

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Локтионова Оксана Геннадьевна

Должность: проректор по учебной работе

Дата подписания: 22.12.2021 15:52:49

Уникальный программный ключ:

0b817ca911e6668abb13a5d426d39e5f1c11ea0b73a943dfaa4851fa9560089

МИНИСТЕРСТВО НАУКИ И РОССИИ

**Федеральное государственное бюджетное образовательное
учреждение высшего образования
«Юго-Западный государственный университет»
(ЮЗГУ)**

Кафедра технологии материалов и транспорта

УТВЕРЖДАЮ
Проректор по учебной работе
О.Г. Локтионова
« 09 » 12 2021 г.



**КОМПЬЮТЕРНЫЕ ТЕХНОЛОГИИ В НАУКЕ И
ПРОИЗВОДСТВЕ**

Методические указания к выполнению практических и самостоятельных работ для
студентов направления подготовки 23.04.03 Эксплуатация транспортно-
технологических машин и комплексов очной и заочной форм обучения

Курск 2021

УДК 004. 09

Составители: А.Ю. Алтухов

Рецензент

Кандидат химических наук, доцент кафедры технологии материалов и транспорта Л.П. Кузнецова

Компьютерные технологии в науке и производстве: Методические указания к выполнению практических и самостоятельных работ для студентов направления подготовки 23.04.03 Эксплуатация транспортно-технологических машин и комплексов очной и заочной форм обучения / Юго-Зап. Гос. ун-т; сост.: А.Ю. Алтухов Курск, 2021. 66 с.: ил. 42, Библиогр.: 11.: с. 66.

Методические указания содержат весь необходимый материал для выполнения практических работ по дисциплине «Компьютерные технологии в науке и производстве».

Приведены содержание и объем практических работ, а также методика их проведения, основные методы решения задач, что помогает усвоить и глубже понять теоретические положения курса. В методические указания включены вопросы для повторения пройденного материала.

Предназначены для студентов направления подготовки 23.04.03 Эксплуатация транспортно-технологических машин и комплексов очной и заочной форм обучения.

Текст печатается в авторской редакции

Подписано в печать . Формат 60x84 1/16.

Усл.печ.л . Уч.-изд.л. . Тираж 100 экз. Заказ . Бесплатно.

Юго-Западный государственный университет.

305040, г. Курск, ул. 50 лет Октября, 94.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	4
Практическое занятие №1. ПЕРВИЧНАЯ ОБРАБОТКА ЭМПИРИЧЕСКИХ ДАННЫХ	5
Практическое занятие №2. РАСЧЕТ МЕР ЦЕНТРАЛЬНОЙ ТЕНДЕНЦИИ И МЕР ВАРИАЦИИ ЭМПИРИЧЕСКИХ ДАННЫХ	23
Практическое занятие №3. АНАЛИЗ ВЗАИМОСВЯЗИ ПРИЗНАКОВ	32
Практическое занятие №4. РАСЧЕТ КРИТЕРИЯ СТЬЮДЕНТА. ВЫЧИСЛЕНИЕ КОЭФФИЦИЕНТА ЛИНЕЙНОЙ КОРРЕЛЯЦИИ ПИРСОНА. ЗАПИСЬ УРАВНЕНИЯ ЛИНЕЙНОЙ РЕГРЕССИИ.....	42
Практическое занятие №5. УСТАНОВЛЕНИЕ ФУНКЦИОНАЛЬНОЙ ЗАВИСИМОСТИ ДЛЯ ИССЛЕДУЕМЫХ ПАРАМЕТРОВ	50
БИБЛИОГРАФИЧЕСКИЙ СПИСОК.....	56
Приложение А. ТЕОРЕТИЧЕСКИЕ СВЕДЕНИЯ ПО ОСОБЕННОСТЯМ РАБОТЫ В РЕДАКТОРЕ ЭЛЕКТРОННЫХ ТАБЛИЦ MS EXCEL	57
Приложение Б. ТАБЛИЦА РАСПРЕДЕЛЕНИЯ КРИТЕРИЯ ХИ- КВАДРАТ.....	63
Приложение В. ТАБЛИЦА РАСПРЕДЕЛЕНИЯ КРИТЕРИЯ СТЬЮДЕНТА	64
Приложение Г ПЕРЕЧЕНЬ ВОПРОСОВ ДЛЯ ПОДГОТОВКИ К ЗАЧЕТУ	65

ВВЕДЕНИЕ

Предлагаемые методические указания составлены в соответствии с рабочей программой дисциплины «Компьютерные технологии в науке и производстве».

При изучении курса в высших учебных заведениях большое значение имеет приобретение навыков самостоятельного выполнения различных задач, что является одним из критериев прочного усвоения курса.

Методические указания содержат разработки практических работ по дисциплине «Компьютерные технологии в науке и производстве». Указываются общие теоретические сведения, необходимые для выполнения практических работ, и порядок выполнения работ.

Предназначены для студентов направления 23.04.03 «Эксплуатация транспортно-технологических машин и комплексов».

Практическое занятие №1

ПЕРВИЧНАЯ ОБРАБОТКА ЭМПИРИЧЕСКИХ ДАННЫХ

Цель работы: научиться проводить первичный анализ эмпирических данных, используя возможности редактора электронных таблиц.

Общие сведения

Математическая статистика – это прикладная отрасль математики, основанная на теории вероятностей и предназначенная для систематизации и анализа эмпирических (опытных) данных.

Ключевыми понятиями математической статистики являются следующие:

1. Событие – это один из возможных исходов эксперимента.
2. Вероятность – это определённая количественная оценка объективной возможности появления некоторого события в заданной совокупности условий.
3. Статистическая совокупность (или выборка) – это вся система событий как исходов эксперимента, это ряд случайных значений измеренного признака $x_1, x_2, \dots, x_i, \dots, x_n$, варьирующихся в силу тех или иных статистических закономерностей. Выборка является частью генеральной совокупности.
4. Генеральная совокупность – это вся совокупность измерений эмпирических данных, свойство которых исследуется в научном эксперименте.
5. Варианта (x_i) – это каждое отдельное значение статистической совокупности, результат отдельного измерения.
6. Объем выборки (n) – общее число вариантов в выборке.
7. Частота (f_i) – число, показывающее сколько раз встречается в выборке каждая варианта x_i ($\sum f_i = n$).
8. Относительная частота (частость (ω_i)) – это доля каждой частоты в общем объеме выборки, то есть:

$$\omega_i = f_i / n. \quad (1)$$

Выборка должна обладать свойством качественной однородности, то есть все ее варианты должны представлять собой некие индивидуальные величины одного и того же качества или свойства.

Главная цель эмпирических наблюдений состоит в том, чтобы проверить гипотезы об интересующих исследователя научных явлениях и закономерностях. Однако перед тем как исследователь начинает проверять гипотезы, он пытается резюмировать или описать их по каждой из переменных. При резюмировании измерений одной переменной используется математическая статистика. В процессе первичной обработки эмпирических данных, с целью получения максимума сведений об исследуемом объекте или процессе, над всем материалом, имеющимся в выборке, необходимо провести следующую последовательность действий:

1. Упорядочивание – это расположение вариант выборки в последовательности, удобной для дальнейшего анализа и рассмотрения, например по возрастанию или убыванию.

2. Группировка – это разделение статистической совокупности на классы (группы), однородные по какому-либо признаку.

3. Табулирование – построение таблиц, в которых каждой variante x_i представлена в соответствие ее частота или относительная частота (частость).

