

**МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
ДОНЕЦКОЙ НАРОДНОЙ РЕСПУБЛИКИ
ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
«ДОНЕЦКИЙ НАЦИОНАЛЬНЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

**МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ
к проведению лабораторных работ по дисциплине
«Биометрия»**

Донецк
2021

**МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
ДОНЕЦКОЙ НАРОДНОЙ РЕСПУБЛИКИ
ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
«ДОНЕЦКИЙ НАЦИОНАЛЬНЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

КАФЕДРА «ПРИКЛАДНАЯ ЭКОЛОГИЯ И ОХРАНА ОКРУЖАЮЩЕЙ СРЕДЫ»

**МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ
к проведению лабораторных работ по дисциплине
«Биометрия»**

для обучающихся по направлению подготовки
05.03.06 «Экология и природопользование»,
профиль «Экологическая безопасность»
всех форм обучения

РАССМОТРЕНО

на заседании кафедры
прикладной экологии и охраны
окружающей среды
Протокол № 6 от 21.01.2021 г.

УТВЕРЖДЕНО

на заседании учебно-издательского
совета ДОННТУ
Протокол № 2 от 24.02.2021 г.

Донецк
2021

УДК 57.087.1(076)
М54

Составители:

Ганнова Юлия Николаевна – кандидат химических наук, доцент кафедры прикладной экологии и охраны окружающей среды ГОУВПО «ДОННТУ»;
Горбатко Сергей Витальевич – кандидат технических наук, доцент кафедры прикладной экологии и охраны окружающей среды ГОУВПО «ДОННТУ».

М54 **Методические рекомендации к проведению лабораторных работ по дисциплине «Биометрия»** : для обучающихся по направлению подготовки 05.03.06 «Экология и природопользование» профиль «Экологическая безопасность» всех форм обучения / ГОУВПО «ДОННТУ», каф. «Прикладная экология и охрана окружающей среды»; сост.: Ю.Н. Ганнова, С.В. Горбатко. — Донецк : ДОННТУ, 2021. – Систем. требования: Acrobat Reader. – Загл. с титул. экрана.

Методические рекомендации разработаны с целью оказания помощи обучающимся в усвоении теоретического материала и получении практических навыков по дисциплине «Биометрия История дизайна», которые содержат указания для выполнения лабораторных работ по курсу.

УДК 57.087.1(076)

Содержание

Введение	4
Лабораторная работа №1	9
Лабораторная работа №2	26
Лабораторная работа №3	40

ВВЕДЕНИЕ

Задачи лабораторного практикума

1. Способствовать развитию общего научного мировоззрения.
2. Дать студентам современное научное представление о статистическом анализе групповых свойств.
3. Сообщить определенный комплекс знаний, необходимый для успешного изучения последующих дисциплин и правильного применения в расчетах средних величин, дисперсионного, корреляционного и регрессионного анализа.

Выполняя лабораторные работы, студенты знакомятся с методами проведения измерений и расчетов, влиянием различных условий на применение той или иной методики. Все это дает возможность лучше понять сущность научных исследований, глубже усвоить важнейшие понятия, постулаты и законы, основательнее закрепить в памяти теоретический материал.

Методика выполнения лабораторных работ

1. Прочитать полностью методику выполнения лабораторной работы. При первом чтении не следует задерживаться на математических выводах формул, нужно лишь составить общее представление об излагаемых вопросах и отметить особенно трудные или неясные места.

2. Далее перейти к выполнению задания лабораторной работы. Целесообразным является использование примеров выполнения, которые приведены к каждой лабораторной работе. Чтобы лучше усвоить соответствующий материал, в электронных документах необходимо давать пояснения значений новых независимых терминов и названий, формул и обозначений переменных.

3. Закончив выполнение лабораторной работы, ответьте на контрольные вопросы, которые находятся в конце каждой работы. При ответах на вопросы попытайтесь не пользоваться учебником и конспектом лекций.

К защите лабораторной работы распечатайте ее на бумаге формата А4, с указанием номер варианта, названием дисциплины, группы, фамилии и инициалов студента. При защите, ответы на контрольные вопросы, которые даны в конце каждого лабораторного задания, должны быть коротко мотивированны.

Техника безопасности и организация рабочего места при выполнении лабораторных работ

Организация рабочего места при выполнении лабораторных работ является очень важным вопросом, который нельзя недооценивать. От

правильной организации рабочего места зависит очень многое – удобство работы, и, как следствие, ее результаты, безопасность для собственного здоровья и здоровья окружающих, безопасность и надежность самой аппаратуры.

Вопросами правильной организации рабочего места занимается относительно новое направление в науке – эргономика, которая возникла на стыке многих наук — анатомии, физиологии, психологии, техники безопасности и других.

Основная часть аппаратуры должна располагаться на гладкой горизонтальной поверхности, расположенной на высоте около 70 см от пола. В качестве такой поверхности используется достаточно прочный и устойчивый письменный стол. Устройства, входящие в состав ПК, следует расположить на рабочей поверхности таким образом, чтобы обеспечить удобство работы.

Расположите монитор на столе так, чтобы центр его экрана находился напротив и несколько ниже уровня ваших глаз (обычно на 5-7 см). Расстояние плоскости экрана от глаз должно составлять не менее 55 см (оптимально 60 или 70 см). Компьютер может оказывать негативное влияние на работающего. Степень этого влияния и характер его проявления бывают различными, и в значительной мере определяются возрастом, психическим и общим состоянием здоровья работающего. Для здоровья работающего могут оказаться вредными несколько факторов: высокое напряжение, которое существует в системе, вредные электромагнитные излучения, неправильная организация рабочего места, труда и отдыха, которые должны быть несколько специфичны.

Высокое напряжение. В лаборатории настольные ПК получают электропитание от промышленной сети переменного тока частотой 50 Гц и напряжением 220 вольт. Это напряжение является опасным для жизни и здоровья человека. Неосторожное обращение с аппаратурой, использующей его, может привести к тяжелым травмам.

Напряжение 220 вольт используется в блоках питания. Как правило, последние делаются неразъемными, а их ремонт в условиях эксплуатации не предусмотрен. Конструкция блока питания исключает случайное прикосновение к токоведущим участкам схемы, находящимся под напряжением 220 вольт. Во всех случаях запрещается вскрывать и пытаться ремонтировать блоки питания.

Вредные излучения. При работе компьютера создаются электромагнитные поля. Параметры создаваемого различными приборами электромагнитного поля могут заметно отличаться. Характер и степень влияния электромагнитных полей на здоровье человека изучены еще не достаточно полно. Однако электромагнитные поля оказывают негативное влияние на здоровье человека. Степень этого влияния зависит от характеристик поля. К основным характеристикам поля относятся частота и интенсивность. Некачественное изображение на экране монитора. К сожалению, вредные излучения являются не единственной, а учитывая

прилагаемые изготовителями меры для их ослабления, и не главной причиной опасного влияния монитора на здоровье. С помощью монитора работающий воспринимает огромный объем информации, которая поступает в мозг для обработки.

При этом органы зрения и мозг работают с большим напряжением, что равносильно сильному стрессу, приводящему к общему ослаблению организма. При этом в наибольшей степени страдают глаза. В этой связи существует специальный термин «компьютерный зрительный синдром», которым обозначается комплекс изменений в зрительной системе, вызванный работой за монитором. Симптомы этого заболевания выражаются в покраснении глаз, в боли и рези в них, в головных болях и головокружении.

Сильнее этому воздействию подвержены дети и подростки. Очень большое значение имеет качество изображения на экране монитора. Чем оно выше, тем меньше устают глаза. Быстрое утомление глаз вызывает мерцающий экран, недостаточная четкость изображения, неоптимальное сочетание цветов. Глаза устают от бликов света на экране, от маленьких символов и их недостаточной резкости, от преобладающих цветов (наиболее неприятны красный и насыщенный синий), от повышенной или пониженной яркости экрана, от засветки экрана внешним освещением.

Уменьшить опасность возникновения неприятных последствий на здоровье от работы за монитором может ряд простых мер. Во-первых, нужно правильно расположить монитор на рабочем месте (см. выше). Центр экрана должен находиться примерно на 6-8 см ниже уровня глаз работающего. Экран монитора должен по возможности освещаться как можно меньше внешними источниками света. Лучше всего, когда экран перпендикулярен источнику света.

Выполнение многочасовых лабораторных работ, в особенности, если это сопряжено с отрицательными эмоциями, например из-за сбоев ПК при их выполнении, может привести к стрессу и, как следствие, к головной боли, повышению давления. Поэтому, по поводу всех замеченных нарушениях работы ПК следует информировать преподавателей ведущих лабораторные работы. Причиной быстрого физического утомления может быть шум. Поэтому следует обязательно обращать внимание преподавателей в случае аномально высокого уровня производимого ПК шума.

Иногда работающие за ПК испытывают боль в пояснице, в шее и плечах. Обычно это происходит из-за неправильного положения корпуса работающего или из-за очень долгого пребывания в неизменной позе.

Правильная организация режима работы и отдыха. Многие часы, проведенные за компьютером, приводят к переутомлению, боли в глазах, в спине и кистях рук. Менее заметными могут стать гипертония и другие заболевания. Важно правильно чередовать режим труда и отдыха. Уменьшить опасность возникновения неприятных последствий на здоровье от работы за ПК может ряд простых мер и, в первую очередь, соблюдение режима труда и отдыха.

Следует ограничивать время непрерывной работы за монитором, а главное, во время работы делать достаточно частый отдых. Простейшим отдыхом являются следующие приемы:

- перевод взгляда с экрана монитора на какой-либо посторонний предмет;

- просто закрывать глаза на 1-2 минуты;

- посмотреть на близко расположенный предмет, а затем перевести взор на удаленный предмет. Затем закрыть глаза и не открывать их около 30 сек. Затем поморгать глазами несколько раз. Такое упражнение следует повторить 5-6 раз. Отдых нужен и по иным причинам. Статичная поза приводит к напряжению мышц спины и шеи. Поэтому периодически (хотя бы каждые 40-45 минут) следует вставать из-за рабочего места и немного походить (хотя бы 10-15 минут, компьютер выключать при этом не нужно!). Работа за монитором в течение 2-4 часов в день при соблюдении рассмотренных выше рекомендаций для большинства взрослых людей считается абсолютно безопасной. Считается, что 8 часов является предельно допустимой нормой для взрослого здорового человека при условии систематических перерывов в работе. В любом случае, при наступлении усталости, рези в глазах, головной боли следует прервать работу и отдохнуть.

Контрольные вопросы

1. Перечислите основные требования, которые предъявляются к расположению работающего и устройств, входящих в состав ПК, при выполнении лабораторных работ.

2. Какие негативные воздействия может оказывать ПК на здоровье работающего с ним человека.

3. Приведите примеры упражнений для снятия усталости глаз, болей в спине и кистях рук при длительной работе с ПК.

ЛАБОРАТОРНАЯ РАБОТА №1

Тема: «Расчет показателей разнообразия признака, средних и средних взвешенных величин с помощью стандартных функций редактора электронных таблиц».

Цель работы – приобрести навыки применения стандартных функций редактора электронных таблиц для выполнения основных биометрических расчетов показателей разнообразия признаков в выборках, средних величин (арифметической, геометрической, гармонической, квадратической).

Запуск Редактора электронных таблиц (Microsoft Excel)

После включения компьютера и загрузки ГОС, на экране монитора появляется рабочий стол. Запустить Microsoft Excel можно следующими способами:



После щелчка на кнопке Пуск необходимо подвести указатель мыши к пункту Программы и в открывшемся подменю выбрать программу Microsoft Excel. Для того чтобы запустить эту программу на выполнение, необходимо щелкнуть левой кнопкой мыши.



После двойного щелчка на книге-документе или ссылке документа Microsoft Excel запускается последний и открывается соответствующий документ (Рис.1.1).

