупности СВ, распределенной по нормальному закону, выбрано 100 значений СВ. Найти доверительные интервалы для математического ожидания и среднего квадратического отклонения с надежностью 0,95.

$x_i - x_{i+1}$	100–120	120–140	140–160	160–180	180–200
$m_i$	17	40	32	8	3

$$\left( \begin{array}{l} 125,78 < a < 127,42 \\ 3,64 < \sigma < 4,82 \end{array} \right)$$

3.2.7. С целью определения средней суммы  $Q$  вкладов в банке произведена выборка.

Сумма, млн руб.	10–30	30–50	50–70	70–90	90–110	110–130
$m_i$	1	3	10	30	60	7

Найти границы среднего вклада с надежностью 0,95.

$$(86,45 < Q < 93,37)$$

### 3.3. Задачи для самостоятельной работы

3.3.1. Станок-автомат штампует валики. По выборке объемом 100 вычислены выборочная средняя  $\bar{x}_в = 12,5$  и  $s = 2,1$ . Найти с надежностью 0,95 доверительные интервалы для математического ожидания и среднего квадратического отклонения.

$$\left( \begin{array}{l} 12,08 < a < 12,92 \\ 1,84 < \sigma < 2,44 \end{array} \right)$$

3.3.2. Найти минимальный объем выборки, при котором с надежностью 0,975 точность оценки математического ожидания  $a$  по выборочной средней равна  $\delta = 0,3$ , если  $\sigma = 1,2$ .

$$(n = 81)$$



3.3.3. Найти с надежностью 0,99 доверительные интервалы для математического ожидания и среднего квадратического отклонения по данным выборки:

а)

$x_i$	12	14	16	18	20	22
$m_i$	5	15	50	16	10	4

$$\left( 15,88 < a < 17,04 \right)$$

$$\left( 1,88 < \sigma < 2,71 \right)$$

б)

$x_i - x_{i+1}$	46– 50	50– 54	54– 58	58– 62	62– 66	67– 70	70– 74	74– 78	78– 82	82– 66
$m_i$	2	4	6	8	12	30	18	8	7	5

$$\left( 65,99 < a < 70,25 \right)$$

$$\left( 6,84 < \sigma < 9,86 \right)$$

## Занятие 4. Статистическая проверка гипотез. Критерии согласия Пирсона и Колмогорова

### 4.1. Краткие теоретические сведения

**Статистической** называется гипотеза о предполагаемом виде неизвестного распределения СВ или о значениях параметров известного вида распределения. **Нулевой гипотезой**  $H_0$  называется выдвинутая гипотеза. *Конкурирующей (альтернативной)* называется гипотеза, которая противоречит нулевой гипотезе.

При проверке статистической гипотезы могут быть допущены ошибки двух родов. **Ошибка первого рода** – будет отклонена верная гипотеза. **Ошибка второго рода** – будет принята неверная гипотеза.

Вероятность допустить ошибку первого рода называется *уровнем значимости*. Для проверки статистической гипотезы используют специальную статистику, которая называется критерием. По рассчитанному значению критерия определяют принимать или отвергать нулевую гипотезу. **Критерий согласия** – это проверка гипотезы о виде распределения СВ.

Основными критериями согласия являются критерии Пирсона  $\chi^2$  и Колмогорова.

При проверке гипотезы с помощью критерия Пирсона поступают следующим образом:

- из генеральной совокупности извлекают выборку объемом  $n$ ;
- по выборке вычисляют  $\bar{x}_B$  и  $\sigma_B$ ;
- переходят к нормированной СВ по формуле  $U_i = \frac{x_i - \bar{x}_B}{\sigma_B}$ ;
- находят вероятности попадания в интервал  $(U_i, U_{i+1})$ ,  
 $P_i = \Phi(U_{i+1}) - \Phi(U_i)$ ;
- вычисляют теоретические частоты  $m'_i = nP_i$ ;
- вычисляют статистику Пирсона  $\chi^2_{\text{набл}} = \sum_{i=1}^k \frac{(m_i - m'_i)^2}{m'_i}$ ;

– из таблицы критических точек распределения Пирсона (прил. 3) по уровню значимости  $\alpha$  и числу степеней свободы  $\nu = k - 1 - r$  определяют  $\chi_{кр}^2$ , где  $k$  – число интервалов в вариационном ряде;  $r$  – количество параметров закона распределения, которые оцениваются по выборке (для нормального закона  $r = 2$ );

– если  $\chi_{набл}^2 \leq \chi_{кр}^2$ , то нет необходимости отвергать нулевую гипотезу, т. е. эмпирические и теоретические частоты согласуются;

– если  $\chi_{набл}^2 > \chi_{кр}^2$ , то гипотеза отвергается, т. е. расхождение между теоретическими и эмпирическими частотами существенно.

Если исследуется дискретная СВ, распределенная по нормальному закону, то теоретические вероятности определяются по формуле  $p_i = \frac{h}{\sigma_B} \varphi(U_i)$ , где  $h$  – шаг,  $U_i = \frac{x_i - \bar{x}_B}{\sigma_B}$ ,  $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ .

**Задача 4.1.1.** Пользуясь критерием Пирсона, при уровне значимости  $\alpha = 0,05$  проверить, согласуется ли гипотеза о нормальном распределении генеральной совокупности с данными выборки.

$x_i - x_{i+1}$	3–8	8–13	13–18	18–23	23–28	28–33	33–38
$m_i$	6	8	15	40	16	8	7

### Решение

По данным выборки методом произведений вычислим  $\bar{x}_B$  и  $D_B$ .

$x_i^*$	$m_i$	$U_i$	$U_i m_i$	$U_i^2 m_i$	$(U_i + 1)^2 m_i$
5,5	6	–3	–18	54	24
10,5	8	–2	–16	32	8
15,5	15	–1	–15	15	0
20,5	40	0	0	0	40
25,5	16	1	16	16	64
30,5	8	2	16	32	72
35,5	7	3	21	63	112
$\Sigma$	100	0	4	212	320

Проверка:

$$320 = 100 + 8 + 212 = 320.$$

$$M_1^* = \frac{4}{100} = 0,04; \quad M_2^* = \frac{212}{100} = 2,12;$$

$$\bar{x}_B = M_1^* h + c = 20,5 + 0,04 \cdot 5 = 20,7;$$

$$D_B = \left( M_2^* - (M_1^*)^2 \right) h^2 = (2,12 - 0,0016) \cdot 25 = 52,96;$$

$$\sigma_B = \sqrt{D_B} = 7,28.$$

Вычислим вероятности попадания в интервалы.

$x_i$	$x_i - \bar{x}_B$	$\frac{x_i - \bar{x}_B}{\sigma_B}$	$\Phi\left(\frac{x_i - \bar{x}_B}{\sigma_B}\right)$
$-\infty$	$-\infty$	$-\infty$	0,5
8	-12,7	-1,74	0,4591
13	-7,7	-1,06	0,3554
18	-2,7	-0,37	0,1443
23	2,3	0,32	0,1255
28	7,3	1,00	0,3413
33	12,3	1,69	0,4545
$+\infty$	$+\infty$	$+\infty$	0,5

$$P_1 = -0,4591 + 0,5 = 0,0409;$$

$$P_2 = -0,3554 + 0,4591 = 0,1037;$$

$$P_3 = -0,1443 + 0,3554 = 0,2111;$$

$$P_4 = 0,1255 + 0,1443 = 0,2698;$$

$$P_5 = 0,3413 - 0,1255 = 0,2158;$$

$$P_6 = 0,4545 - 0,3413 = 0,1132;$$

$$P_7 = 0,5 - 0,4545 = 0,0455.$$

Вычислим  $\chi^2_{\text{набл}}$ .

$m_i$	$P_i$	$m'_i = nP_i$	$m_i - m'_i$	$(m_i - m'_i)^2$	$(m_i - m'_i)^2 / m'_i$
6	0,0409	4,09	1,91	3,648	0,892
8	0,1037	10,37	-2,37	5,617	0,542
15	0,2111	21,11	-6,11	37,332	1,768
40	0,2698	26,98	13,02	169,520	6,283
16	0,2158	21,58	-5,58	31,136	1,443
8	0,1132	11,32	-3,32	11,022	0,974
7	0,0455	4,55	2,45	6,003	1,319
					$\chi^2_{\text{набл}} = 13,221$

Определим число степеней свободы:

$$\nu = k - 1 - r = 7 - 1 - 2 = 4 .$$

По уровню значимости  $\alpha = 0,05$  и числу степеней свободы  $\nu = 4$  найдем критическую точку правосторонней критической области распределения Пирсона (прил. 3):

$$\chi^2_{\text{кр}} = \chi^2(0,05; 4) = 9,5.$$

Так как  $\chi^2_{\text{набл}} > \chi^2_{\text{кр}}$ , то гипотеза о нормальном распределении совокупности отвергается.