4. Графическое представление статистического распределения. В математической статистике принято два вида:

а) полигон частот – это ломаная линия, соединяющая точки с координатами (x_i, f_i) или (x_i, ω_i) ;

б) гистограмма – это ряд смежных прямоугольников, построенных на одной прямой: площадь каждого из них пропорциональна частоте нахождения данной величины в интервале, на котором построен прямоугольник. Необходимо отметить, что для проведения упорядочивания и группировки эмпирических данных во многих отраслях науки, опирающихся на эмпирию, «конструируют» матрицу типа объект-признак.

Матрицей типа объект-признак называется совокупность данных следующего вида:

$$M = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1k} \\ x_{21} & x_{22} & \dots & x_{2k} \\ \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & x_{nk} \end{pmatrix},$$

где строки соответствуют n-объектам, а столбцы k-показателям.

Порядок выполнения работы

Рассмотрим первичную обработку эмпирических данных, полученных в ходе выполнения исследования по выявлению уровня зависимости между будущей профессией студента и степенью удовлетворенности учебой. В данном исследовании рассматривались три признака (три переменных): будущая профессия студента, степень удовлетворенности учебой и продолжительность затрат времени на учебу. Будем приводить данные модельного характера, то есть в реальном исследовании они не были получены. В исследовании принимали участие 100 студентов разных специальностей. Каждому студенту была предоставлена отдельная анкета для ответов на вопросы исследователя, то есть в матрице каждому студенту соответствует анкета заданного номера.

1. Упорядочивание и группировка

Постройте матрицу типа объект-признак (рисунок 1). Названия столбцов матрицы соответствуют рассматриваемым признакам (будущая профессия студента, степень удовлетворенности учебой и продолжительность затрат времени на учебу), а строк – номерам анкет опрашиваемых студентов (объектам). Первая переменная (признак) измерялась по номинальной шкале и кодировалась исследователем следующим образом: политолог – 1, социолог – 2, культуролог – 3, филолог – 4, психолог – 5, историк – 6, то есть первая переменная (признак) может принимать числовые значения от 1 до 6. Вторая переменная (признак) измерялась по шкале порядка и может принимать значения от 1 до 5, отражающие степень удовлетворенности учебой каждого студента. Третья переменная (признак) измерялась по интервальной шкале и кодировалась исследователем следующим образом: если студент тратит на учебу от 0 часов до 1 часа в день, то этому интервалу соответствует 1; от 1 часа до 2,5 часа – 2; от 2,5 до 4 часов – 3; от 4 до 7 часов – 4; от 7 до 8 часов – 5; от 8 до 9 часов – 6, то есть третья переменная (признак) может принимать числовые значения от 1 до 6.

Для того чтобы, используя редактор электронных таблиц, построить матрицу типа объект-признак выполните следующие действия.

Рисунок 1 – Матрица типа объект-признак

- Введите в ячейку A1 «№ анкеты (объекты) / № вопроса (показатели)». Увеличьте ширину столбца А: наведите указатель мыши на разделитель между столбцами А и В в строке с именами ячеек. Указатель станет двунаправленной стрелкой. Нажмите левую кнопку мыши и, удерживая ее, переместите границу вправо. Введите в ячейку A2 номер первой анкеты «№1».

- Введите в ячейку А3 номер второй анкеты «№2». Используйте функцию автозаполнения для размножения номеров анкет от ячейки А2 до ячейки А101: выделите ячейки А2 и А3, наведите указатель мыши на маркер заполнения – квадратик в нижнем правом углу рамки (см. рис. 40 приложения А), выделяющий ячейки А2 и А3. Нажмите и, удерживая левую клавишу мыши, обведите рамкой ячейки А2:А101 и после этого отпустите клавишу мыши.

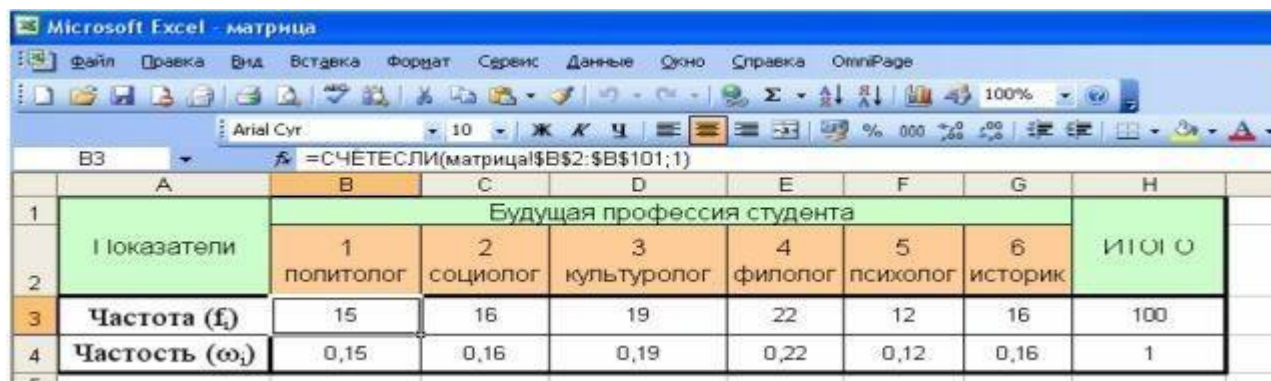
- Введите в ячейку B1 название первой переменной (признака) – «1 Будущая профессия студента».
- Введите в ячейку C1 название второй переменной (признака) – «2 Степень удовлетворенности учебой».
- Введите в ячейку D1 название третьей переменной (признака) – «3 Продолжительность затрат времени на учебу».
- Введите в ячейки B2:B101 значения первой переменной (признака) для каждого объекта (в данном исследовании анкеты студента), то есть значения от 1 до 6, которыми кодируются соответствующие будущие профессии студентов (см. выше). *Значения придумайте самостоятельно.*
- Введите в ячейки C2:C101 значения второй переменной (признака) для каждого объекта (в данном исследовании анкеты студента), то есть значения от 1 до 5, отражающие соответствующую степень удовлетворенности студента учебой. *Значения придумайте самостоятельно.*
- Введите в ячейки D2:D101 значения третьей переменной (признака) для каждого объекта (в данном исследовании анкеты студента), то есть значения от 1 до 6 – номера временных интервалов. *Значения придумайте самостоятельно.*
- Выделите ячейки B1:D1. Выполните команду *Формат/Ячейки*, выберите вкладку *Граница*, выберите тип линии – *жирная линия*, нажмите кнопку *Внешние*, затем нажмите ОК.
- Выделите ячейки B2:D101. Выполните команду *Формат/Ячейки*, выберите вкладку *Граница*, выберите тип линии – *жирная линия*, нажмите кнопку *Внешние*, затем выберите тип линии – *тонкая линия*, нажмите кнопку *Внутренние*, нажмите ОК.
- По аналогии с вышестоящими пунктами задайте соответствующие границы для группы ячеек A2:A101 и для ячейки A1, как показано на рисунке 1.
- Выделите ячейки B1:D1. Выполните команду *Формат/Ячейки*, выберите вкладку *Вид*, выберите *светло-зеленый*, нажмите ОК.
- По аналогии с вышестоящим пунктом задайте соответствующую заливку для группы ячеек A2:A101 и для ячейки A1, как показано на рисунке 1.
- Для ячеек B1:D1 задайте перенос по словам, выравнивание по центру и размер шрифта 14. Для этого выделите ячейки B1:D1,

выполните команду *Формат/Ячейки*, выберите вкладку *Выравнивание*, установите флажок *переносить по словам* и выберите выравнивание *по центру*. Затем выберите вкладку *Шрифт* и размер шрифта 14. Нажмите ОК.

• Аналогично задайте выравнивание по центру и размер шрифта 14 для группы ячеек A2:A101, выравнивание по центру и размер шрифта 12 для группы ячеек B2:D101. Измените название листа «Лист1» на «матрица». Для этого щелкните по названию листа правой кнопкой мыши, в появившемся контекстном меню выберите пункт *переименовать*, наберите имя листа «матрица». Нажмите клавишу Enter.