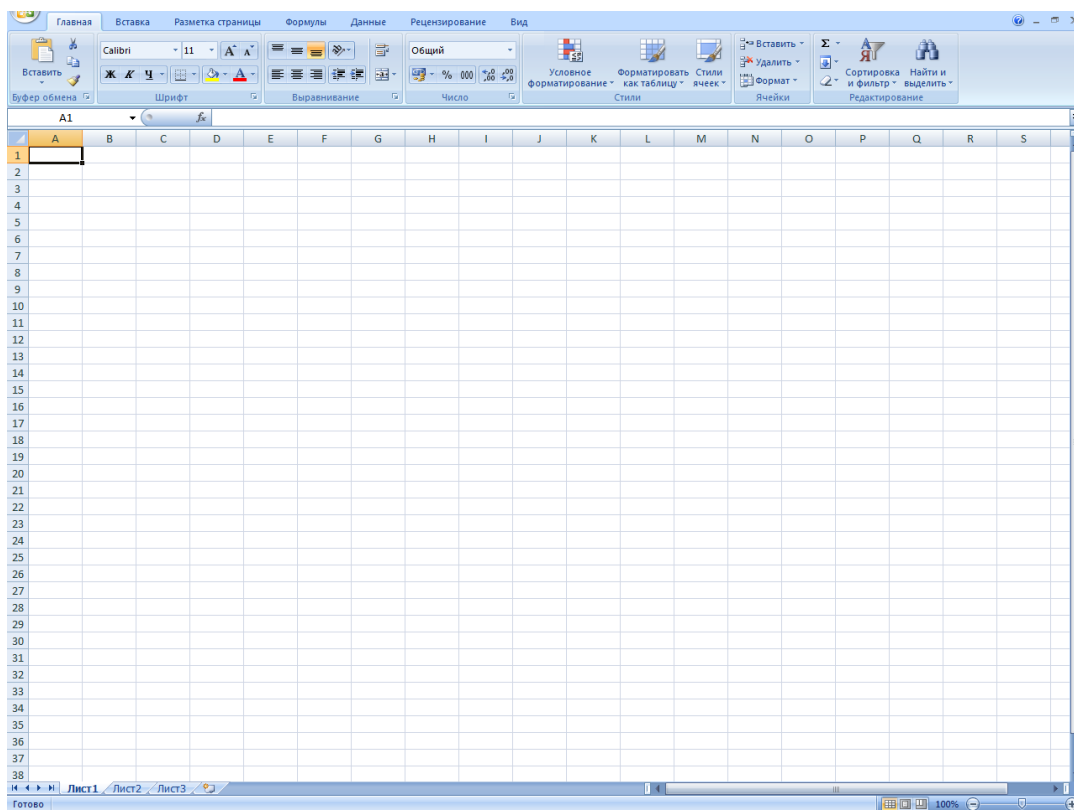


Рисунок 1.1. Фрагмент типового окна редактора электронных таблиц Microsoft Excel

Элементы управления окна Microsoft Excel

После запуска Редактора таблиц Microsoft Excel на экране появляется один из листов электронной таблицы рабочей книги.

Первая строка содержит название документа и является строкой заголовка.

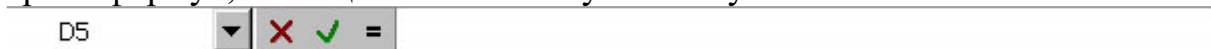
2. Вторая полоса содержит список меню, каждое из которых выдает меню нижнего уровня, которые используют для выбора команд Microsoft Excel.

3. Третья полоса содержит множество инструментов, которые можно запустить, щелкая по ним мышью.

4. Обычно в четвертой строке содержатся инструменты форматирования, с помощью которых можно устанавливать стили, размеры, шрифты, атрибуты шрифта (полужирный, курсив, подчеркнутый), выбирать способы выравнивания (по левому, по правому краю, центрирование по обоим краям, объединение ячеек) для информации содержащейся в текущей ячейке(-ках).

Следующие элементы присущи только окну Microsoft Excel.

5. Пятая полоса является строкой формул, в которой выводится содержимое выделенной в настоящий момент (активной) ячейки и ее координаты - адрес (в левой части). В этой строке можно вводить и редактировать числа, строки символов и формулы. Все, что содержится в строке формул, помещается в активную ячейку листа.



6. Область электронного листа рабочей книги, которая ограничена сверху - названиями столбцов, а слева - названиями строк (рис.2.1).

7. Строка с ярлыками листов книги и строка статуса Microsoft Excel.



Выбирая команду Панель инструментов из меню Вид, можно сделать панели стандартных инструментов и форматирования видимыми или невидимыми.

Структура рабочей книги и виды адресации в Microsoft Excel Окно рабочей книги. Каждый документ Microsoft Excel представляет собой рабочую книгу. Одновременно можно открыть несколько рабочих книг – каждую в своем окне. Каждая книга состоит из листов, а листы – из ячеек, которые имеют уникальный адрес, формируемый из названия столбца и строки, на пересечении которых они располагаются.

Таким образом, каждая ячейка в книгах имеет свой уникальный адрес, который может быть использован как ссылка в формулах. Полный адрес ячейки в пределах книги следующий:

[Книга1.xls]Лист1!Название_столбца Номер_строки

Заголовки столбцов.

A	B	C	D
---	---	---	---

 Столбцы обозначаются буквами латинского алфавита от A до Z и далее от AA до IV (всего 255 столбцов). Они располагаются вдоль верхней их части.

1

Заголовки строк. Строки различаются номерами от 1 до 65536, расположенными в левой части, образуя столбец заголовков строк.

2

3 Ссылки на ячейку. Ссылкой однозначно определяется ячейка или
4 группа ячеек листа, а также упрощается поиск значений или данных,
5 используемых в формуле. С помощью ссылок можно использовать в формуле данные, находящиеся в различных местах листа, книги, других книг или данные другого приложения, а также значение одной и той же ячейки в нескольких формулах.

Чтобы сослаться на диапазон ячеек, введите ссылку на верхний левый угол диапазона, поставьте двоеточие (:), а затем — ссылку на правый нижний угол диапазона, например: A5:C22.

Различают абсолютные и относительные ссылки. Разница между ними проявляется во время копирования ячеек содержащих эти ссылки. При указании мышкой на ячейку(ки), по умолчанию, создаются относительные ссылки, например, A1. При копировании ячейки содержащей такие ссылки они автоматически изменяются на необходимое число колонок и/или строк. Если при копировании необходимо чтобы не было изменений в ссылках - их задают как абсолютные:

\$A\$1 – полностью абсолютная ссылка;

A\$1 – A-относительная, 1-абсолютная;

\$A1 – A-абсолютная, 1-относительная.

Для обозначения абсолютных ссылок на ячейки используют знак “\$”.

Название ячеек и их диапазонов. К ячейке или диапазонам можно обращаться, кроме адреса, по имени, которое задают в поле Имя слева в строке формул (<Ctrl+F3>). Название имени может состоять не более чем из 255 символов – букв, цифр, специальных символов без пробелов. Ссылки на имена всегда абсолютные.

Ввод данных и работа с ячейками

Существует несколько типов данных для работы в Microsoft Excel:

1. текстовые; 2. числовые; 3. даты; 4. времени; 5. формулы.

Чтобы ввести любой тип данных необходимо:

1. выбрать требуемую ячейку (левый щелчок или клавишами стрелками);

2. набрать с клавиатуры символы. При этом они появятся в ячейке и строке формул;

3. нажать <Enter>, <Tab> или клавиши-стрелки. <Esc> - отменяет набор.

Текстовые данные представляют собой последовательности цифр и нецифровых символов. Обычно текст выравнивается (прижимаются) в ячейке по левому краю.

Если часть текста не помещается по длине ячейки, она будет закрыта соседними ячейками.

Для принудительного обозначения последовательности символов как текст - ввод начинают с символа <'> (одинарная кавычка).

Чтобы расположить в активной ячейке текст в несколько строк выполните Формат=>Ячейки=>Выравнивание=>Переносить по словам.

Чтобы принудительно начать в ячейке новую строку, нажмите <Alt+Enter> (в English раскладке клавиатуры).

Для пояснения содержимого ячейки используют возможность добавления примечания: Вставка=>Примечание.

Числовые данные вводят аналогично тексту только в виде последовательности цифр от 0 до 9 с использованием специальных символов
% + - () , / \$ % . E e.

При вводе дробей пользуются или десятичной или натуральной записью (с нулем через пробел при отсутствии целой части).

Числовые данные автоматически выравниваются по правому краю.

Независимо от количества отображаемых разрядов числа хранятся с точностью до 15 разрядов. Если число имеет больше 15 значащих знаков, то разряды после 15-го преобразуются в нули (0).

Дата и время. При вводе значений даты или времени происходит их автоматическое распознавание, и общий формат ячейки заменяется на встроенный формат даты или времени.

Для времени и даты в качестве разделителей используются символы двоеточие (:) и косая (обратный слеш) (/) или точка. Чтобы ввести дату и время суток в одну и ту же ячейку, в качестве разделителя даты и времени используйте пробел. Чтобы ввести время с использованием 12-часового формата, после значения времени вводят отделенные пробелом буквы AM или PM (A или P). Иначе, время будет интерпретировано с использованием 24-часового формата.

Время суток и даты можно складывать, вычитать и производить с ними другие вычисления. В формулы они могут быть введены в виде текста, заключенного в двойные кавычки. Например, формула: ="05.12.94"- "03.05.94" возвращает 216 дней. В Microsoft Excel используется система отсчета дат, в которой числу 1 соответствует 1 января 1900 г.

Формулы. Формула должна начинаться со знака равенства (=), за которым следует набор вычисляемых величин. Результат выполнения отобразится в ячейке, в которой создана формула. С помощью формул можно складывать, умножать и сравнивать данные, а также объединять значения. Значения необходимых ячеек вводятся в формулу с помощью ссылки на ячейку. Подробно о формулах будет сказано далее.

С каждой ячейкой связано две ее характеристики: содержание и значение.

Содержание - это то, что вводится в ячейку с клавиатуры.

Значение – то, что видно на экране.

Для текста и чисел содержание и значение ячейки одинаковы. Для формул, содержание представляет собой введенную формулу, а значение – итог вычислений, произведенных с применением данной формулы.

Редактирование данных ячейки

После ввода данных их можно исправить в самой ячейке или в строке формул. Чтобы перейти в режим редактирования активной ячейки:

1. или <F2>;
2. или двойной щелчок;
3. или щелчок в строке формул.

Для перемещения по редактируемой в ячейке строке символов используют клавиши-стрелки.

Для удаления символов - <Delete> и <Backspace>.




Отменить внесенные изменения и вернуться к первоначальному варианту – <Esc> или щелкнуть на кнопке с "крестиком" в левой части строки формул.

Закончить редактирование с сохранением внесенных изменений - <Enter>, <Tab> или щелкнуть на кнопке с "птичкой" в левой части строки формул.



После редактирования или других действий их можно отменить - <Ctrl+z> или кнопкой на панели инструментов.

Виды курсоров, указателей (маркеров) и их назначение

Для набора текста в Microsoft Excel предназначена клавиатура. Ячейка, в которой появляется текст, называется текущей (активной) при этом она выделена табличным курсором (контурной рамкой с точкой в нижнем правом углу). 

В режиме редактирования (<F2>) содержимого ячейки появляется курсор, который имеет вид горизонтальной мигающей черточки.

Указатели мыши:



Контурный крестик(1) – служит для:

- а) выбора активной ячейки;
- б) отметки (выделения) диапазона ячеек, строк, столбцов;

Контурный крестик (1) появляется в области рабочего листа.



Маленький крестик(2) – служит для:


- а) копирования текста из активной ячейки в соседние;
- б) быстрого формирования набора однотипных формул в соседних ячейках по образцу формулы, находящейся в активной ячейке;


в) автозаполнения ячеек последовательностями чисел с необходимой прогрессией. Здесь необходимо задать два члена последовательности, использовать правую кнопку мыши и появляющееся контекстное меню или совместно с <Ctrl> (здесь к указателю добавляется еще один крестик).

Стрелка(3) - служит для перемещения данных ячейки(ек) в другое место. Появляется при позиционировании контурного крестика к границе активной ячейки или выделенного блока ячеек.

При перемещении автоматически корректируются адреса перемещаемых ячеек во всех формулах, где они используются.

Этот указатель (3) с нажатой <Ctrl> (при этом появляется дополнительный маленький крестик) - служит для копирования данных. Двойной щелчок с этим указателем выделяет диапазон ячеек таблицы.

 Черточка(4) (вертикальная/горизонтальная) со стрелками – служит для изменения ширины столбца или высоты строки. Этот указатель появляется при позиционировании указателя (1) к границе названий столбцов или строк. Нажав левую кнопку и перемещая мышью границы, выполняют изменения. Двойной щелчок – выполняет автоматический подбор под самую длинную строку или высокий столбец.