**Критерий согласия Колмогорова** применяется для проверки гипотезы о законе распределения непрерывной СВ. Для статистической проверки гипотезы с помощью критерия согласия Колмогорова поступают следующим образом:

- выбирают из генеральной совокупности выборку;
- по выборке составляют эмпирическую функцию распределения  $F^*(x)$ ;
- записывают теоретическую функцию распределения  $F(x)$ ;
- вычисляют величину  $D = \max |F^*(x) - F(x)|$ ;

– вычисляют статистику Колмогорова  $\lambda = D\sqrt{n}$ , где  $n$  – объем выборки. СВ  $\lambda$  имеет функцию распределения

$$K(x) = \sum_{i=-\infty}^{\infty} (-1)^i e^{-2i^2 x^2}, \quad x > 0,$$

которая называется функцией Колмогорова;

– находим  $\lambda_{\alpha}$  по уровню значимости  $\alpha$  (прил. 7);

– если  $\lambda \geq \lambda_{\alpha}$ , то гипотеза о законе распределения СВ отклоняется;

– если  $\lambda < \lambda_{\alpha}$ , то нет оснований отклонять нулевую гипотезу.

Рассмотрим применение критерия Колмогорова на примере.

**Задача 4.1.2.** Проверить по критерию Колмогорова гипотезу о нормальном распределении СВ по данным выборки при уровне значимости  $\alpha = 0,05$ .

$x_{i-1}-x_i$	0–0,5	0,5–1	1–1,5	1,5–2	2–2,5	2,5–3	3–3,5	3,5–4
$m_i$	17	11	9	8	2	1	1	1

### Решение

Вычислим выборочную среднюю  $\bar{x}_B$  и исправленное среднее квадратическое отклонение  $s$ .

$$\bar{x}_B = \frac{1}{n} \sum_{i=1}^n x_i m_i = \frac{1}{50} (0,25 \cdot 17 + 0,75 \cdot 11 + 1,25 \cdot 9 + 1,75 \cdot 8 +$$

$$+ 2,25 \cdot 2 + 2,75 + 3,25 + 3,75) = 1,04;$$

$$\overline{x^2} = \frac{1}{n} \sum_{i=1}^n x_i^2 m_i = 1,7625;$$

$$D_B = \overline{x^2} - (\bar{x}_B)^2 = 1,7625 - 1,04^2 = 0,6809;$$

$$D_n = \frac{n}{n-1} D_B = \frac{50}{49} \cdot 0,6809 = 0,6948;$$

$$s = \sqrt{D_n} = \sqrt{0,6948} = 0,834.$$

Тогда теоретическая функция распределения в предположении, что СВ распределена по нормальному закону, имеет вид:

$$F(x) = \frac{1}{2} + \Phi\left(\frac{x - \bar{x}_e}{s}\right) = \frac{1}{2} + \Phi\left(\frac{x - 1,04}{0,834}\right),$$

где  $\Phi(x)$  – функция Лапласа.

Эмпирическую функцию распределения определим по формуле:

$$F^*(x) = \frac{m_x}{n},$$

где  $m_x$  – сумма частот вариант меньших  $x$ .

$$F^*(x) = \begin{cases} 0, & x \leq 0; \\ 0,34, & 0 < x \leq 0,5; \\ 0,56, & 0,5 < x \leq 1; \\ 0,74, & 1 < x \leq 1,5; \\ 0,90, & 1,5 < x \leq 2; \\ 0,94, & 2 < x \leq 2,5; \\ 0,96, & 2,5 < x \leq 3; \\ 0,98, & 3,0 < x \leq 3,5; \\ 1, & x > 3,5. \end{cases}$$

Вычислим величину  $D = \max |F^*(x) - F(x)|$ .

$x_i$	$x_i - \bar{x}_B$	$\frac{x_i - \bar{x}_B}{s}$	$\Phi\left(\frac{x_i - \bar{x}_B}{s}\right)$	$F(x_i) = \frac{1}{2} + \Phi\left(\frac{x_i - \bar{x}_B}{s}\right)$	$F^*(x_i)$	$ F(x_i) - F^*(x) $
0	-1,04	-1,25	-0,3944	0,1056	0	0,1056
0,5	-0,55	-0,65	-0,2422	0,2578	0,34	0,0822
1,0	-0,04	-0,05	-0,0199	0,4801	0,56	0,0799
1,5	0,46	0,55	0,2088	0,7088	0,74	0,0312
2,0	0,96	1,15	0,3749	0,8749	0,90	0,0251
2,5	1,46	1,75	0,4599	0,9599	0,94	0,0199
3,0	1,96	2,35	0,4908	0,9906	0,96	0,0306
3,5	2,46	2,95	0,4985	0,9985	0,98	0,0185
4,0	2,96	3,55	0,4999	0,9999	1	0,0001

$$D = 0,1056.$$

Вычислим статистику Колмогорова:

$$\lambda = D\sqrt{n} = \sqrt{50} \cdot 0,1056 = 0,747.$$

По уровню значимости  $\alpha = 0,05$  найдем по таблице (прил. 7)  $\lambda_{\alpha} = 1,358$ .

Т. к.  $\lambda < \lambda_{\alpha}$ , то нет оснований отвергать гипотезу о нормальном распределении.

## 4.2. Задачи для аудиторной работы

4.2.1. Используя критерий Пирсона, при уровне значимости  $\alpha$  установить случайно или значимо расхождение между теоретическими и эмпирическими частотами, которые вычислены исходя из гипотезы о нормальном распределении СВ.

а)  $\alpha = 0,01$

$m_i$	8	16	40	72	36	18	10
$m'_i$	6	18	36	76	39	18	7

(Случайно)

б)  $\alpha = 0,05$

$m_i$	5	10	20	8	7
$m'_i$	6	14	18	7	5

(Случайно)

4.2.2. Используя критерий Пирсона, при уровне значимости  $\alpha = 0,05$  проверить, согласуется ли гипотеза о нормальном распределении совокупности с данными выборки.

а)

$x_i - x_{i-1}$	-20-(-10)	-10-0	0-10	10-20	20-30	30-40	40-50
$m_i$	20	47	80	89	40	16	8

(Согласуется)



б)

$x_i$	0,3	0,5	0,7	0,9	1,1	1,3	1,5	1,7	1,9	2,1	2,3
$m_i$	6	9	26	25	30	26	21	24	20	8	5

(Согласуется)

4.2.3. Наблюдения за межремонтными интервалами  $T$  (в месяцах) работы зерноуборочного комплекса дали результаты:

0,000; 0,001; 0,003; 0,012; 0,044; 0,156; 0,534;  
 0,802; 0,007 0,822; 0,873; 0,838; 0,170; 0,476;  
 0,322; 0,648; 0,991; 0,107; 0,726; 0,393; 0,827  
 0,419; 0,071; 0,659; 0,309; 0,927; 0,778; 0,327;  
 0,961; 0,826; 0,308; 0,414; 0,707; 0,515; 0,729;  
 0,742; 0,884; 0,632; 0,835; 0,318; 0,394; 0,502;  
 0,471; 0,306; 0,600; 0,846; 0,678; 0,454; 0,623;  
 0,648.

Проверить при уровне значимости  $\alpha = 0,01$  с помощью критерия Колмогорова гипотезу о показательном распределении совокупности.

(Согласуется)

4.2.4. Даны результаты измерения 1000 деталей.

$x_i - x_{i+1}$	97,25- 98,25	98,25- 98,75	98,75- 99,25	99,25- 99,75	99,75- 100,25	100,25- 100,75	100,75- 101,25	101,25- 101,75	101,75- 102,25	102,25- 102,75
$m_i$	21	47	87	158	181	201	142	97	41	25

При уровне значимости  $\alpha = 0,05$  проверить с помощью критерия Колмогорова, согласуются ли данные выборки с гипотезой о нормальном распределении.

(Не согласуются)

### 4.3. Задачи для самостоятельной работы

4.3.1. Используя критерий Пирсона, при уровне значимости  $\alpha = 0,05$  установить, случайно или значимо расхождение между эмпирическими ( $m_i$ ) и теоретическими ( $m'_i$ ) частотами, которые

вычислены в предположении, что генеральная совокупность распределена по нормальному закону.

$m_i$	14	18	32	70	20	36	10
$m'_i$	10	24	34	80	18	22	12

(Значимо)

4.3.2. Используя критерий Пирсона, при уровне значимости  $\alpha = 0,05$  проверить, согласуется ли гипотеза о нормальном распределении совокупности с данными выборки.

$x_i - x_{i+1}$	6-16	16-26	26-36	36-46	46-56	56-66	66-76	76-86
$m_i$	8	7	16	35	15	8	6	5

(Не согласуется)

4.3.3. При уровне значимости  $\alpha = 0,05$  с помощью критерия Колмогорова проверить, согласуется ли гипотеза о нормальном распределении СВ с данными выборки.

$x_i - x_{i-1}$	-20-(-10)	-10-0	0-10	10-20	20-30	30-40	40-50
$m_i$	20	47	80	89	40	16	8

(Согласуется)

## Занятие 5.

### Выборочный коэффициент корреляции и его свойства. Проверка гипотезы о равенстве нулю коэффициента корреляции

#### 5.1. Краткие теоретические сведения

Для вычисления выборочного коэффициента корреляции данные представляются в виде корреляционной таблицы. Корреляционная таблица представляет собой таблицу следующего вида: в первой строке записаны наблюдаемые значения СВ  $X$ , в первом столбце записаны наблюдаемые значения СВ  $Y$ , на пересечении  $i$ -й строки и  $j$ -го столбца записывается частота  $m_{ij}$  появления пары  $(y_i, x_j)$ . В последнем столбце записывается частота появления варианты  $y_i$ , в последней строке – частота появления варианты  $x_j$ , на пересечении последней строки и последнего столбца записывается суммарное количество наблюдений. Корреляционная таблица имеет следующий вид.