2. Табулирование

На основе эмпирических данных, записанных в матрице типа объект-признак, постройте таблицы распределения для каждого из трех признаков (рисунки 2, 6, 7).



	показатели	будущая профессия студента						итоги
		1 политолог	2 социолог	3 культуролог	4 филолог	5 психолог	6 историк	
3	Частота (f_i)	15	16	19	22	12	16	100
4	Частость (ω_i)	0,15	0,16	0,19	0,22	0,12	0,16	1

Рисунок 2 – Таблица распределения студентов по их будущей профессии

В таблицах распределения с помощью редактора электронных таблиц должны быть рассчитаны следующие показатели: частота, частость. Для построения таблиц распределения и автоматизации расчетов частоты и частости в MS Excel выполните следующие действия.

• На *Листе2* книги Excel в которой построена матрица типа объект-признак, постройте и оформите таблицу распределения студентов по их будущей профессии, как показано на рисунке 2, при этом ячейки для значений частоты, частости и итоговых значений

пока оставьте пустыми. *Форматирование таблицы* (заливка ячеек, расстановка границ и т.д.) осуществляется по аналогии с форматированием матрицы типа объект-признак (см. пункт 1).

- Для того чтобы оформить название таблицы «по центру», выделите ячейки B1:G1 и нажмите на панели инструментов кнопку



– *Объединить и поместить в центре*. Аналогично объединяются ячейки A1:A2 и H1:H2.

- Вычислите количество политологов из всей совокупности опрашиваемых студентов, то есть вычислите частоту появления политологов в совокупности опрашиваемых студентов. Для этого проделайте следующие действия:

- 1) выделите ячейку B3;

- 2) щелкните левой кнопкой мыши по команде f_x – *Вставка функции* на стандартной панели инструментов. Появится диалоговое окно *Мастер функций – шаг 1 из 2* (рисунок 3);

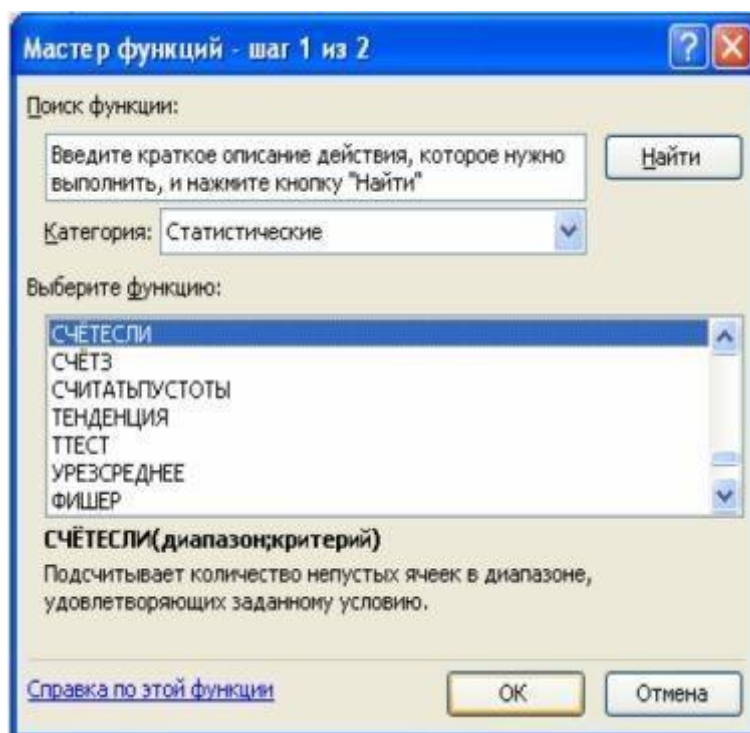


Рисунок 3 – Мастер функций MS Excel

- 3) в диалоговом окне выберите категорию *Статистические*. В выпадающем списке выберите название функции *СЧЁТЕСЛИ*. Нажмите ОК. Появится окно *Аргументы функции* (рисунок 4);

4) в поле *Диапазон* введите: *матрица!B2:B101*, в поле *Критерии* введите *1* (рисунок 4). Нажмите ОК;

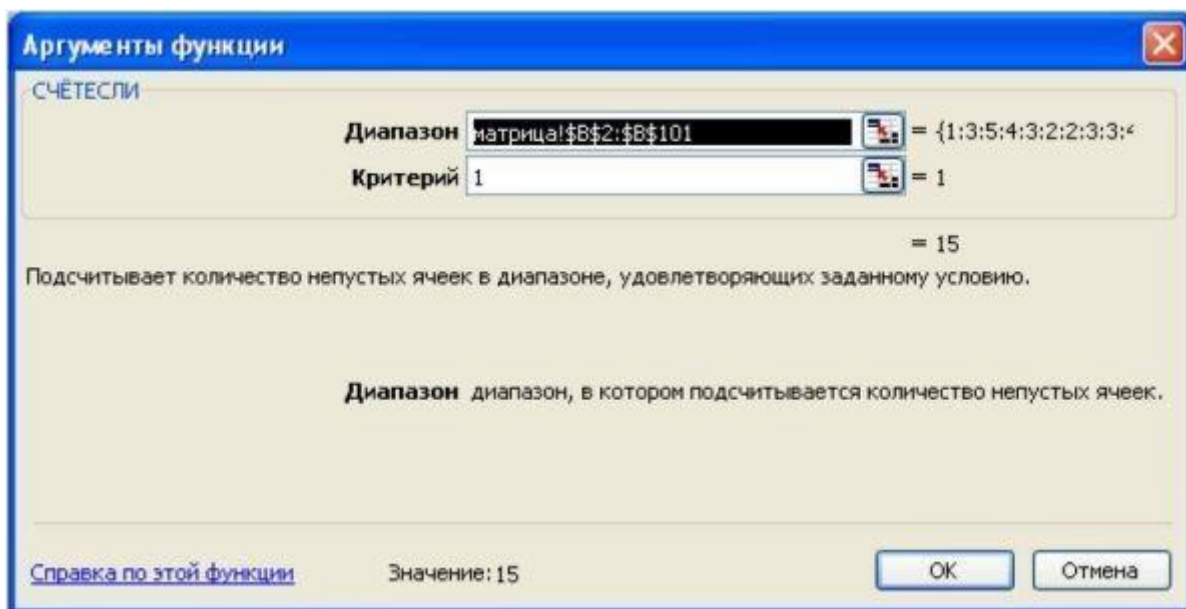


Рисунок 4 – Аргументы функции СЧЁТЕСЛИ

5) аналогичным образом рассчитайте частоты появления студентов всех остальных специальностей в общей совокупности опрошенных студентов (2 – социолог. 3 – культуролог. 4 – филолог. 5 – психолог, 6 – историк). *Значения частот, полученные в вашей таблице распределения, могут отличаться от значений частот, приведенных в таблице рисунка 2, так как матрица типа объект-признак заполнялась вами самостоятельно, то есть значения в вашей матрице отличаются от значений в матрице, приводимой в методических указаниях.*

• Автоматизируйте расчет общего количества студентов (анкет), принимавших участие в исследовании. Для этого:

- 1) выделите ячейку Н3;
- 2) щелкните левой кнопкой мыши по команде *fx – Вставка функции* на стандартной панели инструментов. Выйдет диалоговое окно *Мастер функций – шаг 1 из 2* (см. рисунок 3);
- 3) в диалоговом окне выберите категорию *Статистические*. В выпадающем списке выберите название функции *СЧЁТ*. Нажмите ОК. Появится окно *Аргументы функции* (см. рисунок 5);

4) в поле *Значение1* введите: *матрица!B2:B101* (рисунок 5). Нажмите ОК.