 Двойная черточка(5) (вертикальная/горизонтальная) со стрелками - служит для отображения скрытых столбцов или строк. Этот указатель появляется при позиционировании указателя (1) к границе названий выделенных столбцов или строк, между которыми находятся скрытые элементы. Нажав левую кнопку (или правую с контекстным меню) и перемещая мышью, выполняют отображение скрытых строк или столбцов (или <Ctrl+Shift> + <9>/<0>, или Формат=>Строка (Столбец)=>Отобразить).

Создание и применение формул для расчетов

Формула является основным средством для анализа данных. С помощью формул можно складывать, умножать и сравнивать данные, а также объединять значения и осуществлять поиск. В следующем примере складывается значение ячейки B4 с числом 25. Полученный результат делится на сумму ячеек D5, E5 и F5.

$$=(B4+25)/СУММ(D5:F5)$$

Синтаксис формулы. Формула должна начинаться со знака равенства (=), за которым следует набор вычисляемых величин. Результат выполнения отобразится в ячейке, в которой указана формула. Формула может включать следующие операнды: операторы, ссылки на ячейки, непосредственные значения (константы), имена функций.

Включить или Выключить режим отображения формул в ячейках можно с помощью <Ctrl+`>.

При копировании формул ссылки на ячейки, которые в нее входят изменяются соответственно их новому расположению. Чтобы при копировании формул, содержащих ссылки на ячейки не изменялись, применяют абсолютные ссылки.

Применение операторов в формулах. Операторами обозначаются операции, которые следует выполнить над операндами формулы. В Microsoft Excel включено четыре вида операторов: арифметические, текстовые, операторы сравнения и адресные операторы.

Арифметические операторы (табл. 1.1) используются для выполнения основных математических вычислений над числами. Результатом

выполнения арифметической операции всегда является число. Старшинство операций следующее: -, %, ^, (*, /), (+, -), &, операции сравнения.

Текстовый оператор «&» используется для обозначения операции объединения последовательностей символов в единую последовательность.

Например: A1="Северный" A2=" ветер".

Результатом объединения =A1&A2 будет: "Северный ветер"

Операторы сравнения используются для обозначения операций сравнения двух чисел. Результатом выполнения операции сравнения является логическое значение ИСТИНА или ЛОЖЬ.

Таблица 1.1 Арифметические операторы

Арифметический оператор	Значение	Пример
+	Сложение	=4+4
-	Вычитание	=5-1
*	Умножение	=2*2
/	Деление	=7/7
%	Процент	20 %
^	Возведение в степень	8^2

Адресные операторы (табл. 1.2) объединяют диапазоны ячеек для осуществления вычислений.

Таблица 1.2 Адресные операторы

Адресный оператор	Значение	Пример
: (двоеточие)	Оператор диапазона, который ссылается на все ячейки между границами диапазона включительно	B5:B15
, (запятая)	Оператор объединения, который ссылается на объединение ячеек диапазонов.	СУММ(B5:B15, D5:D15)
(пробел)	Оператор пересечения, который ссылается на общие ячейки диапазонов. Ячейка B7 является общей для двух диапазонов.	СУММ(B5:B15 A7:D7)

Таблица 1.3 Оператор сравнения

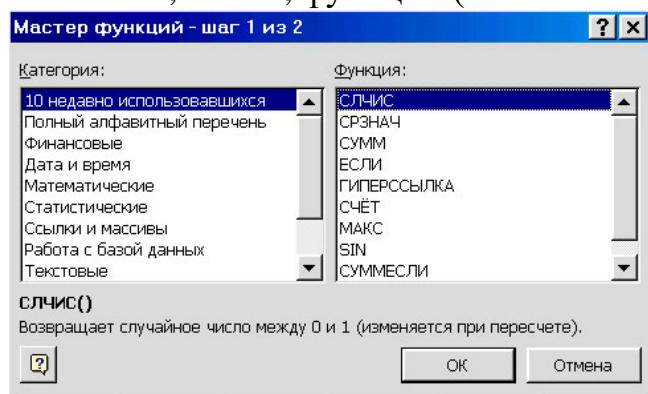
Оператор сравнения	Значение	Пример
=	Равно	A1=B1
>	Больше	A1>B1
<	Меньше	A1<B1

>=	Больше или равно	A1>=B1
<=	Меньше или равно	A1<=B1
<>	Не равно	A1<>B1

Функции. В Microsoft Excel содержится более 1000 стандартных формул, называемых функциями. Функции используются для простых или сложных вычислений. Обращение к функции имеет вид:


ИМЯ_ФУНКЦИИ(арг1; арг2; ...; аргN)

В качестве аргументов (арг1, арг2, аргN) могут выступать: ссылки, константы, имена, функции (вложенность до 7 включительно).



Если аргумент функции выходит за пределы аргументов, принимаемых этой функцией, то она возвращает ошибку #ЧИСЛО!. Если тип аргумента не соответствует данной функции, ею возвращается ошибка #ЗНАЧ!. Ввести функцию можно с помощью:

Рисунок 1.2. Первое диалоговое окно Мастера функций

1. клавиатуры, если известно ИМЯ ФУНКЦИИ.
2. команды Вставка=>Функция; 
3. кнопки с изображением функции на панели инструментов;

В появившемся диалоговом окне Мастер функций (рис.1.2), среди представленного перечня функций необходимо выбрать Категорию, к которой может относиться функция, и среди представленного перечня функций в правой части окна найти нужную, а затем щелкнуть на кнопке Далее.

Во втором диалоговом окне (рис.1.3) с клавиатуры или с помощью мышки указать константы или ячейки, которые будут являться ее аргументами, а затем щелкнуть кнопку ОК.

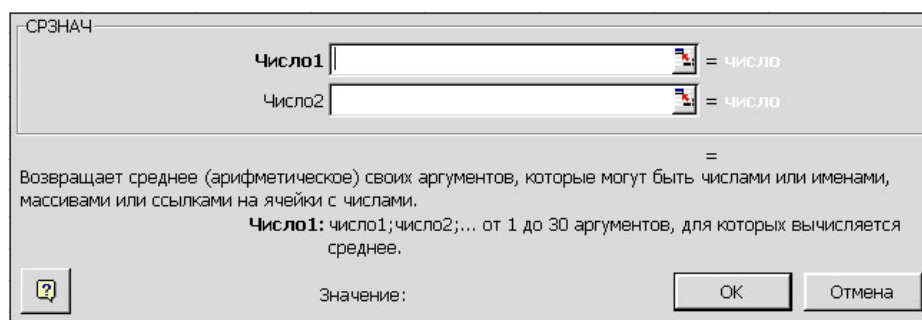



Рисунок 1.3. Второе диалоговое окно Мастера функций

Для доступа к пространству рабочего листа, с целью выбора мышкой нужных ячеек, это окно (см. рис.1.3) можно минимизировать, щелкнув по кнопке  в строке ввода параметров.

В MS Excel существуют два основных типа функций: обычные, которые возвращают одно значение и функции массива – возвращающие набор (массив) значений.

Чтобы вставить функцию массива в диапазон ячеек необходимо:

1. выделить эти ячейки;
2. ввести функцию массива и ее аргументы в верхнюю левую ячейку этого диапазона;
3. нажать комбинацию клавиш <Ctrl+Shift+Enter> или щелкнуть на кнопке ввода (кнопка с изображением зеленой галочки в строке формул).

В результате выполненных действий каждая ячейка выделенного диапазона будет содержать функцию массива, заключенную в фигурные скобки {}, а на листе отобразятся значения возвращаемые функцией.

Для выполнения биометрических расчетов среди предлагаемых в Microsoft Excel 9 групп функций наибольший интерес будут представлять функции относящиеся к математическим и статистическим (табл. 1.4).

Основные расчетные формулы, используемые в лабораторной работе №1

Средняя величина — основной показатель который широко используется в науке и практике.

Средняя величина признака определяется различными способами в зависимости от объекта наблюдения и поставленных целей. Поэтому имеется не один, а несколько видов средних, что приводит к определению типа средней: арифметической, геометрической, гармонической и т.д., и общепринятому разграничению между аналитическими и позиционными средними.

Самым распространенным показателем среднего качества является средняя арифметическая. Вычисляется она по следующим формулам:

Средняя арифметическая СРЗНАЧ(): Взвешенная средняя арифметическая:

$$M = \frac{\sum_{i=1}^n V_i}{n}$$

$$M_{\text{взв}} = \frac{\sum_{i=1}^n (V_i \cdot P_i)}{\sum_{i=1}^n P_i}$$

где V — дата, отдельное значение изучаемого признака, n — число использованных значений признака, P — математический вес усредняемого значения.

Средняя геометрическая применяется во всех случаях, когда необходимо узнать или спланировать средние приросты за определенный период. Вычисляется она по следующим формулам:

Средняя геометрическая СРГЕОМ(): Взвешенная средняя геометрическая:

$$G = \sqrt[n]{\prod_{i=1}^n V_i}$$

$$G_{\text{взв}} = \sqrt[n]{\prod_{i=1}^n V_i^{P_i}}$$

или через логарифмическую форму:

$$G = \sqrt[n]{\exp\left(\sum_{i=1}^n \ln V_i\right)}$$

$$G_{\text{взв}} = \sqrt[n]{\exp\left(\sum_{i=1}^n P_i \cdot \ln V_i\right)}$$

где Π — произведение дат V_i .

В отличие от других, при расчете средней квадратической получаются завышенные значения. Вычисляется она по следующим формулам:

Средняя квадратическая:

Взвешенная средняя квадратическая:

$$S = \sqrt{\frac{\sum_{i=1}^n V_i^2}{n}}$$

$$S_{\text{взв}} = \sqrt{\frac{\sum_{i=1}^n (P_i \cdot V_i^2)}{\sum_{i=1}^n P_i}}$$

Применяется средняя квадратическая при расчете средних радиусов окружностей.

Средняя гармоническая применяется при усреднении меняющихся скоростей с целью последующего расчета среднего времени движения.

При применении среднегармонической не должно быть нулевых дат. Если имеются и отрицательные и положительные значения, необходимо чтобы знаменатель не равнялся нулю. Вычисляется она по следующим формулам:

Средняя гармоническая СРГАРМ(): Взвешенная средняя гармоническая:

$$H = \frac{n}{\sum_{i=1}^n V_i^{-1}}$$

$$H_{\text{взв}} = \frac{\sum_{i=1}^n P_i}{\sum_{i=1}^n \frac{P_i}{V_i}}$$

Всякая группа состоит из особей, отличающихся друг от друга по каждому из признаков. Различия эти иногда очень велики, иногда они почти незаметны, но они всегда имеются, так как невозможно найти даже двух особей абсолютно одинаковых.

При изучении общих свойств совокупностей невозможно ограничиться одними средними величинами, требуется еще привлечь и такие показатели, которые характеризовали бы степень разнообразия особей в группе. Такими показателями являются:

1. лимиты \lim – самое максимальное МАКС() и минимальное МИН() значение признака;

2. среднее квадратическое отклонение σ :

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (V_i - M_i)^2}{n - 1}}$$

$v=n-1$ – число степеней свободы.

Этот показатель используется в качестве абсолютной меры разнообразия и, кроме того, положен в основу почти всех характеристик изменчивости, распределения, корреляции, регрессии, дисперсионного анализа.

3. коэффициент вариации CV:

$$CV = \frac{\sigma}{M} \cdot 100\%$$

Этот показатель является функцией обоих основных показателей — среднего квадратического отклонения и средней арифметической, выражается отвлеченным числом и поэтому очень удобен для сравнения разнообразия любых признаков.

4. нормированное отклонение НОРМАЛИЗАЦИЯ():

$$x = \frac{V - M}{\sigma}$$

Нормированное отклонение показывает, на сколько сигм отклоняется значение признака от средней для соответствующей группы.

Нормированное отклонение — величина неименованная, что представляет большое удобство при сравнении развития различных признаков.

Таблица 1.4 Некоторые встроенные стандартные функции Microsoft Excel

Функция	Действие	Пример
+, -, *, /, ^, (,)	Сложить, вычесть, умножить, разделить, возвести в	1+2; 5-8; A1/A2; A1^B2; A1/(B1+B2) и т.д.