$Y \backslash X$	$x_1$	$x_2$	$x_3$	...	$x_k$	$n_y$
$y_1$	$m_{11}$	$m_{12}$	$m_{13}$	...	$m_{1k}$	$n_{y1}$
$y_2$	$m_{21}$	$m_{22}$	$m_{23}$	...	$m_{2k}$	$n_{y2}$
...	...	...	...	...	...	...
$y_\ell$	$m_{\ell 1}$	$m_{\ell 2}$	$m_{\ell 3}$	...	$m_{\ell k}$	$n_{y\ell}$
$n_x$	$n_{x1}$	$n_{x2}$	$n_{x3}$	...	$n_{xk}$	$n$

Основной оценкой тесноты связи между случайными величинами  $X$  и  $Y$  служит выборочный коэффициент корреляции  $r_b$ , который определяется так:

$$r_b = \frac{\overline{XY} - \overline{x_b} \cdot \overline{y_b}}{\sigma_x \cdot \sigma_y},$$

где  $\overline{XY}$  – среднее арифметическое произведений значений СВ  $X$ ,  $Y$ .

Свойства выборочного коэффициента корреляции аналогичны свойствам коэффициента корреляции между СВ:

1.  $-1 \leq r_b \leq 1$ .

2. Если переменные  $X$  и  $Y$  умножить на одно и то же число, то коэффициент корреляции не изменится.

3. Если  $r_e = \pm 1$ , то корреляционная связь между значениями  $X$  и  $Y$  представляет собой линейную функциональную зависимость.

Для вычисления выборочного коэффициента корреляции применяется формула:

$$r_b = \frac{\sum_i \sum_j x_i y_j m_{ij} - n \overline{x_b} \overline{y_b}}{n \sigma_x \sigma_y}. \quad (5.1)$$

Если  $r_b = 0$ , то между наблюдаемыми значениями  $X$  и  $Y$  корреляционная зависимость отсутствует, чем ближе к единице приближается модуль коэффициента корреляции, тем теснее связь между переменными  $X$  и  $Y$ . Т. к. выборочный коэффициент корреляции вычисляется по данным выборки, то в отличие от коэффициента корреляции генеральной совокупности  $r_b$  является случайной величиной. Если  $r_b \neq 0$ , то возникает вопрос, объясняется ли это действительно существующей связью между СВ  $X$  и  $Y$  или вызвано случайными факторами. Для выяснения этого вопроса проверяется гипотеза  $H_0$  о равенстве нулю коэффициента корреляции  $r$  генеральной совокупности.

Для того чтобы при уровне значимости  $\alpha$  проверить нулевую гипотезу о равенстве нулю коэффициента корреляции генеральной двумерной нормальной совокупности, вычисляют статистику

$$T_{\text{набл.}} = \frac{r_b \sqrt{n-2}}{\sqrt{1-r_b^2}}$$

и по таблице критических точек распределения Стьюдента (прил. 4) по уровню значимости  $\alpha$  и числу степеней свободы  $\nu = n - 2$  находят критическую точку двусторонней критической области

$t_{кр} = t\left(\frac{\alpha}{2}, \nu\right)$ . Если  $|T_{набл.}| < t_{кр}$  – нет оснований отвергать нулевую гипотезу, т. е.  $r = 0$ ; если  $|T_{набл.}| > t_{кр}$  – нулевую гипотезу отвергают, т. е.  $r \neq 0$ . Рассмотрим вычисление выборочного коэффициента корреляции и проверку гипотезы о равенстве нулю коэффициента корреляции генеральной совокупности на примере.

**Задача 5.1.1.** По данной корреляционной таблице вычислить выборочный коэффициент корреляции и при уровне значимости  $\alpha = 0,05$  проверить гипотезу о равенстве нулю коэффициента корреляции генеральной совокупности.

$y_j \backslash x_i$	12,5	17,5	22,5	27,5	$n_y$
20,5	1				1
21,5		2			2
22,5		1	2		3
23,5			3	3	6
24,5				8	8
$n_x$	1	3	5	11	20

### Решение

Вычислим компоненты, входящие в формулу (5.1), для вычисления  $r_B$ .

$$\sum x_i n_i = 12,5 \cdot 1 + 17,5 \cdot 3 + 22,5 \cdot 5 + 27,5 \cdot 11 = 480;$$

$$\sum y_j n_j = 468;$$

$$\sum x_i^2 n_i = 11925; \quad \sum y_j^2 n_j = 10979;$$

$$\begin{aligned} \sum \sum x_i y_j n_{ij} &= 20,5 \cdot 12,5 + 21,5 \cdot 17,5 \cdot 2 + 22,5 \cdot 17,5 + \\ &+ 22,5 \cdot 22,5 \cdot 2 + 23,5 \cdot 22,5 \cdot 3 + 23,5 \cdot 27,5 \cdot 3 + \\ &+ 24,5 \cdot 27,5 \cdot 8 = 11330; \end{aligned}$$

$$\bar{x}_B = \frac{1}{n} \sum x_i n_i = \frac{480}{20} = 24;$$

$$\bar{y}_B = \frac{1}{n} \sum y_j n_j = \frac{468}{20} = 23,4;$$

$$\sigma_x = \sqrt{\frac{1}{n} \sum x_i^2 n_i - (\bar{x}_B)^2} = \sqrt{596,25 - 576} = \sqrt{20,25} = 4,5;$$

$$\sigma_y = \sqrt{\frac{1}{n} \sum y_j^2 n_j - (\bar{y}_B)^2} = \sqrt{548,95 - 547,56} = \sqrt{1,39} = 1,18.$$

Вычислим выборочный коэффициент корреляции:

$$r_e = \frac{\sum \sum x_i y_j m_{ij} - n \bar{x}_B \bar{y}_B}{n \sigma_x \sigma_y} = \frac{11330 - 20 \cdot 24 \cdot 23,4}{20 \cdot 4,5 \cdot 1,18} = 0,923.$$

Проверим гипотезу о равенстве нулю коэффициента корреляции генеральной совокупности. Вычислим:

$$T_{\text{набл.}} = \frac{r_B \sqrt{n-2}}{\sqrt{1-r_B^2}} = \frac{0,923 \cdot \sqrt{18}}{\sqrt{1-0,923^2}} = 10,25.$$

По таблице критических точек распределения Стьюдента (прил. 4) по уровню значимости  $\alpha = 0,05$  и числу степеней свободы  $\nu = n - 2 = 18$  найдем  $t_{\text{кр}} = t(0,025; 18) = 2,101$ .

Так  $|T_{\text{набл.}}| > t_{\text{кр}}$ , то гипотеза о равенстве нулю коэффициента корреляции генеральной совокупности отвергается, т. е. выбранный коэффициент корреляции значим.

## 5.2. Задачи для аудиторной работы

5.2.1. Определить тесноту связи и значимость общего веса  $X$  (г) растения и веса  $Y$  (г) его семян на основании данных.

$x_i$	40	50	60	70	80	90	100
$y_i$	20	25	28	30	35	40	45

$$(r_B = 0,992; T_{\text{набл.}} = 17,6; \text{значим})$$

5.2.2. Для исследования влияния объема капиталовложений  $X$  (млрд руб.) на полученную годовую прибыль  $Y$  (млрд руб.) была собрана статистика по 20 крупным предприятиям, которая сведена в корреляционную таблицу.

$Y \backslash X$	0–10	10–20	20–30	30–40	40–50	$n_y$
1,5–2,5	1					1
2,5–3,5	2	5	2			9
3,5–4,5		3	3	2		8
4,5–5,5					2	2
$n_x$	3	8	5	2	2	20

Вычислить выборочный коэффициент корреляции.

$$(r_B = 0,782)$$

5.2.3. По выборке объема  $n = 100$ , извлеченной из двумерной нормальной совокупности  $(X, Y)$ , вычислить выборочный коэффициент корреляции и при уровне значимости  $\alpha = 0,05$  проверить гипотезу о равенстве нулю коэффициента корреляции генеральной совокупности.

$Y \backslash X$	10	15	20	25	30	35	$n_y$
35	5	1					6
45		6	2				8
55			5	40	5		50
65			2	8	7		17
75				4	7	8	19
$n_x$	5	7	9	52	19	8	100

$$(r_B = 0,817; T_{\text{набл}} = 14,03; r \neq 0)$$

5.2.4. Определить тесноту связи между себестоимостью продукции  $Y$  (тыс. руб.) и количеством выпускаемой продукции  $X$  (тыс. штук) по данным 7 предприятий.

$X$	2	3	4	5	6	7	8
$Y$	2	1,9	2,2	2,4	2,3	2,5	2,5

Выяснить значимость выборочного коэффициента корреляции при  $\alpha = 0,05$ .

$$(r_B = 0,925; T_{\text{набл}} = 5,44; \text{значим})$$

5.2.5. Для выяснения зависимости урожайности сельхозкультур от почвенной влаги были исследованы 20 одинаковых участков земли в пойме реки ( $X$  – расстояние участка от реки;  $Y$  – урожайность).