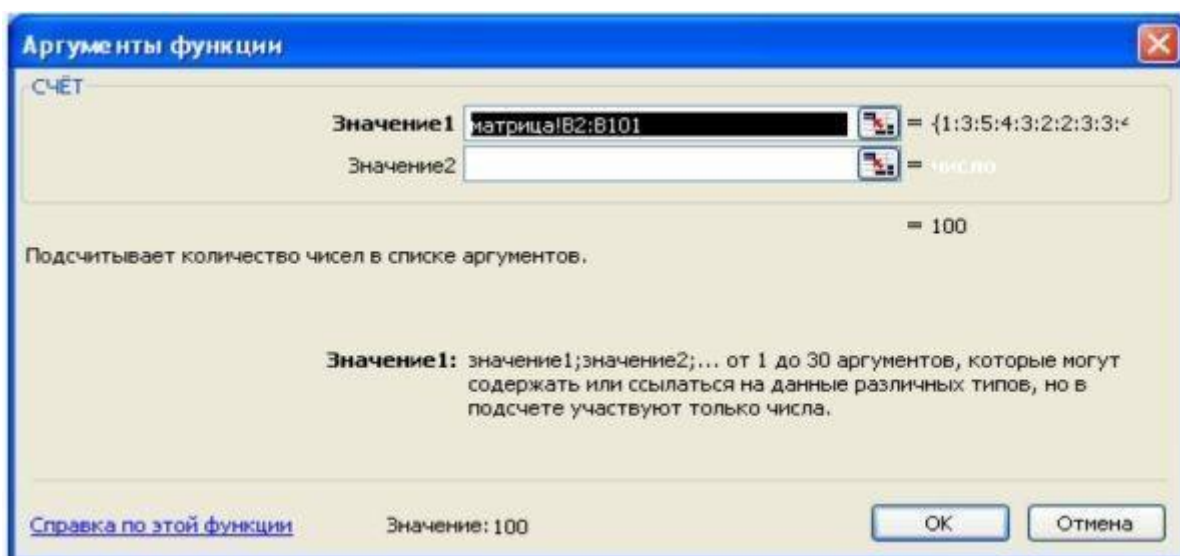


Рисунок 5 – Аргументы функции СЧЁТ

- Рассчитайте частоту совокупности студентов каждой профессиональной группы в общей совокупности студентов, принимавших участие в исследовании. Для этого:

- 1) выделите ячейку B4;
- 2) введите в ячейку B4 следующую формулу: $=СЧЁТЕСЛИ(матрица!$B$2:$B$101;l)/СЧЁТ(матрица!$B$2:$B$101)$. Нажмите Enter. Значок \$ используется для задачи абсолютной ссылки (адресации) на ячейку (см. приложение А):

- 3) используя функцию автозаполнения скопируйте данную формулу на все ячейки диапазона C4:G4.

- Рассчитайте итоговое значение частоты. Для этого:

- 1) выделите ячейку H4;
- 2) введите формулу для подсчета суммы с помощью автосуммы.



Для этого нажмите кнопку. При этом пунктирной линией выделите ячейки B4:G4. Затем нажмите клавишу Enter.

- Измените название листа «Лист2» на «Будущая профессия студентов».

На *Листе3* данной книги Excel постройте и оформите таблицу распределения студентов по степени их удовлетворенности учебой (рисунок 6).

Microsoft Excel - матрица

Файл Правка Вид Вставка Формат Сервис Данные Окно Справка OmniPage

В3 $\text{=СЧЁТЕСЛИ(матрица!\$C\$2:\$C\$101;1)}$

	A	B	C	D	E	F	G	H
1	Показатели	Степень удовлетворенности учебой					ИТОГО	
2		1	2	3	4	5		
3	Частота (f_i)	13	32	23	17	15	100	
4	Частость (ω_i)	0,13	0,32	0,23	0,17	0,15	1	
5								
6								

Рисунок 6 – Таблица распределения студентов по степени их удовлетворенности учебой

- По аналогии с расчетом частот, частостей и итоговых значений для таблицы распределения студентов по их будущей профессии рассчитайте данные показатели в таблице распределения студентов по степени их удовлетворенности учебой.

- Измените название листа «Лист3» на «Удовлетворенность учебой».

- Добавьте в книгу Excel ещё один лист. Для этого выполните команду *Вставка/Лист*.

- Созданный лист разместите в конец книги Excel. Для этого щелчком правой кнопки мыши на вкладке листа вызовите контекстное меню листа, затем выберите команду *Переместить/Скопировать*. На экране появится окно *Переместить или скопировать*, в котором в списке выберите (*переместить в конец*).

- Измените название листа на название «Затраты времени на учебу».

- На листе «Затраты времени на учебу» данной книги Excel постройте и оформите таблицу распределения студентов по времени, затрачиваемому ими на учебу (рисунок 7).

Microsoft Excel - матрица

Файл Правка Вид Вставка Формат Сервис Данные Окно Справка ОффPage

Введите вопрос

100%

Аннотация

В4 =СЧЕТЕСЛИ(матрица!\$D\$2:\$D\$10;1)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Показатели	Время, затрачиваемое на учебу												ИТОГО
2		1	2	3	4	5	6							
3		0 час.	1 час.	1 час.	2,5 час.	2,5	4 час.	4 час.	7 час.	7 час.	8 час.	8 час.	9 час.	
4	Частота (f_i)	12	21	25	22	13	7	100						
5	Частость (ω_i)	0,12	0,21	0,25	0,22	0,13	0,07	1						

Рисунок 7 – Таблица распределения студентов по времени, затрачиваемому ими на учебу

• Для задания формата ячеек В3:М3, задающего подпись «час», необходимо использовать возможности пользовательского форматирования. Для этого:

1) выделите ячейки В3:М3. Выберите пункт меню *Формат/Ячейки*, выберите вкладку *Число*. Вид окна *Формат ячейки* представлен на рисунке 8;

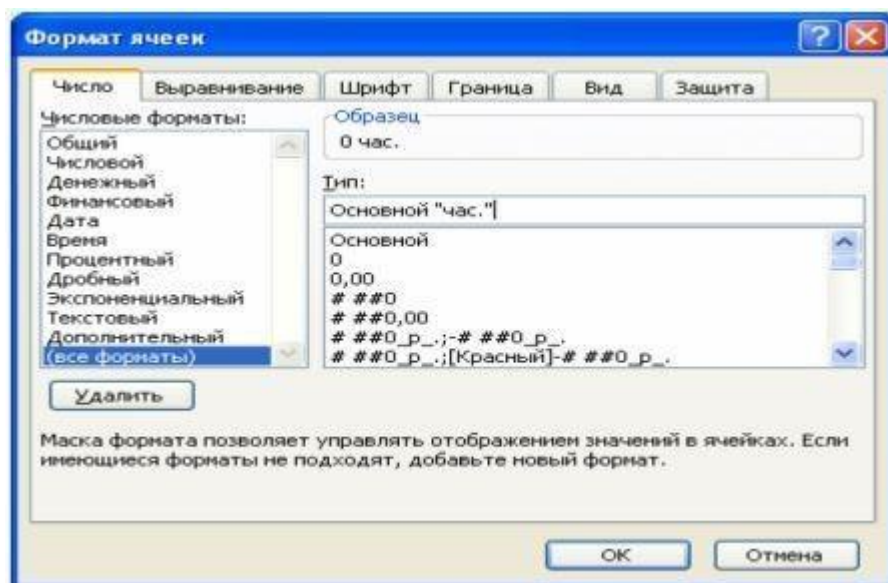


Рисунок 8 – Окно «Формат ячейки»

2) в списке *Числовые форматы* данного окна выберите строку *(все форматы)*;

3) в списке под строкой *Тип* щелчком левой кнопки мыши выберите строку *Основной*, которая помещается в поле ввода нового формата (см. рисунок 8):

4) щелчком левой кнопки мыши в поле ввода нового формата поместите текстовый курсор;

5) после слова *Основной* наберите с клавиатуры подпись "*час.*" (кавычки обязательны!). В поле **Образец** отображается внешний вид ячейки. Вид окна должен соответствовать рисунку 8. Нажмите ОК.