	степень, скобки	
СРЗНАЧ(a1;a2;...)	Возвращает среднее арифметическое	ячейки A1:E1 содержат числа 10, 7, 9, 27 и 2, то: СРЗНАЧ(A1:E1) → 11
СРГЕОМ(a1;a2;...)	Возвращает среднее геометрическое	Ячейки A1:G1 содержат числа 4; 5; 8; 7; 11; 4; 3 СРГЕОМ(A1:G1) → 5,476987
СРОТКЛ(a1;a2;...)	Среднее отклонений от среднего	Ячейки A1:G1 содержат числа 4; 5; 6; 7; 5; 4; 3 СРОТКЛ(A1:G1) → 1,020408
СРГАРМ(a1;a2;...)	Возвращает среднее гармоническое	Ячейки A1:G1 содержат числа 4; 5; 8; 7; 11; 4; 3 СРГАРМ(A1:G1) → 5,028376
КВАДРОТКЛ(a1;a2;...)	Возвращает сумму квадратов отклонений от их среднего	Ячейки A1:G1 содержат числа 4; 5; 8; 7; 11; 4; 3 КВАДРОТКЛ(A1:G1) → 48
СУММ(a1;a2;...)	Возвращает сумму	Ячейки A1:G1 содержат числа 4; 5; 8; 7; 11; 4; 3 СУММ(A1:G1) → 42
ПРОИЗВЕД(a1;a2;...)	Возвращает произведение	Ячейки A1:G1 содержат числа 4; 5; 8; 7; 11; 4; 3 ПРОИЗВЕД(A1:G1) → 147840
СУММПРОИЗВ(a1:aN; b1:bN)	Возвращает сумму произведений	Ячейки A1:G1 содержат числа 4; 5; 8; 7; 11; 4; 3. Ячейки A2:G2 содержат числа 1; 2; 3; 4; 5; 6; 7. СУММПРОИЗВ(A1:G1;A2:G2) → 166
СЧЁТ(a1;a2;...)	Возвращает количество дат (чисел)	Ячейки A1:G1 содержат числа 4; 5; 8; 7; 11; 4; 3 СЧЁТ(A1:G1) → 7
МИН(a1;a2;...)	Возвращает минимальное значение	МИН(A1:J10) → 380
МАКС(a1;a2;...)	Возвращает максимальное значение	МАКС(A1:J10) → 450
НОРМАЛИЗАЦИЯ(W;M;σ)	Возвращает нормированное отклонение x	НОРМАЛИЗАЦИЯ(A12;\$A\$14;\$B\$14) → 2,4315
КОРЕНЬ(a1)	Возвращает	КОРЕНЬ(9) → 3

	корень квадратный числа a1	Аналогично $9^{0,5}$
LN(a1)	Возвращает натуральный логарифм числа a1	Ячейка A1 содержит 100 LN(A1) → 4,60517
EXP(a1)	Возвращает экспоненту числа a1	Ячейка B1 содержит 4,60517 EXP(B1) → 100

Контрольные вопросы

1. Какие основные элементы управления представлены в окне Microsoft Excel. Их назначение.
2. Из каких элементов структурно состоит книга приложения Microsoft Excel.
3. Перечислите типы данных, которые могут быть расположены в ячейках листа Microsoft Excel.
4. В чем отличие абсолютных и относительных ссылок в формулах листа.
5. С помощью каких приемов можно перейти в режим редактирования содержимого ячейки. Укажите клавиши удаления символов в ячейке.
6. Продемонстрируйте основные виды курсоров и указателей. Каково их назначение.
7. Приведите основные этапы создания простых формул и формул массива в ячейках рабочего листа.
8. Какие типы средних используются в биометрических расчетах.
9. Перечислите средние, которые относятся к аналитическим и к позиционным. По какому признаку проводят их разделение.
10. В чем отличие простых и взвешенных средних величин. Дайте определение математическому весу усредняемого признака.
11. Приведите примеры, когда правомерным является применение арифметической, геометрической, квадратической и гармонической средней.
12. Назовите основные свойства, которыми обладают средние величины.
13. В каком порядке должны располагаться все простые средние величины (для одной и той же выборки) согласно их численному абсолютному значению.
14. Приведите примеры основных показателей, которые характеризуют степень разнообразия признаков особей в группе (выборке).
15. Дайте определение понятию «число степеней свободы».

Задание к лабораторной работе №1

1. Используя стандартные функции приложения Microsoft Excel (табл. 2.1), для заданных значений признаков V и весов (частот анализа) P , вычислите среднюю (простую и взвешенную): арифметическую (M), геометрическую (G), гармоническую (H), квадратическую (S). Сравните между собой полученные результаты. Сохраните полученные результаты.

2. Используя стандартные функции приложения Microsoft Excel для заданных значений определите показатели разнообразия признака: среднеквадратическое отклонение (σ), лимиты \lim (максимальное и минимальное), коэффициент вариации CV . Рассчитать нормированное отклонение (x) признака. Сохраните полученные результаты.

3. С помощью известных формул для расчета средних и показателей разнообразия признака, используя арифметические операторы, покажите (подтвердите расчетом), что в п.1 и п.2 получены верные результаты. Окончательно сохраните и распечатайте все полученные результаты.

Порядок выполнения

1. Набрать исходные выданные значения содержания вещества в пробах и частоты (веса) его определения в два рядом стоящих столбца A и B , которые озаглавить, как показано в примере.

2. В следующем столбце D рассчитать натуральные логарифмы ($=LN(V)$) содержания вещества в пробах из первого столбца A , которые будут необходимы для последующего расчета средней геометрической простой и взвешенной.

3. В следующем столбце E рассчитать обратные величины ($=1/V$) для значений из столбца A , которые будут необходимы для последующего расчета средней гармонической простой и взвешенной.

4. Используя стандартную функцию СЧЁТ() в ячейке $K3$ рассчитать число элементов выборки, например, используя столбец A .

5. С помощью функций СРЗНАЧ(), СУММПРОИЗВ() и СУММ() в ячейках $I7$, $I10$ и $K7$ рассчитать арифметические средние простые и взвешенную.

6. В следующем столбце F рассчитать разности ($=V-M$) между значениями столбца A и взвешенной арифметической средней из ячейки $K7$, которые будут необходимы для последующего расчета среднего квадратического отклонения и нормированное отклонение (x) в столбце G .

7. С помощью функций СРГЕОМ(), EXP(), СУММПРОИЗВ(), ПРОИЗВ() и СУММ() в ячейках $I14$, $I17$ и $K14$ рассчитать геометрические средние простые и взвешенную.

8. С помощью функций КОРЕНЬ(), СУММПРОИЗВ() и СУММ() в ячейках $I21$ и $K21$ рассчитать квадратическую среднюю простую и взвешенную.

9. С помощью функций СРГАРМ(), СУММПРОИЗВ() и СУММ() в ячейках I25, I28 и K25 рассчитать гармонические средние простые и взвешенную.

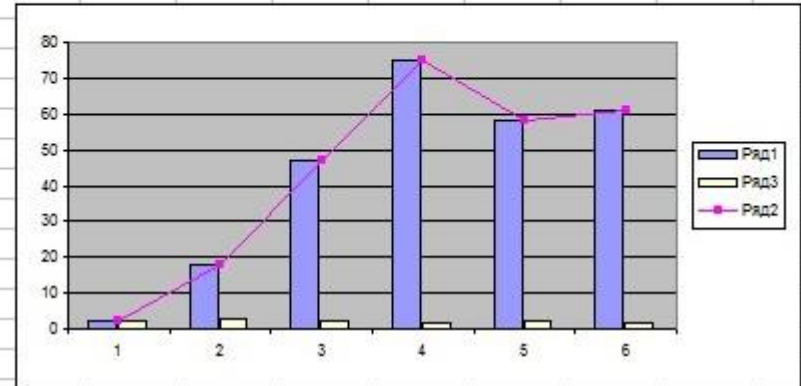
10. Используя функции МАКС() и МИН() найти лимиты для столбца А.

11. С помощью функций КВАДРОТКЛ(), СУММПРОИЗВ() и КОРЕНЬ() в ячейках I35 и K35 рассчитать среднее квадратические отклонения простое и взвешенное.

12. В ячейках I39 и K39 рассчитать коэффициенты вариации через простые и взвешенные средние.

Пример к лабораторной работе № 1

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	Лабораторная работа №1 по курсу "Применение компьютерных технологий в научных исследованиях ТНСМ"																			
2	Выполнил ст. гр. ТТМм-00								Фомовская Е.											
3	Задание										количество элементов= 12									
4	Copper Fluoride																			
5	6-0170	CuOHF*CuF2	Ln(V)	1/V	V-M	x	1. Средняя арифметическая													
6	12,1	2	2,49321	0,08264	5,91012	1,26326	простая													взвешенная
7	18,6	18	2,92316	0,05376	12,4101	2,6526	7,355													6,189884
8	8,35	47	2,12226	0,11976	2,16012	0,46171	проверка													
9	6,17	75	1,8197	0,16207	-0,0199	-0,0043	простая													
10	8,85	58	2,18042	0,11299	2,66012	0,56859	7,355													
11	6,77	61	1,9125	0,14771	0,58012	0,124														
12	0,82	49	-0,1985	1,21951	-5,3699	-1,1478	2. Средняя геометрическая													
13	4,81	73	1,5707	0,2079	-1,3799	-0,2949	простая													взвешенная
14	6,15	54	1,81645	0,1626	-0,0399	-0,0085	5,978													2,749681
15	3,19	55	1,16002	0,31348	-2,9999	-0,6412	проверка													
16	6,17	38	1,8197	0,16207	-0,0199	-0,0043	простая													
17	6,28	76	1,83737	0,15924	0,09012	0,01926	5,978													
18																				
19	3. Средняя квадратичная										5. Лимиты									
20	простая										взвешенная					Максимальное		Минимальное		
21	8,53244										6,88887					18,6		0,82		
22																				
23	4. Средняя гармоническая										6. Среднее квадратичное отклонение									
24	простая										взвешенная					простая		взвешенная		
25	4,13258										3,942828					4,51743		4,67847		
26	проверка																			
27	простая										7. Коэффициент вариации									
28	4,13258										через простую					через взвешенную				
29											61,4198					75,5825				
30																				



	B	C	D	E	F	G	H	I	J	K
1				Лабораторная работа №1 г						
2				Выполнил ст. гр. ТТММ-00				Фомовская Е.		
3								количество элементов=		=СЧЁТ(A6:A17)
4										
5	CuOHF*CuF2	Ln(V)	1/V	V-M	x			1. Средняя арифметическая		
6	2	=LN(A6)	=1/A6	=A6-\$K\$7	=F6/\$K\$36			простая		взвешенная
7	18	=LN(A7)	=1/A7	=A7-\$K\$7	=F7/\$K\$36			=СРЗНАЧ(A6:A17)		=СУММПРОИЗВ(A6:A17;B6:B17)/СУММ(B6:B17)
8	47	=LN(A8)	=1/A8	=A8-\$K\$7	=F8/\$K\$36				проверк	
9	75	=LN(A9)	=1/A9	=A9-\$K\$7	=F9/\$K\$36			простая		
10	58	=LN(A10)	=1/A10	=A10-\$K\$7	=F10/\$K\$36			=СРЗНАЧ(A6:A17)		
11	61	=LN(A11)	=1/A11	=A11-\$K\$7	=F11/\$K\$36					
12	49	=LN(A12)	=1/A12	=A12-\$K\$7	=F12/\$K\$36			2. Средняя геометрическая		
13	73	=LN(A13)	=1/A13	=A13-\$K\$7	=F13/\$K\$36			простая		взвешенная
14	54	=LN(A14)	=1/A14	=A14-\$K\$7	=F14/\$K\$36			=СРГЕОМ(A6:A17)		=EXP(СУММПРОИЗВ(B6:B17;D6:D17)^(1/СУММ(B6:B17)))
15	55	=LN(A15)	=1/A15	=A15-\$K\$7	=F15/\$K\$36				проверк	
16	38	=LN(A16)	=1/A16	=A16-\$K\$7	=F16/\$K\$36			простая		
17	76	=LN(A17)	=1/A17	=A17-\$K\$7	=F17/\$K\$36			=ПРОИЗВЕД(A6:A17)^(1/K3)		
18										
19								3. Средняя квадратичная		
20								простая		взвешенная
21								=КОРЕНЬ(СУММПРОИЗВ(A6:A17;A6:A17)/K3)		=КОРЕНЬ(СУММПРОИЗВ(A6:A17;A6:A17;B6:B17)/СУММ(B6:B17))
22										
23								4. Средняя гармоническая		
24								простая		взвешенная
25								=СРГАРМ(A6:A17)		=СУММ(B6:B17)/СУММПРОИЗВ(E6:E17;B6:B17)
26									проверк	
27								простая		
28								=K3/СУММ(E6:E17)		
29										
30								5. Лимиты		
31								Максимальное		Минимальное
32								=МАКС(A6:A17)		=МИН(A6:A17)
33										
34								6. Среднее квадратичное отклонение		
35								простая		взвешенная
36								=КОРЕНЬ(КВАДРОТКЛ(A6:A17)/(K3-1))		=КОРЕНЬ(СУММПРОИЗВ(F6:F17;F6:F17)/(K3-1))
37										
38								7. Коэффициент вариации		
39								через простую		через взвешенную
40								=I36/I7*100		=K36/K7*100
41										

ЛАБОРАТОРНАЯ РАБОТА №2

Тема: «Отображение и выравнивание распределения признака в выборках с помощью графических и аналитических методов»

Цель работы – с помощью стандартных функций и встроенных инструментов редактора электронных таблиц Microsoft Excel научиться выполнять анализ, графическое и аналитическое выравнивания показателей распределения признаков в выборках.