$Y \backslash X$	0,4	0,8	1,0	1,2	1,8	2,0	$n_y$
3,0	2	3					5
3,5		4	2	1			7
4,5			1	2	2		5
5,0					2	1	3
$n_x$	2	7	3	3	4	1	20

Вычислить выборочный коэффициент корреляции.

$$(r_B = 0,871)$$

### 5.3. Задачи для самостоятельной работы

5.3.1. При отладке токарного станка были измерены погрешности обработки  $X$  (мкм) для разных диаметров обрабатываемых деталей  $Y$  (см).

$X$	0,5	1	1,5	2	2,5	3	3,5	4	4,5	5
$Y$	3	3	4	4	4	5	5	5	6	8

Найти выборочный коэффициент корреляции между  $X$  и  $Y$  и определить его значимость при  $\alpha = 0,05$ .

$$(r_B = 0,921; T_{\text{набл}} = 6,68; \text{значим})$$



5.3.2. Для изучения надежности машин был собран статистический материал зависимости времени непрерывной работы  $Y$  (в месяцах) и количества предшествующих ремонтов  $X$ .

$Y \backslash X$	0	1	2	3	$n_y$
2–6			1	2	3
6–10		1	3	1	5
10–14	1	2	1		4
14–18	2	1	1		4
18–22	1	3			4
$n_x$	4	7	6	3	20

Вычислить  $r_e$  и установить тесноту связи при  $\alpha = 0,05$ .

( $r_B = -0,812$ ;  $T_{\text{набл}} = 5,9$ ; значим)

5.3.3. По данным корреляционной таблицы вычислить  $r_B$ .

$Y \backslash X$	5	10	15	20	$n_y$
10	2				2
20	5	4	4		13
30	3	8	3	3	17
40		3	6	6	15
50			2	1	3
$n_x$	10	15	15	10	50

**Занятие 6.**  
**Линейная регрессия.**  
**Определение параметров линейной регрессии**

**6.1. Краткие теоретические сведения**

Если обе линии регрессии  $Y$  на  $X$  и  $X$  на  $Y$  являются прямыми, то в этом случае корреляцию называют линейной. Выборочное уравнение прямой линии регрессии  $Y$  на  $X$  имеет вид:

$$\bar{y}_x - \bar{y}_B = r_B \frac{\sigma_y}{\sigma_x} (x - \bar{x}_B). \quad (6.1)$$

Уравнение прямой регрессии  $X$  на  $Y$  имеет вид:

$$\bar{x}_y - \bar{x}_B = r_B \frac{\sigma_x}{\sigma_y} (y - \bar{y}_B). \quad (6.2)$$

Здесь  $x, y$  – значения СВ  $X, Y$ ;  $\bar{y}_x, \bar{x}_y$  – их выборочные средние.

Коэффициент уравнений (6.1)–(6.2) можно также определить по формулам, полученным методом наименьших квадратов. Например, если уравнение (6.1) взять в виде  $\bar{y}_x = ax + b$ , то параметры  $a$  и  $b$  линейной регрессии имеют вид:

$$a = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2}; \quad (6.3)$$
$$b = \frac{\sum_{i=1}^n y_i \sum_{i=1}^n x_i - \sum_{i=1}^n x_i \sum_{i=1}^n x_i y_i}{n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2}.$$

**Задача 6.1.1.** Распределение 40 заводов отрасли по количеству слесарей  $X$  и числу станко-смен  $Y$  задано корреляционной таблицей.

$Y \backslash X$	10–15	15–20	20–25	25–30	30–35	35–40	$n_y$
0–0,2	4						4
0,2–0,4	2	2					4
0,4–0,6			2				2
0,6–0,8		6		4	4		14
0,8–1,0					6	6	12
1,0–1,2						4	4
$n_x$	6	8	2	4	10	10	40

Составить уравнение прямой регрессии  $Y$  на  $X$ .

### Решение

По корреляционной таблице вычислим:

$$\bar{x}_B = \frac{1}{n} \sum_{i=1}^n x_i n_{xi} = \frac{1}{40} (12,5 \cdot 6 + 17,5 \cdot 8 + 22,5 \cdot 2 + 27,5 \cdot 4 + 32,5 \cdot 10 + 37,5 \cdot 10) = 26,75;$$

$$\bar{y}_B = \frac{1}{n} \sum_{i=1}^n y_i n_{yi} = \frac{1}{40} (0,4 \cdot 4 + 0,3 \cdot 4 + 0,5 \cdot 2 + 0,7 \cdot 14 + 0,9 \cdot 12 + 1,1 \cdot 4) = 0,69;$$

$$\sigma_x = 0,29; \quad \sigma_y = 9,25; \quad r_B = 0,85.$$

Подставим вычисленные значения в уравнение (6.1):

$$\bar{y}_x - 0,69 = 0,85 \cdot \frac{9,25}{0,29} (x - 26,75);$$

$$\bar{y}_x - 0,69 = 22,8(x - 26,75);$$

$$\bar{y}_x = 22,8x - 609,2.$$

**Задача 6.1.2.** При эталонировании медного термометра изучалась зависимость электрического сопротивления  $Y$  от температуры  $X$ . Были получены следующие результаты.

$i$	1	2	3	4	5	6
$x_i$	0	10	20	30	40	50
$y_i$	0,533	0,553	0,574	0,596	0,619	0,645

Оценить параметры уравнения регрессии с помощью метода наименьших квадратов и записать уравнение регрессии  $Y$  на  $X$ .

### Решение

Сведем результаты вычисления в таблицу.

Номер опыта $i$	Исходные данные		Расчетные данные		
	$x_i$	$y_i$	$x_i y_i$	$x_i^2$	$y_i^2$
1	0	0,533	0	0	0,2841
2	10	0,553	5,53	100	0,3047
3	20	0,574	11,48	400	0,3295
4	30	0,596	17,88	900	0,3552
5	40	0,619	24,75	1600	0,3832
6	50	0,645	32,25	2500	0,4160
$n = 6$	$\sum x_i$	$\sum y_i$	$\sum x_i y_i$	$\sum x_i^2$	$\sum y_i^2$
	150	3,519	91,89	5500	2,0727

Параметры линейной регрессии определим по формулам (6.3):

$$a = \frac{6 \cdot 91,89 - 150 \cdot 3,519}{6 \cdot 5500 - 150^2} = 0,002237;$$

$$b = \frac{3,519 \cdot 5500 - 150 \cdot 91,89}{6 \cdot 5500 - 150^2} = 0,53067.$$

Эмпирическое уравнение регрессии  $Y$  на  $X$  примет вид:

$$\bar{y}_x = 0,53067 + 0,002237x.$$

## 6.2. Задачи для аудиторной работы

6.2.1. Найти выборочное уравнение регрессии  $Y$  на  $X$  по данным, приведенным в корреляционной таблице.

$Y \backslash X$	5	10	15	20	25	30	35	40	$n_y$
100	2	1							3
120	3	4	3						10
140			5	10	8				23
160				1		6	1	1	9
180							4	1	5
$n_x$	5	5	8	11	8	6	5	2	50

$$(\bar{y}_x = 0,92x + 112,72)$$

6.2.2. Для исследования зависимости годового объема производства  $Y$  от основных фондов  $X$  получены статистические данные по 20 предприятиям.

$Y \backslash X$	12,5	17,5	22,5	27,5	$n_y$
20–21	1				1
21–22		2			2
22–23		1	2		3
23–24			3	3	6
24–25				8	8
$n_x$	1	3	5	11	20

Составить выборочное уравнение регрессии  $Y$  на  $X$ .

$$(\bar{y}_x = 17,524 + 0,2447x)$$

6.2.3. По данным измерениям двух переменных величин найти уравнение линейной регрессии  $Y$  на  $X$ .

$x_i$	66	70	75	80	82	85	90	92	95	98
$y_i$	60	78	65	87	74	70	78	95	88	90

$$(\bar{y}_x = 12,25 + 0,8x)$$

6.2.4. В таблице приводятся данные о распаде 10 г радиоактивного вещества, где  $t$  – время (мес.),  $X$  – количество (г) оставшегося вещества в момент  $t$ .

$t$	1	2	3	4	5	6	7	8	9	10	11	12
$X$	8,45	7,67	5,08	3,63	3,46	2,43	1,91	1,17	0,98	0,81	0,76	0,72

Составить уравнение линейной регрессии  $X$  на  $t$ .

### 6.3. Задачи для самостоятельной работы

6.3.1. В таблице приведены данные о связи между ценой на нефть  $X$  (ден. ед.) и индексом нефтяных компаний  $Y$  (усл. ед.).

$X$	11,0	11,5	12,0	12,5	13,0	13,5
$Y$	1,5	1,5	1,6	1,7	1,9	1,9

Составить уравнение прямой регрессии  $Y$  на  $X$ .

$$(\bar{y}_x = 0,189x - 0,677)$$

6.3.2. Найти выборочные уравнения прямых регрессии  $Y$  на  $X$  и  $X$  на  $Y$  по данным, приведенным в корреляционной таблице.

$Y \backslash X$	5	10	15	20	25	30	35	40	$n_y$
100	2	1							3
120	3	4	3						10
140			5	10	8				23
160				1		6	1	1	9
180							4	1	5
$n_x$	5	5	8	11	8	6	5	2	50

$$(\bar{y}_x = 1,92x + 100,9; \quad \bar{x}_y = 0,42y - 38,3)$$

## ТИПОВОЙ РАСЧЕТ ПО МАТЕМАТИЧЕСКОЙ СТАТИСТИКЕ

В задачах 1–20 дан интервальный статистический ряд распределения частот экспериментальных значений случайной величины  $X$ .