- По аналогии с расчетом частот, частостей и итоговых значений для таблицы распределения студентов по их будущей профессии рассчитайте данные показатели в таблице распределения студентов по времени, затрачиваемому ими на учебу.

3. Графическое представление статистического распределения

Постройте полигоны частот и гистограммы распределения студентов по их будущей профессии, по степени удовлетворенности учебой и по времени, затрачиваемому ими на учебу. Для этого выполните следующие действия.

- Откройте таблицу распределения студентов по их будущей профессии. Для этого перейдите на лист «Будущая профессия студентов», щелкнув по нему левой кнопкой мыши.

- Постройте полигон частот распределения студентов по их будущей профессии по частоте. Для этого:

- 1) выберите команду *Вставка/Диаграмма*. Появится диалоговое окно *Мастер функций (шаг 1 из 4): тип диаграммы* (рисунок 9);

- 2) выберите *Тип: График, Вид: График с маркерами, помечающими точками данные* (рисунок 9). Нажмите Далее;

- 3) на вкладке *Диапазон данных* в поле *Диапазон* введите B3:G3, то есть частоты распределения студентов по их будущей профессии (рисунок 10);

- 4) перейдите на вкладку *Ряд*. В строке *Подписи оси X* введите диапазон ячеек B2:G2, то есть названия будущих профессий студентов. В строке *Имя* наберите фразу «*частота распределения*» (рисунок 11). Нажмите Далее;

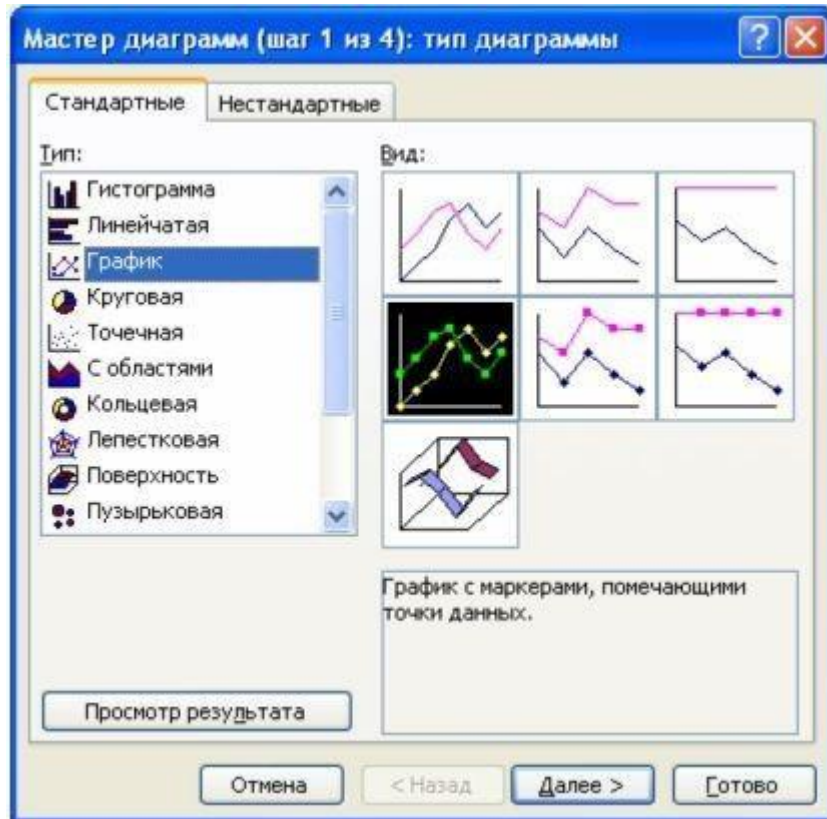


Рисунок 9 – Мастер диаграмм (шаг 1 из 4): тип диаграммы

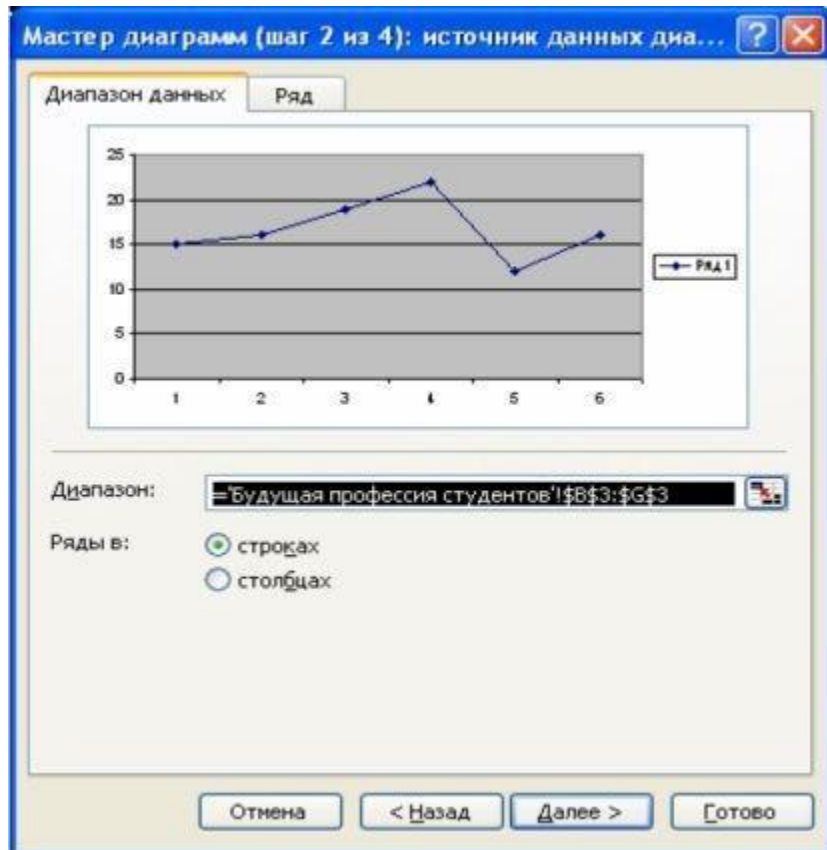


Рисунок 10 – Вкладка Диапазон данных Мастера диаграмм MS Excel

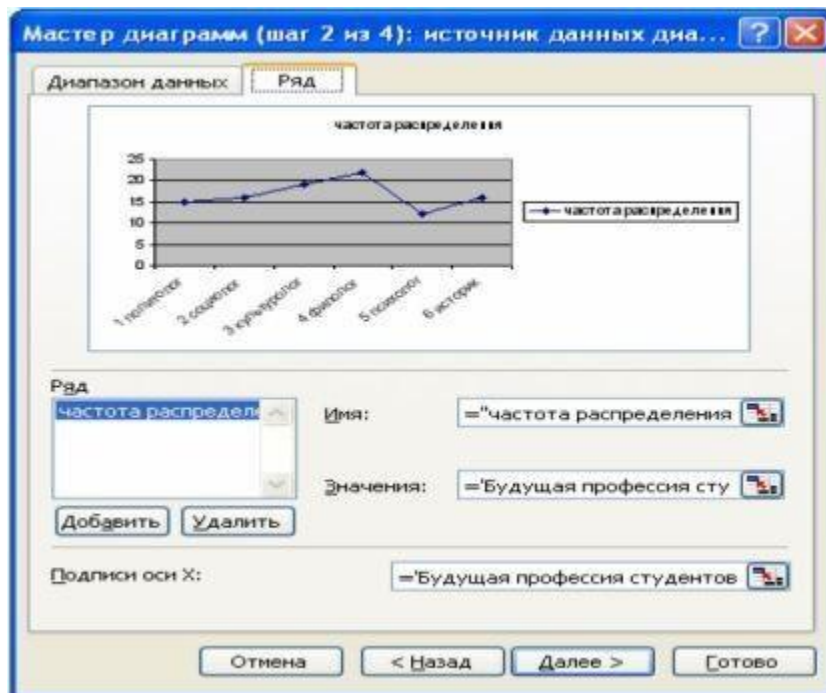


Рисунок 11 – Вкладка Ряд Мастера диаграмм MS Excel

5) на вкладке *Заголовки* в строке *Название диаграммы* наберите «Полигон частот распределения студентов по их будущей профессии». В строке *Ось X (категорий)* наберите «Будущая профессия студента». В строке *Ось Y (значений)* наберите «Частота» (рисунок 12). Нажмите *Далее*;