Общие сведения и понятия о распределениях признака

Разнообразие объектов составляющих группу — это основное свойство всякой совокупности. В малочисленных группах трудно определить какую-либо закономерность в разнообразии дат. По мере увеличения численности изучаемых групп все больше проявляются закономерности в разнообразии, которые скрыты (незаметны) в малочисленных группах.

Если имеется многочисленная группа, то различные значения признака встречаются в этой группе разное число раз. Это явление называется распределением признака.

При изучении многих совокупностей по различным признакам определяется несколько типов распределения признака в группе. В исследованиях наибольшее значение имеют:

- нормальное распределение;
- биномиальное распределение;
- распределение редких событий (Пуассона).

Изобразить распределение признака можно следующими основными способами:

- вариационным рядом
- 2) гистограммой;
- 3) вариационной кривой;
- 4) кумулятой.

Вариационный ряд – это упорядоченное отражение реально существующего распределения значений признака, по отдельным особям изученной группы. Вариационный ряд представляет собой двойной ряд чисел, состоящий из обозначений классов и соответствующих им частот.

Вариационный ряд позволяет без конкретных вычислений определить величину среднего уровня признака и разнообразия.

Составление вариационного ряда.

При составлении вариационного ряда все величины признака разбиваются на равные интервалы—классы. Предварительно необходимо установить число классов, их величину, границы, середины и частоты.

Число классов. Весь размах значений признака от минимума до максимума разделяется обычно на 5—12 равных интервалов. При точных исследованиях устанавливается число классов:

$$R = 1 + 3,3 \cdot \lg(n)$$

с последующим округлением до ближайшего целого вверх.

Величина классов или величина классового промежутка равна размаху значений от минимума до максимума, деленному на принятое число классов. Обычно величина классов устанавливается по формуле:

$$K = \frac{V_{\max} - V_{\min}}{1 + 3,3 \cdot \lg(n)}$$

где V_{\max} — максимальное значение, V_{\min} — минимальное значение.

Полученное дробное число при делении округляют до ближайшего целого числа. Например, если получено 43,4, то за величину классового промежутка K нужно взять 45.

Границы классов. Конец каждого класса должен быть меньше начала следующего на величину, равную принятой точности измерения (ξ).

$$\begin{aligned} W_{s(l)} &= V_{\max} + 0,5 \cdot K - \xi \\ W_{s(l)} &= V_{\max} - 0,5 \cdot K \end{aligned}$$

Например, если измеряется длина с точностью до $\xi = 1$ см и установлена величина классового промежутка 5 см, то границы классов, начиная с нижнего минимального будут такими: 100—104, 105—109, 110—114, 115—119 и т. д.

Средины классов устанавливаются двумя способами. Если признак может быть выражен любым числом — и целым и дробным, то для установления середины класса нужно к началу класса прибавить половину классового промежутка.

В тех случаях, когда признак выражается только целыми числами, середина классов равна полусумме начала и конца класса.

Частоты классов устанавливаются путем разности дат по классам, которые определены описанным выше способом. Обозначаются частоты классов символом f . Каждая дата, попав в соответствующий класс, приравнивается по величине ко всем другим датам, попавшим в этот класс.

Гистограмма — вариационный ряд представленный в виде диаграммы, в которой различная величина частот изображается различной высотой столбцов. На гистограмме наглядно проявляются особенности распределения.

Вариационная кривая — это изображение вариационного ряда в виде кривой, ординаты которой пропорциональны частотам вариационного ряда. Вариационная кривая — это удобный и наглядный способ иллюстрации

вариационного ряда в тех случаях, когда на одном графике нужно расположить или изобразить несколько распределений.

Кумулята – это изображение распределения признака в виде кривой, ординаты которой пропорциональны накопленным частотам вариационного ряда. Чтобы сопоставить ряд накопленных частот, нужно к частотам наименьшего класса прибавить частоту следующего класса, т.е. определить сумму частот по всем классам.

Кумулята имеет преимущество перед вариационной кривой в случае изучения процесса накопления какого-либо признака. Один из этих методов – метод накопления (пробит) – показывает детали действия ядов и отравляющих веществ в живых организмах и в природе в целом.

Нормальное распределение

В большинстве распределений, с которыми приходится встречаться проявляется определенная закономерность:

1. Крайние значения — наименьшее и наибольшее — появляются редко;

2. Чем ближе значение признака к средней арифметической, тем оно чаще встречается;

3. В центре распределения имеются такие значения, которые встречаются наиболее часто и образуют в вариационном ряду модальный класс.

Такое распределение значений признака так часто встречается в самых различных областях науки и практики, что первоначально оно принималось за норму всякого массового случайного проявления признаков и в соответствии с этим получило особое название - нормальное.

Закон нормального распределения выражается формулой:

$$p^* = \frac{n \cdot k}{\sigma} \cdot \frac{1}{\sqrt{2 \cdot \pi}} \cdot \exp\left(-\frac{x^2}{2}\right)$$

где: p^* — теоретическая частота каждого класса распределения;

n — объем группы, число объектов исследования;

k — классовый промежуток (величина классов);

σ — среднее квадратичное отклонение; $\sigma = \sqrt{\frac{\sum_{i=1}^n (V_i - M)^2}{n - 1}}$

$x = \frac{W - M}{\sigma}$ — нормированное отклонение средин каждого класса распределения W от среднего M .

Биномиальное распределение

Группа значений может изучаться не только по количественным признакам, которые могут иметь различную степень своего проявления и измеряются именованными величинами — в килограммах, литрах, сантиметрах и других единицах измерения. Есть признаки, которые обычно не имеют градаций (мужской пол, красная масть и др.). У каждого отдельного элемента распределения такой признак может присутствовать или отсутствовать. Такие признаки называются качественными или альтернативными.

Характеристика группы по качественному признаку заключается в указании, сколько в этой группе имеется элементов с наличием данного признака и у скольких элементов его нет. Для такой характеристики употребляются следующие обозначения:

n — общее количество элементов в группе;

n_+ — количество элементов, имеющих изучаемый признак;

n_- — количество элементов, не имеющих данного признака;

$p = \frac{n_+}{n}$ - доля элементов, имеющих признак;

$q = \frac{n_-}{n}$ - доля элементов, не имеющих признак;

очевидно следующее равенство: $p + q = 1$.

Получаемое распределение называется биномиальным. Такое название объясняется следующим:

1. признак может иметь только два варианта: он есть или его нет;
2. закономерности таких распределений имеют количественное выражение, связанное с коэффициентами разложения бинома Ньютона, который в применении к этому типу распределений может быть выражен следующим образом:

$$1 = (p + q)^n = \sum_{i=0}^n \frac{n!}{i!(n-i)!} \cdot p^i q^{n-i}$$

где i — количество элементов в группе, которые имеют изучаемый признак;

или в развернутом виде:

$$\sum_{i=0}^n \cdot (p + q)^n = \frac{1}{1} p^0 \cdot q^n + \frac{n}{1} \cdot p q^{n-1} + \frac{n \cdot (n-1)}{1 \cdot 2} \cdot p^2 q^{n-2} + \frac{n \cdot (n-1)(n-2)}{1 \cdot 2 \cdot 3} p^3 q^{n-3} \dots + \frac{1}{1} \cdot p^n q^0$$

Решая формулу бинома относительно величины p , можно получить следующие значения:

$p^0 \cdot q^n$ — нулевой член бинома (содержащий p в нулевой степени), дает ожидаемую долю таких равночисленных групп, в которых из n элементов ни один не имеет изучаемого признака;

$np^1 q^{n-1}$ — первый член бинома (с p^1), дает долю групп, в которых только один элемент имеет ожидаемый признак;

$\frac{n(n-1)}{2} p^2 q^{n-2}$ - второй член бинома (с p^2), дает долю групп, в которых изучаемый признак имеют два элемента;

p^n — последний член бинома, дает долю равночисленных групп, в которых все n элементов имеют изучаемый признак.

Распределение редких событий (Пуассона)

События, происходящие редко, один или небольшое число раз на 1000, 10000 и больше обычных явлений, могут быть сведены в особое распределение, в котором вариациями являются различные число редких случаев, а частотами — количество больших групп, среди которых редкое событие произошло определенное число раз.

Распределения редких, случайных событий обычно подчиняются определенному закону, который выражается формулой, предложенной Пуассоном:

$$p_x^* = p \cdot e^{-a} \cdot \frac{a^x}{x!}$$

где p - теоретическая частота распределения, ожидаемое число больших групп, среди которых редкое событие произошло x раз;

$p = \sum p_x$ - общее количество исследованных больших групп;

p_x - фактическая (эмпирическая) частота распределения;

x - число редких событий, происшедших в каждой большой группе; обычно x равно небольшому целому числу: 0, 1, 2, 3 и т. д.;

$x!$ — произведение натуральных чисел от 1 до x (факториал).

Считается, что факториал нуля равен единице: $0! = 1$;

$a = \frac{\sum x \cdot p_x}{p}$ - средняя встречаемость или среднее число редких случаев

на каждую большую группу.

Теоретическое распределение редких событий имеет одну особенность: в нем значение средней величины примерно равно квадрату сигмы (девиате). Поэтому, если в других распределениях основных величин две — M и σ , то в распределении редких событий обе основные величины сведены к одной — a .

Из этой особенности распределения редких событий вытекают два следствия:

1. все теоретическое распределение может быть построено на основании только одной средней;

2. при определении достоверности отличия теоретического распределения от эмпирического при помощи критерия χ^2 число степеней свободы равно числу классов без одного.

Для определения вида распределения (нормальное, биномиальное или Пуассона) изучаемого признака производят сопоставление эмпирических и теоретических частот этого распределения.

Нахождение ряда теоретических частот для имеющегося эмпирического распределения называется выравниванием эмпирических кривых по нормальному или другому закону. Этот процесс имеет очень большое теоретическое и практическое значение. Выравнивание эмпирических кривых вскрывает закономерность распределения, которая обычно скрыта под случайной формой своего проявления.

Достоверность различия распределений

Для практических и научных работ необходимо установить сильно или слабо расходятся эмпирический и теоретический ряды. Необходимо установить такой предел, не достижение которого означает что расхождение между эмпирическим и нормальным (биномиальным или Пуассона) распределением еще не настолько велико, чтобы с ним считаться, и что данный эмпирический ряд еще можно практически принять за искомый закон распределения.

Для этой цели применяются особые показатели, из которых в исследованиях используются критерий χ^2 (хи-квадрат) и критерий λ (лямбда).

Критерий χ^2 (хи-квадрат)

Критерий χ^2 применяется во всех случаях, когда необходимо определить степень отличия фактического распределения частот от теоретического. Определяется величина χ^2 по следующей формуле:

$$\chi^2 = \sum_{j=1}^m \frac{(f_j - p^*)^2}{p_j^*},$$

где f — эмпирическая частота, p^* — теоретическая частота.

Если крайние классы распределения имеют теоретические частоты меньше единицы, то при вычислении χ^2 их необходимо объединить в один класс вместе с ближайшим классом, имеющим частоты $p^* > 1$. Вместе с теоретическими надо объединить и соответствующие фактические частоты.

После нахождения величины χ^2 требуется определить, велика или мала она для данного распределения. Для этого пользуются таблицей предельных значений χ^2 .