Требуется:

- 1) построить полигон и гистограмму частот (относительных частот) СВ  $X$ ;
- 2) по виду полигона и гистограммы и исходя из механизма образования СВ сделать предварительный выбор закона распределения;
- 3) вычислить выборочную среднюю  $\bar{x}_e$  и исправленное среднее квадратическое отклонение  $s$ ;
- 4) записать гипотетичную функцию распределения и плотность распределения;
- 5) найти доверительные интервалы для математического ожидания и среднего квадратического отклонения при доверительной вероятности  $\gamma = 0,95$ ;
- 6) найти теоретические частоты нормального закона распределения и проверить гипотезу о нормальном распределении СВ с помощью критерия Пирсона при уровне значимости  $\alpha = 0,05$ .

1. Даны результаты испытания стойкости 200 удлиненных сверл диаметра 4 мм, ч:

Стойкость сверла $x_i$	3–3,2	3,2–3,4	3,4–3,6	3,6–3,8	3,8–4
Частота $m_i$	16	50	70	44	20

2. Даны результаты исследования 100 напыленных образцов на прочность напыленного слоя, кг/мм<sup>2</sup>:

Прочность $x_i$	2,0–2,2	2,2–2,4	2,4–2,6	2,6–2,8	2,8–3,0
Частота $m_i$	7	22	38	23	10

3. Даны результаты исследования на разрыв 100 образцов дюра-люминия, кг/мм<sup>2</sup>:

Предел прочности $x_i$ , кг/мм <sup>2</sup>	42–43	43–44	44–45	45–46	46–47
Частота $m_i$	8	25	36	22	9

4. Даны результаты содержания фосфора (6 %) в 100 чугуновых образцах:

Содержание фосфора $x_i$	0,1–0,2	0,2–0,3	0,3–0,4	0,4–0,5	0,5–0,6
Частота $m_i$	7	22	38	24	9

5. Даны результаты испытания стойкости 100 сверл, ч:

Стойкость $x_i$ , ч	17,5–22,5	22,5–27,5	27,5–32,5	32,5–37,5	37,5–42,5
Частота $m_i$	7	20	44	20	9

6. Даны данные о среднесуточном пробеге 100 автомобилей автocolонны, сотни км:

$x_i$ , сотни км	1,2–1,6	1,6–2,0	2,0–2,4	2,4–2,8	2,8–3,2
Частота $m_i$	8	19	47	20	6

7. С автомата обрабатывающего втулки диаметра  $d = 40$  мм взята выборка изделий объемом 100. Результаты измерения диаметров втулок приведены в таблице:

Диаметр $x_i$ , мм	40,00–40,04	40,04–40,08	40,08–40,12	40,12–40,16	40,16–40,20
Частота $m_i$	8	19	44	20	9



8. В таблице приведены статистические данные о трудоемкости, мин, операции «Контроль механического состояния автомобиля после возвращения в гараж»:

Трудоемкость $x_i$ , мин	3–4	4–5	5–6	6–7	7–8	8–9
Частота $m_i$	6	8	33	35	11	7

9. В таблице приведены статистические данные о трудоемкости, мин, операции «ремонт валика водяного насоса автомобиля»:

Трудоемкость $x_i$ , мин	0–10	10–20	20–30	30–40	40–50
Частота $m_i$	17	47	70	46	20

10. Даны результаты испытания стойкости 100 фрез (в часах):

Стойкость $x_i$ , ч	21–26	26–31	31–36	36–41	41–46
Частота $m_i$	8	21	43	21	7

11. Даны сведения о расходе воды, используемой цехом для технических нужд в течение 100 дней, м<sup>3</sup>:

Расход $x_i$ , м <sup>3</sup>	8–12	12–16	16–20	20–24	24–28
Частота $m_i$	7	25	36	22	10

12. Даны квартальные данные о среднесуточном пробеге 100 автомобилей, км:

Среднесуточный пробег $x_i$	120–140	140–160	160–180	180–200	200–220
Частота $m_i$	9	21	40	18	12

13. Даны значения температуры масла в двигателе автомобиля БелАЗ при средних скоростях:

Температура $x_i$ , град.	40–45	45–50	50–55	55–60	60–65
Частота $m_i$	8	17	46	18	11

14. Даны размеры внутреннего диаметра гайки, мм:

Диаметр $x_i$ , мм	10,00–10,02	10,02–10,04	10,04–10,06	10,06–10,08	10,08–10,10
Частота $m_i$	9	16	47	21	7

15. Даны размеры диаметров 100 отверстий, просверленных одним и тем же сверлом:

Диаметр $x_i$ , мм	8,02–8,07	8,07–8,12	8,12–8,17	8,17–8,22	8,22–8,27
Частота $m_i$	10	19	38	21	12

16. Даны результаты измерения диаметра валика, обработанного одношпиндельным автоматом:

Диаметр $x_i$ , мм	19,80–19,85	19,85–19,90	19,90–19,95	19,95–20,00	20,05–20,10	20,10–20,15
Частота $m_i$	6	15	27	32	14	6

17. Даны результаты исследования грануляции партии порошка, мкм:

Грануляция $x_i$ , мкм	0–40	40–80	80–120	120–160	160–200
Частота $m_i$	7	23	35	26	9

18. Даны результаты наблюдений за сроком службы 150 однотипных станков до выхода за пределы норм (в месяцах двухсменной работы):

Срок $x_i$ , мес.	18–20	20–22	22–24	24–26	26–28
Частота $m_i$	15	27	61	29	18

19. Даны результаты измерения толщины, см, 100 слюдяных прокладок:

Толщина $x_i$ , см	0,20–0,26	0,26–0,32	0,32–0,38	0,38–0,44	0,44–0,50
Частота $m_i$	13	19	48	12	8

20. Даны диаметры 100 валиков после шлифовки, мм:

Диаметр $x_i$ , мм	20,0–20,1	20,1–20,2	20,2–20,3	20,3–20,4	20,4–20,5
Частота $m_i$	11	23	49	10	7

В задачах 21–40 приводятся результаты наблюдений над СВ  $X$  и  $Y$ . Используя эти экспериментальные данные, необходимо:

1) построить корреляционное поле. Подобрать математическую модель регрессионной зависимости  $Y$  от  $X$  (рекомендуется использовать модель линейной регрессии);

2) оценить параметры  $a$  и  $b$  модельного уравнения регрессии (6.1) методом наименьших квадратов;

3) записать эмпирическое уравнение регрессии  $Y$  на  $X$ .

21. СВ  $X$  и СВ  $Y$  – уровни жидкости в различных цилиндрах одной гидросистемы после контрольных испытаний.

$x_i$ , см	12,1	11,2	9,8	10,4	9,2	8,5	8,8	7,4
$y_i$ , см	10,5	9,3	8,3	9,6	8,6	7,1	6,9	5,8

22. СВ  $X$  – величина напряжения стального бруса; СВ  $Y$  – величина нагрузки при сжатии стального бруса.

$x_i$ , Кн	5	10	20	40	60
$y_i$ , МПа	51,33	78,00	144,3	263,6	375,2

23. СВ  $X$  – углубление резца; СВ  $Y$  – удельная энергия.

$x_i$	4	8	10	14	16	20	19	23
$y_i$	41	50	81	104	120	139	154	180

24. СВ  $X$ , СВ  $Y$  – уровни жидкости в различных цилиндрах одной гидросистемы после контрольных испытаний.

$x_i$ , см	7,4	6,6	7,0	6,4	6,0	6,5	5,8	5,4
$y_i$ , см	5,8	5,2	5,0	5,1	4,6	5,0	4,4	3,9

25. СВ  $X$  – электрическое сопротивление молибдена; СВ  $Y$  – температура.

$x_i$ , см	61,97	57,32	52,70	47,92	37,72	32,09	28,09
$y_i$ , К	2289	2132	1988	1830	1489	1286	1178

26. СВ  $X$  – электровооруженность труда на одного рабочего; СВ  $Y$  – выпуск продукции на одного рабочего.

$x_i$ , кВт·ч	3,0	3,5	4,0	4,5	5,0	5,5	6,0	6,5
$y_i$ , тыс. руб.	4,3	4,8	5,0	5,7	6,5	7,0	7,5	8,1

27. СВ  $X$  – температура; СВ  $Y$  – сопротивление медного термометра.

$x_i$ , °С	0	10	20	30	40	50	60	70
$y_i$ , Ом	0,533	0,552	0,574	0,596	0,619	0,645	0,687	0,690

28. СВ  $X$  – масса детали; СВ  $Y$  – время, затрачиваемое на закрепление детали на токарном станке.

$x_i$ , кг	0,7	0,8	1,0	1,2	1,3	1,4	1,5	1,7
$y_i$ , с	2,2	2,3	2,4	2,5	2,6	2,7	2,8	3,0

29. СВ  $X$  – плотность брикетов из карбонильного железного порошка; СВ  $Y$  – предел прочности на стане двух таких брикетов.

$x_i$ , %	75	76	77	80	82	85	88	90
$y_i$ , ГПа	2,1	2,0	2,5	2,4	3,6	4,0	4,1	5,0

30. СВ  $X$  – скорость движения автомобиля ЗИЛ-130; СВ  $Y$  – длина его тормозного пути.