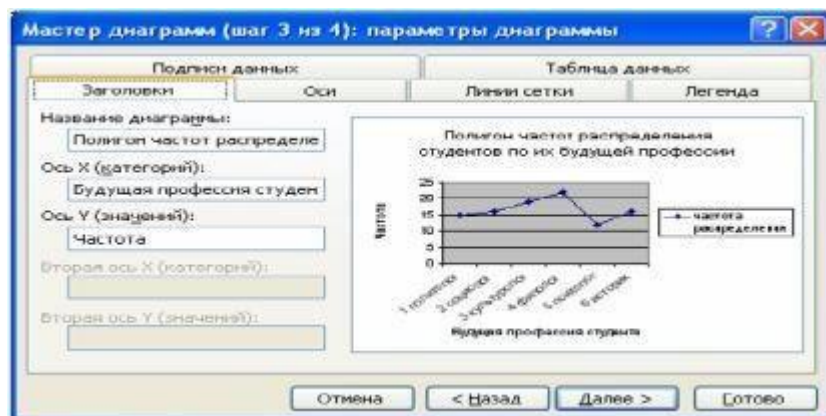


Рисунок 12 – Мастер диаграмм (шаг 3 из 4): параметры диаграммы

6) в появившемся диалоговом окне (рисунок 13) выберите размещение диаграммы на имеющемся листе. Нажмите *Готово*:

7) в итоге получите полигон частот распределения студентов по их будущей профессии (рисунок 14).

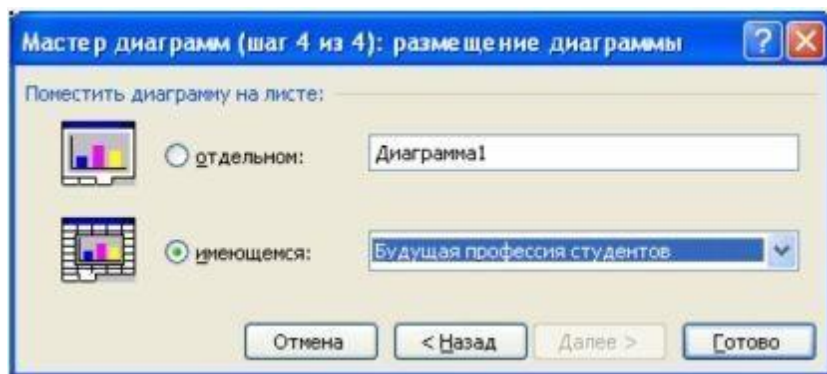


Рисунок 13 – Мастер диаграмм (шаг 4 из 4): размещение диаграммы



Рисунок 14 – Полигон частот распределения студентов по их будущей профессии

- Постройте гистограмму распределения студентов по их будущей профессии по частотам. Алгоритм действий по построению гистограммы будет аналогичен алгоритму действий по построению полигона частот, с той лишь разницей, что нужно выбрать *Тип: гистограмма, Вид: Обычная* (рисунок 15), а на вкладке **Диапазон** данных в поле *Диапазон* ввести B4:G4, то есть частоты распределения студентов по их будущей профессии. В итоге получите гистограмму распределения студентов по их будущей профессии по частотам (рисунок 16).

- Аналогичным образом постройте полигон частот распределения студентов по степени их удовлетворенности учебой по частоте и гистограмму распределения студентов по степени их удовлетворенности учебой по частотам.