Для того чтобы пользоваться этой таблицей, необходимо предварительно установить число степеней свободы для изучаемого

распределения. Если в качестве теоретического распределения берется нормальное, все детали которого определяются двумя постоянными величинами M и σ , то число степеней свободы в таких случаях равно числу классов без двух. За число классов берется то, которое получилось после объединения классов с дробными теоретическими частотами.

$$\chi^2 = \sum_{j=1}^m \frac{(f_j - P_j^*)^2}{P_j^*} \geq \chi^2_{st}$$

{ b_1 = при малой ; b_2 – при обычной; b_3 – при большой }
ответственности исследования

Для каждого числа степеней свободы указаны три цифры предельных значения χ^2 , соответствующие трем стандартным степеням вероятности ($P_3=0,999$, $P_2=0,99$ и $P_1=0,95$) того, что распределения, показавшие такие значения χ^2 или большие, различаются достоверно. Под достоверным различием понимается такое расхождение распределений, которое не может произойти в порядке обычных случайных отклонений фактических частот от теоретических.

Если фактическое значение χ^2 больше третьего значения, соответствующего $P_3=0,999$, то во всех случаях можно считать различие между распределениями достоверным.

Критерий λ (лямбда)

Критерий λ применяется для определения достоверности расхождения между фактическими и теоретическими распределениями, а также различий между любыми двумя распределениями частот одного и того же признака даже в том случае, когда число классов и число дат у этих распределений неодинаково. Для применения критерия λ не требуется определять число степеней свободы и не нужны таблицы для определения, трех предельных значений критерия, так как для любого числа классов эти предельные значения одинаковы: 1,36; 1,63; 1,95 и соответствуют обычным трем степеням вероятности достоверного различия — $P_1=0,95$; $P_2=0,99$; $P_3=0,999$.

Для сравнения эмпирического распределения с теоретическим, при одинаковом числе классов и при одинаковой общей численности групп критерий лямбда (λ) определяется по формуле:

$$\lambda = \frac{|d|}{\sqrt{n}} = \frac{\left| \sum_{i=1}^m f_i - \sum_{i=1}^m p_i^* \right|_{\max}}{\sqrt{n}}$$

где d — максимальная разность (без учета ее знака) между накопленными частотами в эмпирическом и теоретическом распределениях

для одного и того же класса; n — общее число дат, образовавших эмпирическое распределение.

Для определения критерия лямбда требуется составить ряды накопленных частот для обоих сравниваемых распределений взять наибольшую разность (без учета ее знака) между этими величинами полученную разность разделить на \sqrt{n} .

Основные сведения о построении диаграмм и графиков в Microsoft Excel

В Microsoft Excel имеется возможность графического представления данных в виде диаграмм и графиков. Построенные таким образом графические образы поддерживают «горячую» связь с данными листа, на основе которых они были созданы, и изменяются каждый раз, когда изменяются эти исходные данные. Можно создать внедренную диаграмму или лист диаграммы.

В Microsoft Excel существует возможность построения различных типов диаграмм. Тип диаграммы выбирается в зависимости от используемых данных и от того, каким образом необходимо их представить. В ходе выполнения лабораторных работы будут использованы три стандартных типа диаграмм:

1) Гистограмма (показывает изменение данных за определенный период времени и иллюстрирует соотношение отдельных значений данных) - для построения гистограммы;

2) График (отражает тенденции изменения данных за равные промежутки времени) - для построения вариационной кривой;

3) Точечная диаграмма (отображает взаимосвязь между нечетными интервалами числовых значений, расположенных в нескольких рядах, и представляет две группы чисел в виде одного ряда точек в координатах XY) – для построения куммуляты. Создание любой диаграммы в Microsoft Excel можно выполнить с помощью кнопки на панели инструментов или команды Вставка=> Диаграмма. Для этого необходимо выполнить следующее:

1) выделить диапазоны ячеек с данными;

2) вызвать Мастер диаграмм (кнопка или Вставка=> Диаграмма);

3) в появившемся окне диалога (шаг 1 из 4, рис. 2.1) выбрать нужный Тип и Вид диаграммы. Щелкнуть Далее;

4) во втором окне диалога (шаг 2 из 4, рис. 2.2) можно изменить диапазоны данных для построения, указать их расположение (по столбцам/строкам), задать имена, и подписи по оси X. Щелкнуть Далее;

5) в третьем окне диалога (шаг 3 из 4, рис. 2.3) можно задать подписи значений (над столбцами), добавить внизу таблицу с исходными данными, легенду (название рядов) и ее расположение, линии сетки осей X и Y, заголовки (название диаграммы и осей). Щелкнуть Далее;

6) в последнем окне (шаг 4 из 4, рис. 2.4) указать место расположения: отдельный лист (диаграмм) или текущий. Построенную диаграмму можно:

- 1) перемещать и изменять ее размеры;
- 2) улучшать внешний вид;
- 3) распечатывать.

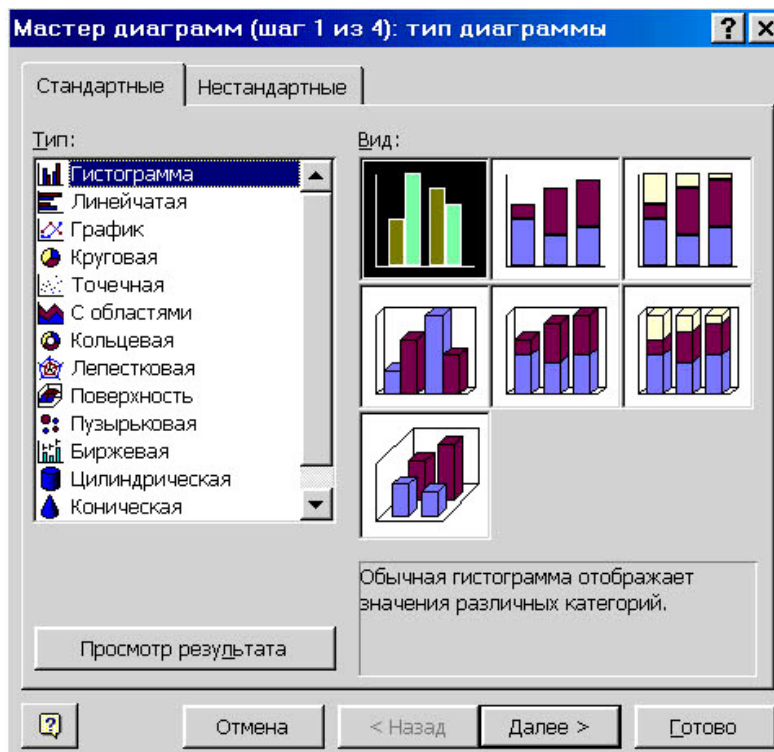


Рисунок 2.1. Первое окно Мастера диаграмм

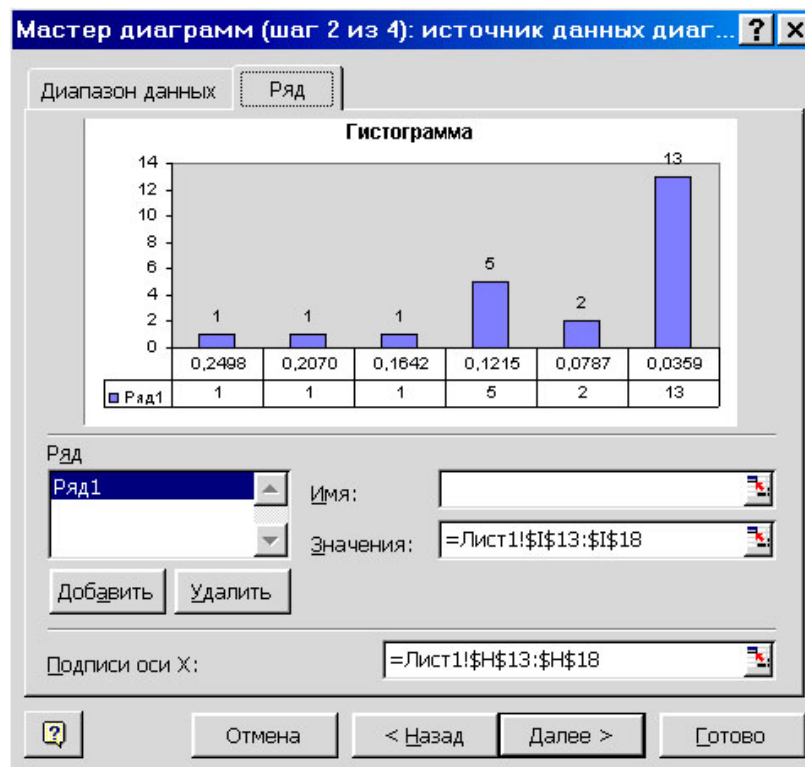


Рисунок 2.2. Второе окно Мастера диаграмм



Рисунок 2.3. Третье окно Мастера диаграмм

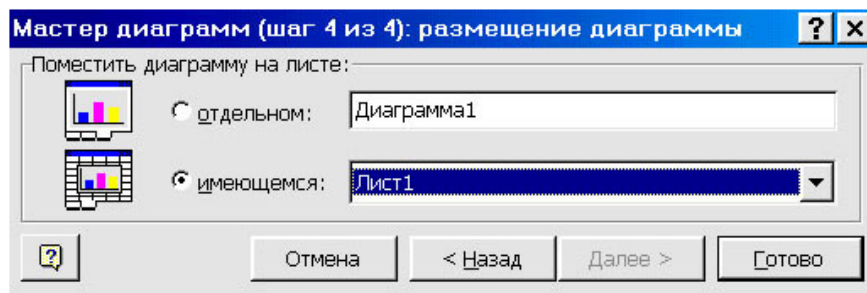


Рисунок 2.4. Четвертое окно Мастера диаграмм

Улучшение внешнего вида диаграмм

Каждая диаграмма состоит из нескольких независимых друг от друга элементов. Чтобы изменить какой-либо элемент необходимо щелкнуть по нему или выбрать его в списке панели инструментов Диаграмма. Выбранный элемент ограничивается прямоугольником с расположенными на нем маркерами (черными квадратными точками). С помощью маркеров можно изменять размеры элементов.

Изменение типа диаграммы. Выделить полностью всю диаграмму (или Элементы диаграммы>Область диаграммы), а затем Тип диаграммы (или правая кнопка в контекстном меню).

Добавление легенды, заголовков, и др. элементов. Для этого:

- 1) выбрать Область диаграммы;
- 2) в контекстном меню (правая кнопка) выбрать Параметры диаграммы>Заголовки (или Легенда/Оси/Линии сетки/Таблица данных/Подписи данных). Чтобы добавить текст, который не является стандартным элементом диаграммы, используют кнопку Надпись панели Рисование. Аналогично текст можно добавить и в область рабочего листа.

Форматирование текстовых и числовых данных. Данные, помещенные на диаграмму, находятся в собственном прямоугольном поле. Чтобы улучшить их отображение изменяют их формат. Для этого:

- 1) выбрать элемент диаграммы;
- 2) в контекстном меню выполнить Формат выбранного элемента (название зависит от элемента);
- 3) далее можно использовать вкладки Вид (для меток делений и линий осей), Шкала (начальное и конечное значение, разворот оси, точка пересечения осей), Шрифт (гарнитура, размер, нижний, верхний), Число (формат значений), Выравнивание (наклон текста).

Для трехмерных диаграмм можно изменять их угол наклона или вращать. Для этого используют элемент Углы.

Чтобы добавить дополнительные (вспомогательные) оси необходимо выбрать элемент Ряд и далее Формат рядов данных. В окне диалога использовать вкладку Ось. Необходимые стандартные функции Microsoft Excel

Для выполнения расчетов в этой лабораторной работе наибольший интерес будут представлять стандартные статистические и математические функции представленные в таблице ниже.