$x_i$ , км/ч	10	15	20	25	30	35	40	45
$y_i$ , м	1,8	2,7	2,5	4,5	4,4	6,3	6,5	6,5

31. СВ  $X$  – скорость движения автомобиля ВАЗ-2301; СВ  $Y$  – длина его тормозного пути.

$x_i$ , км/ч	31	30	42	40	55	48	64	59
$y_i$ , м	2,0	2,6	3,9	5,2	7,0	6,2	7,5	8,6

32. СВ  $X$  – давление гелия; СВ  $Y$  – объем одного моля гелия.

$x_i \cdot 10^8$ , Па	3,0	3,6	4,0	4,5	5,2	5,6	6,0	6,4
$y_i \cdot 10^{-2}$ , м <sup>3</sup>	1,98	1,92	1,93	1,81	1,83	1,70	1,73	1,68

33. СВ  $X$  – масса груза, подвешенного на эластичном шнуре; СВ  $Y$  – удлинение этого шнура.

$x_i$ , кг	0,05	0,07	0,100	0,125	0,150	0,175	0,200	0,250
$y_i$ , см	0,005	0,052	0,012	0,016	0,017	0,025	0,027	0,034

34. СВ  $X$  – температура при прессовании болтов из стекловолокнита; СВ  $Y$  – предел их прочности.

$x_i$ , °C	1,30	1,35	1,40	1,45	1,50	1,55	1,60	1,65
$y_i \cdot 10^8$ , Па	10,8	10,2	9,2	8,9	8,3	8,3	8,0	7,3

35. СВ  $X$  – ударная вязкость инструментальных быстрорежущих сталей; СВ  $Y$  – коэффициент их обрабатываемости.

$x_i \cdot 10^{-3}$ , Дж/м <sup>2</sup>	0,7	0,8	0,9	1,0	1,1	1,2	1,3	1,4
$y_i$	0,6	0,62	0,64	0,67	0,69	0,73	0,75	0,8

36. СВ  $X$  – стаж работы; СВ  $Y$  – среднегодовое перевыполнение нормы.

$x_i$ , год	2	3	4	5	6	7
$y_i$ , %	6	6	7	8	9	10

37. СВ  $X$  – отклонение размеров валиков от номинала при черновой обработке; СВ  $Y$  – при чистовой обработке.

$x_i$ , мкм	–30	–25	–20	–15	–10	–5	0
$y_i$ , мкм	–8	–4	0	2	4	8	12

38. СВ  $X$  – скорость движения автомобиля БелАЗ; СВ  $Y$  – температура смазочного масла в двигателе этого автомобиля.

$x_i$ , км/ч	20	25	30	35	40	45	50	55
$y_i$ , °C	43,5	43,9	44,2	45,0	46,0	46,9	47,5	49,0

39. СВ  $X$  – скорость резания; СВ  $Y$  – площадь поперечного сечения стружки при обработке.

$x_i$ , м/мин	25,0	22,7	22,1	19,8	17,0	12,3	10,7	10,0	8,2
$y_i$ , мм <sup>2</sup>	1,1	1,4	1,7	2,1	2,6	4,7	6,1	7,0	10,0

40. СВ  $X$  – температура; СВ  $Y$  – коэффициент трения в подшипнике.

$x_i$ , °C	60	70	80	90	100	110	120
$y_i$	0,0148	0,0124	0,0102	0,0085	0,0071	0,0059	0,0051

## ЛИТЕРАТУРА

1. Боровиков, В. П. STATISTICA: Статистический анализ и обработка данных в среде Windows / В. П. Боровиков, И. П. Боровиков. – М. : Информационно-издательский дом «Филинь», 1998. – 608 с.
2. Гмурман, В. Е. Теория вероятностей и математическая статистика / В. Е. Гмурман. – М. : Высшая школа, 2003. – 479 с.
3. Гмурман, В. Е. Руководство к решению задач по теории вероятностей и математической статистике / В. Е. Гмурман. – М. : Высшая школа, 1997. – 400 с.
4. Микулик, Н. А. Теория вероятностей и математическая статистика / Н. А. Микулик, А. В. Метельский. – Минск : Пион, 2002. – 192 с.
5. Математика для инженеров / под науч. ред. Н. А. Микулика. – Минск : Элайда, 2006. – Т. 2. – 496 с.
6. Микулик, Н. А. Решение технических задач по теории вероятностей и математической статистике / Н. А. Микулик, Г. Н. Рейзина. – Минск : Вышэйшая школа, 1966. – 163 с.
7. Белько, И. В. Теория вероятностей и математическая статистика. Примеры и задачи / И. В. Белько, Г. Л. Свирид. – Минск : Новое знание, 2002. – 250 с.
8. Теория вероятностей и математическая статистика : учебник / сост.: М. А. Матальцкий, Г. А. Хацкевич. – Минск : Вышэйшая школа, 2017. – 591 с. : ил.
9. Сборник задач по теории вероятностей, случайных процессов и математической статистике : учебное пособие / Ю. С. Харин, Г. А. Хацкевич, В. И. Лобач. – Минск : БГУ, 1995. – 99 с.



Приложение 1

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

x	0	1	2	3	4	5	6	7	8	9
0,0	0,3989	3989	3989	3988	3986	3084	3982	3980	3977	3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3025	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3697
0,4	3683	3668	3652	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2804	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	0,2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1,7	0940	0925	0909	0893	0878	0863	0846	0833	0818	0804
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
2,0	0,0540	0529	0519	0508	0498	0488	0478	0468	0459	0449
2,1	0440	0431	0422	0413	0404	0396	0387	0379	0371	0363
2,2	0355	0347	0339	0332	0325	0317	0310	0303	0297	0290
2,3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229
2,4	0224	0219	0213	0208	0203	0198	0194	0189	0184	0180
2,5	0175	0171	0167	0163	0158	0154	0151	0147	0143	0139

Окончание табл.

$x$	0	1	2	3	4	5	6	7	8	9
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061
2,9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0046
3,0	0,0044	0043	0042	0040	0039	0038	0037	0036	0035	0034
3,1	0033	0032	0032	0030	0029	0028	0027	0026	0025	0025
3,2	0024	0023	0022	0022	0021	0020	0020	0019	0018	0018
3,3	0017	0017	0012	0016	0015	0015	0014	0014	0013	0013
3,4	0012	0012	0010	0011	0011	0010	0010	0010	0009	0009
3,5	0009	0008	0008	0008	0008	0007	0007	0007	0007	0006
3,6	0006	0006	0006	0005	0005	0005	0005	0005	0005	0004
3,7	0004	0004	0004	0004	0004	0004	0003	0003	0003	0003
3,8	0003	0003	0003	0003	0003	0002	0002	0002	0002	0002
3,9	0002	0002	0002	0002	0002	0002	0002	0002	0001	0001

Значения функции Лапласа  $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$

x	Φ(x)	x	Φ(x)	x	Φ(x)	x	Φ(x)
0,00	0,0000	0,32	0,1255	0,64	0,2389	0,96	0,3315
0,01	0,0040	0,33	0,1293	0,65	0,2422	0,97	0,3340
0,02	0,0080	0,34	0,1331	0,66	0,2454	0,98	0,3365
0,03	0,0120	0,35	0,1368	0,67	0,2486	0,99	0,3389
0,04	0,0160	0,36	0,1406	0,68	0,2517	1,00	0,3413
0,05	0,0199	0,37	0,1443	0,69	0,2549	2,01	0,3438
0,06	0,0239	0,38	0,1480	0,70	0,2580	1,02	0,3461
0,07	0,0279	0,39	0,1517	0,71	0,2611	1,03	0,3485
0,08	0,0319	0,40	0,1554	0,72	0,2642	1,04	0,3508
0,09	0,0359	0,41	0,1591	0,73	0,2673	1,05	0,3531
0,10	0,0398	0,42	0,1628	0,74	0,2703	1,06	0,3554
0,11	0,0438	0,43	0,1664	0,75	0,2734	1,07	0,3577
0,12	0,0478	0,44	0,1700	0,76	0,2764	1,08	0,3599
0,13	0,0517	0,45	0,1736	0,77	0,2794	1,09	0,3621
0,14	0,0557	0,46	0,1772	0,78	0,2823	1,10	0,3643
0,15	0,0596	0,47	0,1808	0,79	0,2852	1,11	0,3665
0,16	0,0636	0,48	0,1844	0,80	0,2881	1,12	0,3686
0,17	0,0675	0,49	0,1879	0,81	0,2910	1,13	0,3708
0,18	0,0714	0,50	0,1915	0,82	0,2939	1,14	0,3729
0,19	0,0753	0,51	0,1950	0,83	0,2967	1,15	0,3749
0,20	0,0793	0,52	0,1985	0,84	0,2995	1,16	0,3770
0,21	0,0832	0,53	0,2019	0,85	0,3023	1,17	0,3790
0,22	0,0871	0,54	0,2054	0,86	0,3051	1,18	0,3810
0,23	0,0910	0,55	0,2088	0,87	0,3078	1,19	0,3830
0,24	0,0948	0,56	0,2123	0,88	0,3106	1,20	0,3849
0,25	0,0987	0,57	0,2157	0,89	0,3133	1,21	0,3869
0,26	0,1026	0,58	0,2190	0,90	0,3159	1,22	0,3883
0,27	0,1064	0,59	0,2224	0,91	0,3186	1,23	0,3907
0,28	0,1103	0,60	0,2257	0,92	0,3212	1,24	0,3925
0,29	0,1141	0,61	0,2291	0,93	0,3238	1,25	0,3944
0,30	0,1179	0,62	0,2324	0,94	0,3264		
0,31	0,1217	0,63	0,2357	0,95	0,3289		

Окончание табл.