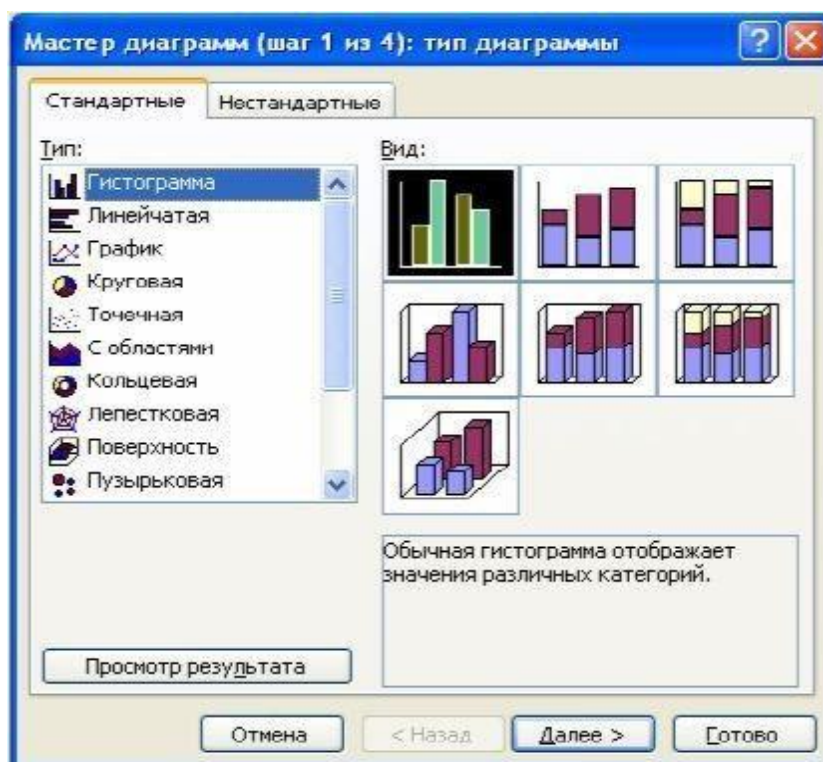


Рисунок 15 – Мастер диаграмм (шаг 1 из 4): тип диаграммы

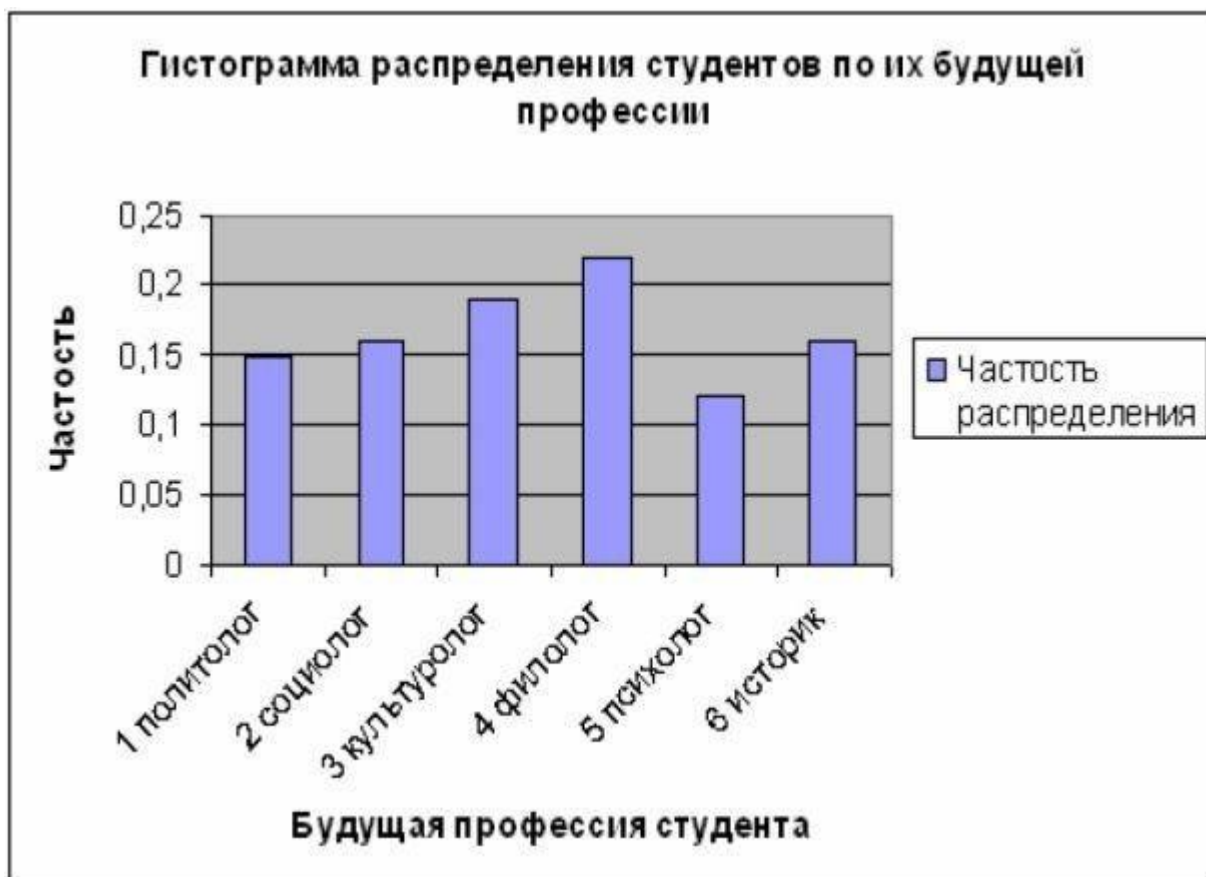


Рисунок 16 – Гистограмма распределения студентов по их будущей профессии

- Постройте полигон частот и гистограмму распределения студентов по времени, затрачиваемому ими на учебу. Так как данный признак (переменная) измерялся по порядковой шкале, следовательно, прежде чем приступить к графическому представлению эмпирических данных, необходимо рассчитать плотности распределения. Плотность распределения рассчитывается по следующей формуле

$$p_i = \frac{f_i}{n_i - m_i}, \quad (2)$$

где f_i – частота распределения i -го интервала;

n_i – левая граница i -го интервала;

m_i – правая граница i -го интервала.

Для расчета плотности распределения студентов в первом интервале (от 0 часов до 1 часа) выделите ячейку A9 и введите формулу $=B4/(C3-B3)$. Аналогично рассчитайте плотности распределения студентов по остальным интервалам (рисунок 17). В качестве диапазона данных для построения полигона частот и гистограммы распределения задайте плотности распределения.

- Сохраните документ под именем «*Матрица*».

Полигон частот распределения студентов по времени, затрачиваемому на учебу

Показатели	1	2	3	4	5	6
Частота (f_i)	12	21	25	22	13	7
Частость (ω_i)	0,12	0,21	0,25	0,22	0,13	0,07

Плотности распределения

12
14
17
7
13
7

Полигон частот распределения студентов по времени, затрачиваемому на учебу

График показывает плотность распределения (Y-ось) по номерам временных интервалов (X-ось). Легенда: Ряд 1.

Практическое занятие №2

РАСЧЕТ МЕР ЦЕНТРАЛЬНОЙ ТЕНДЕНЦИИ И МЕР ВАРИАЦИИ ЭМПИРИЧЕСКИХ ДАННЫХ

Цель работы: освоить методы расчета мер центральной тенденции и мер вариации эмпирических данных с помощью редактора электронных таблиц.

Общие сведения

Меры центральной тенденции – это такие значения признака, вокруг которых группируются отдельные значения элементов, полученных в ходе проведения научного эксперимента.

Рассмотрим следующие меры центральной тенденции: среднее арифметическое, средневзвешенное, мода и медиана.

Среднее арифметическое (M) – это результат деления суммы измерений на их количество. Общая формула для вычисления среднего арифметического выглядит следующим образом:

$$M = \frac{\sum_{i=1}^n X_i}{n}, \quad (3)$$

где X_i – значения отдельных элементов выборки;
 n – объем выборки.

Средневзвешенное применяется для вычисления среднего значения в случае использования интервальных шкал. Общая формула для вычисления средневзвешенного выглядит следующим образом:

$$M = \frac{\sum_{i=1}^k f_i x_i}{n}, \quad (4)$$

где x_i – середина i -го интервала;
 f_i – частота интервала;
 n – объем выборки;
 k – число интервалов.

Мода – это такое значение признака, которое встречается наиболее часто (имеет наибольшую частоту).

В случае когда значения получены с помощью интервальных шкал, говорят о модальном интервале. Точное значение моды в случае интервальных рядов вычисляется по формуле

$$M_0 = x_0 + h \cdot \frac{n_2 - n_1}{(n_2 - n_1) + (n_2 - n_3)}, \quad (5)$$

где x_0 – начальное значение модального интервала;

h – длина модального интервала;

n_1 – частота интервала перед модальным;

n_2 – частота модального интервала;

n_3 – частота интервала после модального.

Медиана – это величина измерения, равноудаленная от начала и конца измерений, выстроенных по возрастанию.

Для интервальных шкал медиана вычисляется по формуле

$$M_e = x_n + \frac{\frac{n}{2} - S_{M_e-1}}{n_{M_e}} \cdot h, \quad (6)$$

где x_n – начало медианного интервала;

n – объем выборки;

S_{M_e-1} – сумма частот интервалов, предшествующих модальному;

n_{M_e} – частота медианного интервала;

h – ширина медианного интервала.

Меры вариации (меры рассеяния) оценивают степень изменчивости вариант и являются одной из характеристик группировки.

Дисперсия – сумма квадратов отклонений каждого значения в выборке от среднего, разделенная на объем выборки:

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - M)^2}{n}, \quad (7)$$

где x_i – значение i -го элемента в выборке;

M – среднее арифметическое (или средневзвешенное),

n – объем выборки.

Среднеквадратическое отклонение (стандартное отклонение) – корень квадратный из дисперсии:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - M)^2}{n}}. \quad (7)$$

Порядок выполнения работы

По эмпирическим данным, записанным в матрице типа объект-признак и таблицах распределения работы №1, рассчитайте меры центральной тенденции и меры вариации для каждого из признаков (переменных).

1. Расчет модального значения

- Откройте документ «Матрица».
- Вычислите модальное значение для распределения студентов по их будущей профессии.
- Выберите лист «Будущая профессия студентов».
- Выделите ячейку В7 и вызовите *Мастер функций* MS Excel с помощью команды f_x – *Вставка функции*.
- В диалоговом окне выберите категорию *Статистические*. В выпадающем списке выберите название функции *МОДА*. Нажмите ОК. Появится окно *Аргументы функции* (рисунок 18).