Таблица 2.1 - Стандартные статистические и математические функции

Функция	Действие	Пример
КВАДРОТКЛ (a1;a2;...)	Возвращает сумму квадратов отклонений от их среднего	Ячейки A1:G1 содержат числа 4; 5; 8; 7; 11; 4; 3 КВАДРОТКЛ(A1:G1) → 48
СЧЁТ(a1;a2;...)	Возвращает количество дат (чисел)	Ячейки A1:E1 содержат числа 4;5;8; 7; 1 СЧЁТ(A1:E1) → 5
ЧАСТОТА (массив выборки; массив классов)	Определяет частоту появления значений в выборке	ЧАСТОТА(A1:J10; A12:H12) → 1, 7, 19, 30, 25, 10, 7, 1
ХИ2ОБР(1-bi;v)	Возвращает χ^2_{st}	ХИ2ОБР(1-0,95;4) → 9,4
НОРМРАСП (W ;M;σ;0)	Возвращает отдельное значение нормального распределения	НОРМРАСП(A12;\$A\$14;\$B\$14;0)* 10 → 1,4970
БИНОМРАСП (i;n;p;0)	Возвращает отдельное значение биномиального распределения	БИНОМРАСП(2;6;0,85;0)*23 → 0,126

ПУАССОН($x;a;0$)	Возвращает отдельное значение распределения Пуассона	ПУАССОН(4;0,499;0) * 1000 → 1,6
НОРМАЛИЗАЦИЯ ($W;M;\sigma$)	Возвращает нормированное отклонение x	НОРМАЛИЗАЦИЯ(A12;\$A\$14;\$B\$14) → 2,4315
ОКРУГЛВВЕРХ ($a1;z$)	Возвращает округленное (до z знаков после запятой) вверх значение $a1$.	ОКРУГЛВВЕРХ(5,6789;0) → 6 ОКРУГЛВВЕРХ(5,6789;2) → 5,68
СРЗНАЧ($a1;a2;..$)	Возвращает среднее арифметическое	Ячейки A1:E1 содержат числа 10, 7, 9, 27 и 2, то: СРЗНАЧ(A1:E1) → 11
LOG($a1$)	Возвращает десятичный логарифм числа $a1$	Ячейка A1 содержит 100 LOG(A1) → 2
EXP($a1$)	Возвращает экспоненту числа $a1$	Ячейка B1 содержит 4,60517 EXP(B1) → 100
МИН($a1;a2;...$)	Возвращает минимальное значение	МИН(A1:J10) → 380
МАКС($a1;a2;...$)	Возвращает максимальное значение	МАКС(A1:J10) → 450

Контрольные вопросы

1. Что называют выборкой и распределением признака в выборке.
2. Какие типы распределений рассматриваются в исследованиях.
3. Перечислите способы, с помощью которых можно изобразить распределение признака в выборке.
4. В выполненной лабораторной работе найдите и покажите вариационный ряд, гистограмму, вариационную кривую и куммуляту.
5. Из каких этапов складывается составление вариационного ряда.
6. Назовите основные свойства нормального распределения.
7. В каких случаях для характеристики распределения признака в выборке используют биномиальное распределение. Приведите примеры.
8. Дайте характеристику распределения Пуассона. Приведите примеры этого распределения.

9. С помощью, каких показателей проводят оценку достоверности различий между эмпирическими и теоретическими распределениями.

10. Сколько порогов вероятности безошибочных прогнозов рассматривают на практике и чему они равны.

11. На основании сравнения, каких данных в выполненной лабораторной работе был, сделан вывод о соответствии (или не соответствии) эмпирического распределения нормальному, биномиальному и Пуассона.

Задание к лабораторной работе №2

1. На основе выданных значений, полученных в результате экспериментального изучения количества химического вещества (*мг*) в пробах (*л*), построить: вариационный ряд, гистограмму, вариационную кривую и кумуляту.

2. Выполнить выравнивание значений вариационного ряда по нормальному закону.

3. Выполнить выравнивание значений вариационного ряда по биномиальному закону. Подобрать долю (вероятность) положительных исходов (*p*) так, чтобы по критерию χ^2 («хи-квадрат») расхождение между эмпирическим и биномиальным распределением было бы наименьшим.

4. Выполнить выравнивание значений вариационного ряда в предположении, что их распределение подчиняется распределению редких событий (Пуассона).

5. Используя критерии χ^2 («хи-квадрат») и λ (лямбда) сделайте вывод о достоверности различий распределения эмпирического от теоретического.

6. Окончательно сохраните и распечатайте все полученные результаты.

Порядок выполнения

Часть 2.1

1. Набрать исходные выданные значения в два рядом стоящих столбца, например, столбцы А и В.

2. В третьем столбце (С) рассчитать содержание вещества в единице объема пробы.

3. Для построения вариационного ряда принять точность исходных данных 0,0001. Используя функции МАКС(), МИН(), СЧЁТ() и СРЗНАЧ() найти минимальное, максимальное, число дат выборки и среднее содержание вещества.

4. По приведенным выше формулам рассчитать размах, размер класса (*K*), число классов (*R*) и границы (верхнюю, нижнюю) классов. С помощью функции массива {ЧАСТОТА()}, которую создают через <F2>, <Ctrl+Shift+Enter>, определить частоты для каждого класса. По частотам рассчитать кумуляту (см. пример).

5. Используя Мастер Диаграмм построить гистограмму и вариационную кривую. Для этого выделить частоты вариационного ряда,

вызвать Мастер диаграмм и выбрать тип Гистограмма. Для построения вариационной кривой используйте тип График.

6. В виде отдельной диаграммы аналогично вариационной кривой строят кумуляту.

7. Далее найти теоретические частоты нормального распределения. Для этого, используя функции СРЗНАЧ() определяют середины классов, а с помощью КВАДРОТКЛ() среднее квадратическое отклонение (сигма - σ). По найденным значениям M , σ , n и серединам классов эмпирического распределения с использованием функции НОРМРАСП() рассчитывают теоретические частоты нормального распределения.

8. Если необходимо выполняют редукцию классов и рассчитывают критерий χ^2 («хи-квадрат») для двух порогов вероятностных прогнозов (0,95 и 0,99). Для определения стандартного χ^2_{st} используют функцию ХИ2ОБР().

9. Используя приведенные выше формулы, рассчитать критерий лямбда (см. пример).

10. На основе полученных критериев сделать вывод о соответствии эмпирического распределения нормальному.

11. Сохраните и распечатайте полученные результаты.

Часть 2.2

12. На втором листе книги аналогично пунктам 1-4, 6, 8-10 выполнить выравнивание значений вариационного ряда по биномиальному закону. С помощью стандартной функции БИНОМРАСП() рассчитывают теоретические частоты биномиального распределения.

13. Методом последовательных приближений (в диапазоне от 0 до 1) подобрать долю (вероятность) положительных исходов (p) так, чтобы по критерию χ^2 («хи-квадрат») расхождение между эмпирическим и биномиальным распределением было бы наименьшим.

14. Используя Мастер Диаграмм построить точечные диаграммы эмпирического и биномиального распределения (см. пример).

15. Сохраните и распечатайте полученные результаты.

Часть 2.3

16. На третьем листе книги аналогично пунктам 1-4, 6, 8-10 выполнить выравнивание значений вариационного ряда по закону распределения редких событий (Пуассона). С помощью стандартной функции ПУАССОН() рассчитывают теоретические частоты распределения Пуассона.

17. Используя Мастер Диаграмм построить точечные диаграммы эмпирического распределения и распределения Пуассона (см. пример).

18. Сохраните и распечатайте полученные результаты.

ЛАБОРАТОРНАЯ РАБОТА №3

Тема: «Корреляционный и регрессионный анализ экспериментальных данных в электронных таблицах»

Цель работы – с помощью стандартных функций и встроенных инструментов редактора электронных таблиц Microsoft Excel научиться выполнять корреляционный и регрессионный анализ эмпирических данных, которые могут быть использованы экспериментальных исследованиях.

Общие сведения и понятия о корреляционном и регрессионном анализе

Во многих исследованиях часто изучается несколько признаков в их взаимной связи. Если вести такое исследование по отношению к двум признакам, то можно заметить, что изменчивость одного признака находится в некотором соответствии с изменчивостью другого.

В некоторых случаях такая зависимость проявляется настолько сильно, что при изменении первого признака на определенную величину всегда изменяется и второй признак на определенную величину, поэтому каждому значению первого признака всегда соответствует совершенно определенное, единственное значение второго признака. Такие связи называются функциональными.

Функциональные связи встречаются в физических и математических обобщениях. Площадь треугольника точно определяется его высотой и основанием, длина окружности — радиусом, скорость падения является функцией времени падения и ускорения силы тяжести, скорость протекания определенной химической реакции находится в зависимости от температуры.

Необходимо учесть, что в чистом виде функциональные связи встречаются только в идеальных условиях, когда предполагается, что никаких посторонних влияний нет.

На практике это недостижимо. Никогда нельзя точно измерить фактически имеющийся радиус круга, причем вычисленная площадь никогда не равна в точности фактической, вследствие практической невозможности начертить точную окружность. Скорость падения реального тела в реальных условиях будет всегда различна при одних и тех же времени и ускорении силы тяжести. На практике всегда действуют посторонние для данной функциональной зависимости факторы, которые нарушают точность этой зависимости в разных случаях по-разному.

Пока такие нарушения остаются настолько незначительными, что их практически можно не учитывать, связь считается функциональной.

При изучении некоторых объектов приходится иметь дело со связями другого рода. Тут каждому определенному значению первого признака соответствует не одно значение второго признака, а целое распределение этих значений при вполне определенных основных показателях этого

частного распределения — средней величины и степени разнообразия. Такая связь называется корреляционной связью или просто корреляцией.

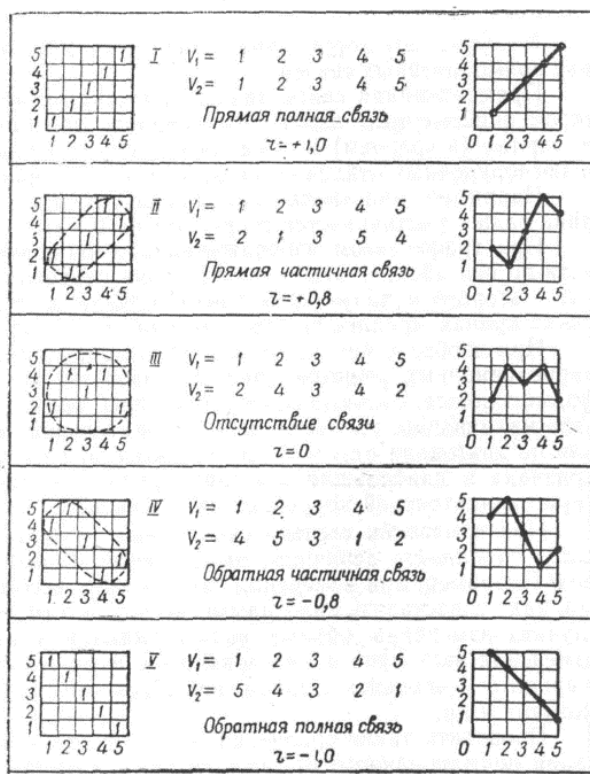
По форме корреляция может быть – прямолинейной и криволинейной. По направлению - прямой и обратной. Степень корреляции устанавливает силу связи между количественными и качественными признаками. Она измеряется следующими показателями:

- 1) коэффициентом корреляции r ;
- 2) корреляционным отношением η ;
- 3) тетракорическим и поликорическим показателями связи;
- 4) частным и множественным коэффициентами корреляции.

Изобразить корреляционную связь двух признаков можно следующими способами (рис. 3.1):

1) При помощи корреляционного ряда, состоящего из ряда пар значений признаков.

2) При помощи линии регрессии, абсциссы которой пропорциональны значениям признаков.



Коэффициент корреляции. Коэффициент корреляции измеряет степень и определяет направление прямолинейных связей. Прямолинейная связь между признаками, это такая связь, при которой равномерным изменениям первого признака соответствуют равномерные (в среднем) изменения второго признака.

При измерении степени связи между разными признаками используют их нормированные отклонения, а коэффициент корреляции (r) имеет следующую простую формулу:

Рисунок 3.1 – Схема корреляционных связей

$$r = \frac{\sum_{j=1}^n x_{1j} \cdot x_{2j}}{v}$$

где $x_1 = \frac{V - M_v}{\sigma_v}$, $x_2 = \frac{Y - M_y}{\sigma_y}$ - нормированные отклонения признаков;

$$v - \text{число степеней свободы } v=n-1; \sigma_v = \sqrt{\frac{\sum_{j=1}^n (V_j - M_v)^2}{v}}, \sigma_y = \sqrt{\frac{\sum_{j=1}^n (Y_j - M_y)^2}{v}}$$

Сумма произведений нормированных отклонений, входящая в формулу для коэффициента корреляции, обладает следующими тремя особыми свойствами:

1) Если оба признака изменяются параллельно, то сумма произведений их нормированных отклонений дает положительную величину. Если при увеличении одного признака другой уменьшается, то вся сумма будет отрицательной.

Поэтому коэффициент корреляции определяет направление связи: при прямых связях он положителен, а при обратных — отрицателен.