$x$	$\Phi(x)$	$x$	$\Phi(x)$	$x$	$\Phi(x)$	$x$	$\Phi(x)$
1,26	0,3962	1,59	0,4441	1,92	0,4726	2,50	0,4938
1,27	0,3980	1,60	0,4452	1,93	0,4732	2,52	0,4941
1,28	0,3997	1,61	0,4463	1,94	0,4738	1,54	0,4945
1,29	0,4015	1,62	0,4474	1,95	0,4744	2,56	0,4948
1,30	0,4032	1,63	0,4484	1,96	0,4750	2,58	0,4951
1,31	0,4049	1,64	0,4495	1,97	0,4756	2,60	0,4953
1,32	0,4066	1,65	0,4505	1,98	0,4761	2,62	0,4956
1,33	0,4082	1,66	0,4515	1,99	0,4767	2,64	0,4959
1,34	0,4099	1,67	0,4525	2,00	0,4772	2,66	0,4961
1,35	0,4115	1,68	0,4535	2,02	0,4783	2,68	0,4963
1,36	0,4131	1,69	0,4545	2,04	0,4793	2,70	0,4965
1,37	0,4147	1,70	0,4554	2,06	0,4803	2,72	0,4967
1,38	0,4162	1,71	0,4564	2,08	0,4812	2,74	0,4969
1,39	0,4177	1,72	0,4573	2,10	0,4821	2,76	0,4971
1,40	0,4192	1,73	0,4582	2,12	0,4830	2,78	0,4973
1,41	0,4207	1,74	0,4591	2,14	0,4838	2,80	0,4974
1,42	0,4222	1,75	0,4599	2,16	0,4846	2,82	0,4976
1,43	0,4236	1,76	0,4608	2,18	0,4854	2,84	0,4977
1,44	0,4251	1,77	0,4616	2,20	0,4861	2,86	0,4979
1,45	0,4265	1,78	0,4625	2,22	0,4868	2,88	0,4980
1,46	0,4279	1,79	0,4633	2,24	0,4875	2,90	0,4981
1,47	0,4292	1,80	0,4641	2,26	0,4881	2,92	0,4982
1,48	0,4306	1,81	0,4649	2,28	0,4887	2,94	0,4984
1,49	0,4319	1,82	0,4656	2,30	0,4893	2,96	0,4985
1,50	0,4332	1,83	0,4664	2,32	0,4898	2,98	0,4986
1,51	0,4345	1,84	0,4671	2,34	0,4904	3,00	0,49865
1,52	0,4357	1,85	0,4678	2,36	0,4909	3,20	0,49931
1,53	0,4370	1,86	0,4686	2,38	0,4913	3,40	0,49966
1,54	0,4382	1,87	0,4693	2,40	0,4918	3,60	0,499841
1,55	0,4394	1,88	0,4699	2,42	0,4922	3,80	0,499928
1,56	0,4406	1,89	0,4706	2,44	0,4927	4,00	0,499968
1,57	0,4418	1,90	0,4713	2,46	0,4931	4,50	0,499997
1,58	0,4429	1,91	0,4719	2,48	0,4934	5,00	0,499997

Значения функции  $\chi^2_{\alpha;v}$ ;  $P(\chi^2 \geq \chi^2_{\alpha;v}) = \alpha$ 

$v \backslash \alpha$	0,20	0,10	0,05	0,02	0,01	0,001
1	1,642	2,706	3,841	5,412	6,635	10,827
2	3,219	4,605	5,991	7,824	9,210	13,815
3	4,642	6,251	7,815	9,837	11,345	16,266
4	5,989	7,779	9,488	11,668	13,277	18,467
5	7,289	9,236	11,070	13,388	15,086	20,515
6	8,558	10,645	12,592	15,033	16,812	22,457
7	9,803	12,017	14,067	16,622	18,475	24,322
8	11,030	13,362	15,507	18,168	20,090	26,125
9	12,242	14,684	16,919	19,679	21,666	27,877
10	13,442	15,987	18,307	21,161	23,209	29,588
11	14,631	17,275	19,675	22,618	24,725	31,264
12	15,812	18,549	21,026	24,054	26,217	32,909
13	16,985	19,812	22,362	25,472	27,688	34,528
14	18,151	21,064	23,685	26,683	29,141	36,123
15	19,311	22,307	24,996	28,259	30,578	37,697
16	20,465	23,542	26,296	29,633	32,000	39,252
17	21,615	24,769	27,587	30,995	33,409	40,790
18	22,760	25,989	28,869	32,346	34,805	42,312
19	23,900	27,204	30,144	33,687	36,191	43,820
20	25,038	28,412	31,410	35,020	37,566	45,315
21	26,171	29,615	32,671	36,343	38,932	46,797
22	27,301	30,813	33,924	37,659	40,289	48,268
23	28,429	32,007	35,172	38,968	41,638	49,728
24	29,553	33,196	36,415	40,270	42,980	51,179
25	30,675	34,382	37,652	41,566	44,312	52,620
26	31,795	35,563	38,885	42,856	45,642	54,052
27	32,912	36,741	40,113	44,140	46,963	55,476
28	34,027	37,916	41,337	45,419	48,278	56,893
29	35,139	39,087	42,557	46,693	49,588	58,302
30	36,250	40,256	43,773	47,962	50,892	59,703

Распределение Стьюдента.

Значения  $t_{\alpha;v}$  удовлетворяют условию  $P(t \geq t_{\alpha;v}) = \int_{t_{\alpha;v}}^{\infty} S(t, v) dt = \alpha$

$v \setminus \alpha$	0,40	0,30	0,20	0,10	0,05	0,025	0,010	0,005	0,001	0,0005
1	0,325	0,727	1,376	3,078	6,314	12,71	31,82	63,66	318,3	636,6
2	0,289	0,617	1,061	1,886	2,920	4,303	6,965	9,925	22,33	32,60
3	0,277	0,584	0,978	1,638	2,353	3,182	4,541	5,841	10,22	12,94
4	0,271	0,569	0,941	1,533	2,132	2,776	3,747	4,604	7,173	8,610
5	0,267	0,559	0,920	1,476	2,015	2,571	3,365	5,032	5,893	6,859
6	0,265	0,553	0,906	1,440	1,943	2,447	3,143	3,707	5,208	5,959
7	0,263	0,549	0,896	1,415	1,895	2,365	2,998	3,499	4,785	5,405
8	0,262	0,546	0,889	1,397	1,860	2,306	2,896	3,355	4,501	5,041
9	0,261	0,543	0,883	1,383	1,833	2,262	2,821	3,250	4,297	4,781
10	0,260	0,542	0,879	1,372	1,812	2,228	2,764	3,169	4,144	4,587
11	0,260	0,540	0,876	1,363	1,796	2,201	2,718	3,106	4,025	4,437
12	0,259	0,539	0,873	1,356	1,782	2,179	2,681	3,055	3,930	4,318
13	0,259	0,538	0,870	1,350	1,771	2,160	2,650	3,012	3,852	4,221
14	0,258	0,537	0,868	1,345	1,761	2,145	2,624	2,977	3,787	4,140
15	0,258	0,536	0,866	1,341	1,753	2,131	2,602	2,947	3,733	4,073
16	0,258	0,535	0,865	1,337	1,746	2,120	2,583	2,921	3,686	4,015
17	0,257	0,534	0,863	1,333	1,740	2,110	2,567	2,898	3,646	3,965
18	0,257	0,534	0,862	1,330	1,734	2,101	2,552	2,878	3,611	3,922
19	0,257	0,533	0,861	1,328	1,729	2,093	2,539	2,861	3,579	3,883
20	0,257	0,533	0,860	1,325	1,725	2,086	2,528	2,845	3,552	3,850
21	0,257	0,532	0,859	1,323	1,721	2,080	2,518	2,831	3,527	3,819
22	0,256	0,532	0,858	1,321	1,717	2,074	2,508	2,819	3,505	3,792
23	0,256	0,532	0,858	1,319	1,714	2,069	2,500	2,807	3,485	3,767
24	0,256	0,531	0,857	1,318	1,711	2,064	2,492	2,797	3,467	3,745
25	0,256	0,531	0,856	1,316	1,708	2,060	2,485	2,787	3,450	3,725

Окончание табл.