2) При полных связях, когда изменения обоих признаков строго соответствуют друг другу и корреляционная связь превращается в функциональную, сумма произведений нормированных отклонений становится равной числу степеней свободы: $\sum_{j=1}^n x_{1j} \cdot x_{2j} = v = n - 1$

Поэтому максимальное значение коэффициента корреляции равно 1:

- для прямых связей: +1

- для обратных связей: -1

3) При полном отсутствии корреляционной связи между признаками коэффициент корреляции равен нулю.

На практике коэффициент корреляции рассчитывают по развернутой формуле:

$$r = \frac{\sum_{i=1}^n ((V_i - M_v) \cdot (Y_i - M_y))}{\sqrt{\sum_{i=1}^n (V_i - M_v)^2} \cdot \sqrt{\sum_{i=1}^n (Y_i - M_y)^2}}$$

где M_v и M_y — среднеарифметические для $\{V\}$ и $\{Y\}$.

Ошибка коэффициента корреляции. Как и всякая выборочная величина, коэффициент корреляции имеет свою ошибку репрезентативности, вычисляемую для малых выборок ($n < 100$) по формуле:

$$m_r = \sqrt{\frac{1 - r^2}{n - 2}}$$

где r — коэффициент корреляции в выборке; n — численность выборки, т.е. число пар значений, по которым вычислялся выборочный коэффициент корреляции.

Критерий достоверности коэффициента корреляции, определяемый по формуле:

$$t_r = \frac{r}{m} \geq t_{st}$$

оценивается путем сравнения фактически полученного значения t_r с фиксированными значениями t_{st} (стандартное значение критерия Стьюдента), соответствующими трем степеням вероятности безошибочных прогнозов. При этом $v = n - 2$.

Если расчетное $t_r \geq t_{st}$, то коэффициент корреляции достоверен, и можно считать, что между исследуемыми признаками существует взаимосвязь.

Если $t_r < t_{st}$, то коэффициент корреляции недостоверен, и нельзя сделать вывод о взаимосвязи между признаками или в генеральной совокупности.

Регрессией называется изменение функции при определенных изменениях одного или нескольких аргументов.

Функция – это признак, зависящий от других признаков - аргументов.

Зависимость функции от аргументов может быть физиологической или условно принятой в исследовании.

Примером физиологической зависимости может служить зависимость веса животного (функции) от его возраста (аргумента).

Если по длине определяется вес животного, считается что вес зависит от длины, если необходимо предусмотреть размеры животных разного веса, то принимается, что длина зависит от веса. Это пример условной зависимости.

Вскрыть функцию – это значит найти закономерности по которым изменяется изучаемый признак в зависимости от изменения одного или нескольких признаков. Если изменения функции исследуется в зависимости от одного признака, то регрессия называется простой:

$$Y = f(x) \Leftrightarrow x = f(Y)$$

Если изучается зависимость изменения функции – от изменения нескольких признаков, регрессия называется множественной.

$$Y = f(x_1, x_2, \dots, x_n)$$

Если при одинаковом приращении аргумента, но при разных его значениях (малых, больших или средних) функция имеет неодинаковое приращение, причем среднее изменение не идет по прямой, то регрессия называется криволинейной. Если при любом значении (малом, среднем или большом) аргумента одинаковое изменение его приводит к одинаковому изменению значения функции, то регрессия называется прямолинейной.

Регрессию принято отображать в виде:

- регрессионного ряда (эмпирического или теоретического);
- линии регрессии (эмпирического или теоретического);
- коэффициентов регрессии образующих уравнение регрессии.

По виду эмпирической линии регрессии, можно установить на каких участках изменения аргумента, признак (функция) развивалась в лучших или худших условиях.

Процесс получения усредненного течения изменения функции при равномерном увеличении значения аргумента – называется выравниванием

эмпирических рядов. В биометрических расчетах наибольшее распространение получили следующие способы выравнивания эмпирических данных:

1. Графический.

2. Способ скользящих средних – используется, если форма функции неизвестна. Состоит в следующем: для каждого значения аргумента берется среднее арифметическое от нескольких средних значений аргумента. Минимальное количество должно быть три значения аргумента.

В случае, когда необходимо сохранить крайние значения используют следующий способ.

3. Способ взвешенной скользящей средней. В этом случае с обоих концов ряда добавляется по два значения ряда и производят выравнивание

$$x_{n+2} = \frac{2 \cdot x_n + x_{n-1} - x_{n-3}}{2}$$

4. Выравнивание с помощью метода наименьших квадратов (МНК). Этот метод предоставляет наиболее универсальную возможность для выравнивания и определения вида аналитической функции (прямолинейной, обратной, параболической, гиперболической, степенной, логарифмической, экспоненциальной, периодической и комбинации их) приближенно заменяющей табличные данные полученные экспериментальным путем. МНК предназначен для выбора из совокупности назначенного типа кривых такой, для которой сумма квадратов отклонений эмпирических данных ($Y_{\text{э}}$) от вычисленных (Y_p) по найденной зависимости является наименьшей:

$$(Y_{\text{э}} - Y_p)^2 \Rightarrow \min$$

Например, определение неизвестных коэффициентов линейной зависимости вида:

$$Y_p = a + b \cdot V$$

МНК предлагает следующие формулы:

$$b = \frac{n \cdot \sum_{i=1}^n V_i \cdot Y_{\text{э}i} - \sum_{i=1}^n V_i \cdot \sum_{i=1}^n Y_{\text{э}i}}{n \cdot \sum_{i=1}^n V_i^2 - (\sum_{i=1}^n V_i)^2} \quad a = \frac{1}{n} \cdot (\sum_{i=1}^n Y_{\text{э}i} - b \cdot \sum_{i=1}^n V_i)$$

где n – количество пар дат.

Для функций отличных от линейной их приводят к нормальному виду используя различные математические приемы:

- Логарифмирование ($\text{LOG}(V)$, $\text{LN}(V)$) и потенцирование (10^V , $\text{EXP}(V)$);

- извлечение корня ($\text{КОРЕНЬ}(V)$, $V^{(1/n)}$) и возведение в степень (V^n);

- замена функций на обратные им ($1/Y$);

- агрегирование аргументов и функций;

- разделение аргументов и функций;

- комбинацию перечисленных выше приемов.

Необходимые стандартные функции Microsoft Excel

Для выполнения расчетов в этой лабораторной работе наибольший интерес будут представлять следующие новые стандартные статистические и математические функции, которые представлены в таблице ниже.

Таблица 3.1 - Стандартные статистические и математические функции

Функции	Действия	Пример
НАКЛОН(массив Y; массив V)	Вычисляет b для линейной функции	НАКЛОН(A2:E2;A1:E1) → 0,564721
ОТРЕЗОК(массив Y; Массив V)	Вычисляет a для линейной функции	ОТРЕЗОК(A2:E2;A1:E1) → 1,15979
КОРРЕЛ(массив V; Массив Y)	Вычисляет коэффициент корреляции	КОРРЕЛ(A1:E1;A2:E2) → 0,9872
СТЮДРАСПОБР (вероятность, степень свободы)	Вычисляет стандартное значение критерия Стьюдента	СТЮДРАСПОБР(0,05;20) → 2,086

Контрольные вопросы

1. Какие типы связей между признаками называют функциональными и корреляционными. Приведите примеры.
2. Как классифицируют корреляцию по форме и направлению. В выполненной лабораторной работе укажите форму и направление найденных связей.
3. С помощью, каких показателей измеряют силу и направление связи между количественными или качественными признаками.
4. Перечислите способы, с помощью которых изображают корреляционную связь между двумя признаками.
5. Какие значения принимает коэффициент корреляции для прямых, обратных связей и при полном отсутствии связей между признаками. Чему равно его наибольшее и наименьшее возможное значение.
6. С помощью какого показателя оценивают достоверность коэффициента корреляции. В выполненной лабораторной работе укажите достоверные и недостоверные коэффициенты корреляции.
7. Дайте определение регрессии и функции.
8. В каких случаях регрессию относят к простой или множественной, прямолинейной или криволинейной.
9. Укажите способы отображения регрессии.

10. Что называют выравниванием эмпирических рядов. Приведите способы выравнивания.

11. МНК его назначение и свойства, использование для функций отличных от прямой. Какие приемы используют для нормализации вида функций.

Задание к лабораторной работе №3

1. На основе выданных значений, которые получены в результате физико-химического анализа проб, выполнить регрессионный и корреляционный анализ: найти коэффициенты линейной регрессии, коэффициенты корреляции.

2. Определить достоверности рассчитанных коэффициентов корреляции и сделать выводы.

3. Используя найденные коэффициенты линейной регрессии выполнить выравнивание эмпирических значений.

4. Для наихудшего варианта линейной регрессии построить график эмпирических и выровненных данных.

5. Подобрать и добавить к построенному графику наилучшую линию тренда – одну из предлагаемых типов сглаживающих функций - логарифмическую, степенную, экспоненциальную или полиномиальную. В качестве критерия отбора использовать автоматически рассчитываемый коэффициент корреляции для данной функции. При этом значения (Y) для найденной сглаживающей функции не должны принимать отрицательные (недопустимые) значения.

Порядок выполнения

1. Набрать исходные выданные значения в три пары рядом стоящих столбцов, например, столбцы А и В, D и E, G и H. Для последующих расчетов желательно оставить вначале пространство свободных ячеек (см. пример), поэтому исходные данные лучше подготовить в строках, например, ниже 11.

2. С помощью стандартных функций КОРРЕЛ(), СЧЁТ(), ОТРЕЗОК() и НАКЛОН() в столбцах С, F, I (для соответствующих данных) рассчитать коэффициенты корреляции, критерии их достоверности, коэффициенты линейной регрессии.

3. Используя стандартные функции СТЬЮДРАСПОБР() и СЧЁТ() определить стандартное значение критерия Стьюдента для заданных выборок.

4. Выделить диапазон ячеек наихудшего варианта линейной регрессии, например, G15:I32 (см. пример) и с помощью Мастера диаграмм построить точечную диаграмму.

5. С помощью щелчка правой кнопкой мышки на любой из точек исходных эмпирических данных в появившемся контекстном меню выбрать пункт Добавить линию тренда.

6. В появившемся окне диалога, во вкладке Тип выбрать форму сглаживающей функции. Во вкладке Параметры установить флажки: показывать уравнение на диаграмме и поместить на диаграмму величину достоверности аппроксимации (R^2).

7. С помощью щелчка правой кнопкой мышки на линии тренда, в последующем можно изменить тип сглаживающей функции.

8. Сохраните и распечатайте полученные результаты.

ПЕРЕЧЕНЬ РЕКОМЕНДОВАННОЙ ЛИТЕРАТУРЫ

1. Колесникова, И. И. Статистика : практикум. Учебное пособие / И. И. Колесникова, Г. В. Круглякова. — Минск : Вышэйшая школа, 2011. — 285 с. — ISBN 978-985-06-1892-4. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/20144.htm>
2. Матальцкий, М. А. Теория вероятностей, математическая статистика и случайные процессы : учебное пособие / М. А. Матальцкий, Г. А. Хацкевич. — Минск : Вышэйшая школа, 2012. — 720 с. — ISBN 978-985-06-2105-4. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/20289.html>
3. Гринь, А. Г. Вероятность и статистика : учебное пособие / А. Г. Гринь. — Омск : Омский государственный университет им. Ф.М. Достоевского, 2013. — 304 с. — ISBN 987-5-7779-1663-1. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/24879.html>
4. Шилова, З. В. Теория вероятностей и математическая статистика : учебное пособие / З. В. Шилова, О. И. Шилов. — Саратов : Ай Пи Ар Букс, 2015. — 158 с. — ISBN 978-5-906-17262-4. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/33863.html>
5. Теория вероятностей и математическая статистика : учебное пособие / Т. А. Гулай, А. Ф. Долгополова, Д. Б. Литвин, С. В. Мелешко. — Ставрополь : Ставропольский государственный аграрный университет, АГРУС, 2013. — 257 с. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/47360.html>

МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ
к проведению лабораторных работ по дисциплине
«Биометрия»

Составители:

Ганнова Юлия Николаевна – кандидат химических наук, доцент кафедры прикладной экологии и охраны окружающей среды ГОУВПО «ДОННТУ»;
Горбатко Сергей Витальевич – кандидат технических наук, доцент кафедры прикладной экологии и охраны окружающей среды ГОУВПО «ДОННТУ».

Ответственный за выпуск:

Шаповалов Валерий Васильевич – доктор химических наук, профессор, заведующий кафедрой «Прикладная экология и охрана окружающей среды» ГОУВПО «ДОННТУ».