$\nu \backslash \alpha$	0,40	0,30	0,20	0,10	0,05	0,025	0,010	0,005	0,001	0,0005
26	0,256	0,531	0,856	1,315	1,716	2,056	2,479	2,779	3,435	3,707
27	0,256	0,531	0,855	1,314	1,703	2,052	2,473	2,771	3,421	3,690
28	0,256	0,530	0,855	1,313	1,701	2,048	2,467	2,763	3,408	3,674
29	0,256	0,530	0,854	1,311	1,699	2,045	2,462	2,756	3,396	3,659
30	0,256	0,530	0,854	1,310	1,697	2,042	2,457	2,750	3,385	3,646
40	0,255	0,529	0,851	1,303	1,684	2,021	2,423	2,704	3,307	3,551
50	0,255	0,528	0,849	1,298	1,676	2,009	2,403	2,678	3,262	3,495
60	0,254	0,527	0,848	1,296	1,671	2,000	2,390	2,660	3,232	3,460
80	0,254	0,527	0,846	1,292	1,664	1,990	2,374	2,639	3,195	3,415
100	0,254	0,526	0,845	1,290	1,660	1,984	2,365	2,626	3,174	3,389
200	0,254	0,525	0,843	1,286	1,653	1,972	2,345	2,601	3,131	3,339
500	0,253	0,525	0,842	1,283	1,648	1,965	2,334	2,586	3,106	3,310
$\infty$	0,253	0,524	0,842	1,282	1,645	1,960	2,326	2,576	3,090	3,291

Значения функции  $t_{\gamma;n}: \bar{x}_B - t_{\gamma;n} \frac{\bar{s}}{\sqrt{n}} < a < \bar{x}_B + t_{\gamma;n} \frac{s}{\sqrt{n}}$

$n \setminus \gamma$	0,95	0,99	0,999	$n \setminus \gamma$	0,95	0,99	0,999
5	2,78	4,60	8,61	20	2,093	2,861	3,883
6	2,57	4,03	6,86	25	2,064	2,797	3,745
7	2,45	3,71	5,96	30	2,045	2,756	3,659
8	2,37	3,50	5,41	35	2,032	2,720	3,600
9	2,31	3,36	5,04	40	2,023	2,708	3,558
10	2,26	3,25	4,78	45	2,016	2,692	3,527
11	2,23	3,17	4,59	50	2,009	2,679	3,502
12	2,20	3,11	4,44	60	2,001	2,662	3,464
13	2,18	3,06	4,32	70	1,996	2,649	3,439
14	2,16	3,01	4,22	80	1,991	2,640	3,418
15	2,15	2,98	4,14	90	1,987	2,633	3,403
16	2,13	2,95	4,07	100	1,984	2,627	3,392
17	2,12	2,92	4,02	120	1,980	2,617	3,374
18	2,11	2,90	3,97	$\infty$	1,960	2,576	3,291
19	2,10	2,88	3,92				



Приложение 6

Значения коэффициентов  $q_1$  и  $q_2$ ;  $q_1S < \sigma < q_2S$

	0,99		0,98		0,95		0,00	
	$q_1$	$q_2$	$q_1$	$q_2$	$q_1$	$q_2$	$q_1$	$q_2$
1	0,356	15,0	0,388	79,8	0,446	31,9	0,510	15,9
2	0,434	14,1	0,466	9,97	0,521	6,28	0,578	4,40
3	0,483	6,47	0,514	5,11	0,566	3,73	0,620	2,92
4	0,519	4,39	0,549	3,67	0,599	2,87	0,649	2,37
5	0,546	3,48	0,576	3,00	0,624	2,45	0,672	2,090
6	0,569	2,98	0,597	2,62	0,644	2,202	0,690	1,916
7	0,588	2,66	0,616	2,377	0,661	2,035	0,705	1,797
8	0,604	2,440	0,631	2,205	0,675	1,916	0,718	1,711
9	0,618	2,277	0,644	2,076	0,688	1,826	0,729	1,645
10	0,630	2,154	0,656	1,977	0,699	1,755	0,739	1,593
11	0,641	2,056	0,667	1,898	0,708	1,698	0,748	1,550
12	0,651	1,976	0,676	1,833	0,717	1,651	0,755	1,515
13	0,660	1,910	0,685	1,779	0,725	1,611	0,762	1,485
14	0,669	1,854	0,693	1,733	0,732	1,577	0,769	1,460
15	0,676	1,806	0,700	1,694	0,739	1,548	0,775	1,437
16	0,683	1,764	0,707	1,659	0,745	1,522	0,780	1,418
17	0,690	1,727	0,713	1,629	0,750	1,499	0,785	1,400
18	0,696	1,695	0,719	1,602	0,756	1,479	0,790	1,385
19	0,702	1,668	0,725	1,578	0,760	1,460	0,794	1,370
20	0,707	1,640	0,730	1,556	0,765	1,444	0,798	1,358
21	0,712	1,617	0,734	1,536	0,769	1,429	0,802	1,346
23	0,722	1,576	0,743	1,502	0,777	1,402	0,809	1,326
24	0,726	1,558	0,747	1,487	0,781	1,391	0,812	1,316
25	0,730	1,541	0,751	1,473	0,784	1,380	0,815	1,308
26	0,734	1,526	0,755	1,460	0,788	1,371	0,818	1,300
27	0,737	1,512	0,758	1,448	0,791	1,361	0,820	1,293
29	0,744	1,487	0,765	1,426	0,796	1,344	0,825	1,279
30	0,748	1,475	0,768	1,417	0,799	1,337	0,828	1,274
40	0,774	1,390	0,792	1,344	0,821	1,279	0,847	1,228
50	0,793	1,336	0,810	1,297	0,837	1,243	0,861	1,199
60	0,808	1,299	0,824	1,265	0,849	1,217	0,871	1,179
70	0,820	1,272	0,835	1,241	0,858	1,198	0,879	1,163
80	0,829	1,250	0,844	1,222	0,866	1,183	0,886	1,151
90	0,838	1,233	0,852	1,207	0,873	1,171	0,892	1,141
100	0,845	1,219	0,858	1,195	0,878	1,161	0,897	1,133
200	0,887	1,15	0,897	1,13	0,912	1,11	0,925	1,09

Критические значения распределения Колмогорова

$$P(\lambda > \lambda_\alpha) = \alpha$$

$\alpha$	0,2	0,1	0,05	0,02	0,01	0,001
$\lambda_\alpha$	1,073	1,224	1,358	1,520	1,627	1,950

## СОДЕРЖАНИЕ

Введение.....	3
Теоретические вопросы для подготовки к экзамену.....	4
Аналитическая система TIBCO Statistica™.....	5
Общая информация.....	5
Запуск STATISTICA.....	7
Главное окно STATISTICA.....	7
Spreadsheet – электронная таблица с исходными данными.....	9
Построение статистического графика.....	11
Статистический анализ.....	14
Вывод результатов анализа.....	19
Занятие 1. Статистическое распределение. Эмпирическая функция распределения и ее свойства. Полигон и гистограмма.	
Числовые характеристики выборки.....	21
1.1. Краткие теоретические сведения.....	21
1.2. Задачи для аудиторной работы.....	26
1.3. Задачи для самостоятельной работы.....	29
Занятие 2. Точечные оценки неизвестных параметров распределения.....	31
2.1. Краткие теоретические сведения.....	31
2.2. Задачи для аудиторной работы.....	33
2.3. Задачи для самостоятельной работы.....	35
Занятие 3. Интервальные оценки.....	36
3.1. Краткие теоретические сведения.....	36
3.2. Задачи для аудиторной работы.....	39
3.3. Задачи для самостоятельной работы.....	40
Занятие 4. Статистическая проверка гипотез.	
Критерии согласия Пирсона и Колмогорова.....	42
4.1. Краткие теоретические сведения.....	42
4.2. Задачи для аудиторной работы.....	48
4.3. Задачи для самостоятельной работы.....	49
Занятие 5. Выборочный коэффициент корреляции и его свойства. Проверка гипотезы о равенстве нулю коэффициента корреляции.....	51
5.1. Краткие теоретические сведения.....	51
5.2. Задачи для аудиторной работы.....	54
5.3. Задачи для самостоятельной работы.....	56

Занятие 6. Линейная регрессия. Определение параметров линейной регрессии.....	58
6.1. Краткие теоретические сведения.....	58
6.2. Задачи для аудиторной работы.....	61
6.3. Задачи для самостоятельной работы.....	62
<b>ТИПОВОЙ РАСЧЕТ ПО МАТЕМАТИЧЕСКОЙ СТАТИСТИКЕ.....</b>	<b>63</b>
<b>ЛИТЕРАТУРА .....</b>	<b>72</b>
Приложение 1.....	73
Приложение 2.....	75
Приложение 3.....	77
Приложение 4.....	78
Приложение 5.....	80
Приложение 6.....	81
Приложение 7.....	82

Учебное издание

**ЧЕПЕЛЕВА** Тереса Иосифовна  
**ЧЕПЕЛЕВ** Николай Иосифович  
**ЧЕПЕЛЕВ** Сергей Николаевич и др.

## **МАТЕМАТИЧЕСКАЯ СТАТИСТИКА**

Учебно-методическое пособие  
для обучающихся по специальностям  
1-37 01 02 «Автомобилестроение (по направлениям)»,  
1-37 01 01 «Двигатели внутреннего сгорания»

Редактор *Н. А. Костешева*  
Компьютерная верстка *Н. А. Школьниковой*

Подписано в печать 05.12.2022. Формат 60×84 <sup>1</sup>/<sub>16</sub>. Бумага офсетная. Ризография.

Усл. печ. л. 4,94. Уч.-изд. л. 3,05. Тираж 100. Заказ 605.

Издатель и полиграфическое исполнение: Белорусский национальный технический университет.  
Свидетельство о государственной регистрации издателя, изготовителя, распространителя  
печатных изданий № 1/173 от 12.02.2014. Пр. Независимости, 65. 220013, г. Минск.