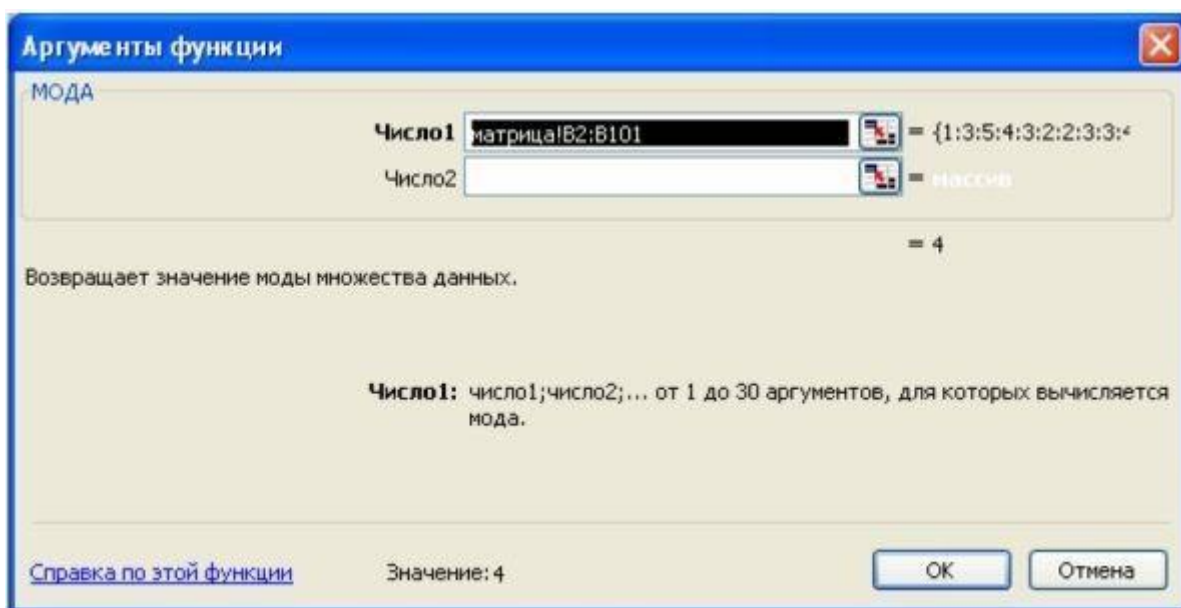


Рисунок 18 – Аргументы функции МОДА

• В строке *Число1* введите: *матрица!B2:B101*. Нажмите ОК. Поскольку нами в работе №1 были рассчитаны частоты и частоты распределения студентов по их будущей профессии, в данном случае возможно было бы определить модальное значение по одному из этих двух показателей. Однако не всегда научные исследования предполагают построение таблиц распределения и, следовательно, расчет частот. При этом расчет мер центральной тенденции необходим для статистического анализа результатов научного эксперимента.

• Аналогично рассчитайте модальное значение для второго признака (переменной) – степень удовлетворенности студентов учебой.

• Рассчитайте модальное значение для третьего признака (переменной) – время, затрачиваемое студентами на учебу. Поскольку значения данной переменной (признака) измерялись по интервальной шкале, то использовать функцию МОДА мастера функций редактора электронных таблиц для данного признака невозможно. Для этого необходимо воспользоваться формулой (5). Для эмпирических данных, приведенных в методических указаниях, данная формула в редакторе электронных таблиц запишется следующим образом: $=F3+(G3-F3)*((F4-D4)/((F4-D4)+(F4-H4)))$ (рисунок 19).

	A	B	C	D	E	F	G	H	I	J	K	L	M
1		Время, затрачиваемое на учебу											
2	Показатели	1	2	3	4	5	6						
3		0 час.	1 час.	1 час.	2,5 час.	2,5	4 час.	4 час.	7 час.	7 час.	8 час.	8 час.	9 час.
4	Частота (fi)	12	21	25	22	13	7						
5	Частость (oi)	0,12	0,21	0,25	0,22	0,13	0,07						
6	плотности распределения	12	14	17	7	13	7						
7													
8													
9													
10	МОДА	3,36											
11	МЕДИАНА	3,52											
12	СРЕДНЕЕВЗВЕШЕННОЕ	4,02											

Рисунок 19 – Пример расчета МОДЫ в MS Excel для переменных, значения которых измеряют по интервальной шкале

2. Расчет медианы

- Вычислите медиану для распределения студентов по степени их удовлетворенности учебой.
- Выберите лист «Удовлетворенность учебой».
- Выделите ячейку B9 и вызовите *Мастер функций* MS Excel с помощью команды f_x – *Вставка функции*.
- В диалоговом окне выберите категорию *Статистические*. В выпадающем списке выберите название функции *МЕДИАНА*. Нажмите ОК. Появится окно *Аргументы функции* (рисунок 20).

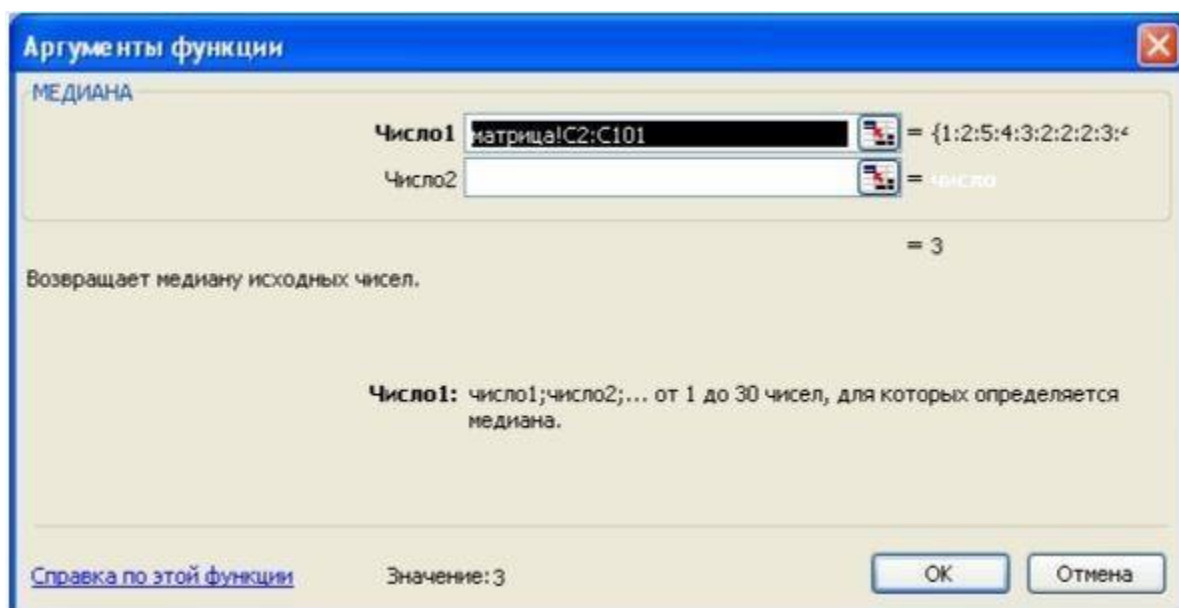


Рисунок 20 – Аргументы функции МЕДИАНА

- В строке *Число1* введите: *матрица!C2:C101*. Нажмите ОК.
- Рассчитайте медиану для третьего признака (переменной) – продолжительность затрат времени на учебу. Поскольку значения данной переменной (признака) измерялись по интервальной шкале, использовать функцию *МЕДИАНА* мастера функций редактора электронных таблиц для данного признака невозможно. Для этого необходимо воспользоваться формулой (6). Для эмпирических данных, приведенных в методических указаниях, данная формула в редакторе электронных таблиц запишется следующим образом:

$$=F3+((N4/2-SUMM(B4:E4))/F4)*(G3-3)$$
 (рисунок 21).

Показатели	1	2	3	4	5	6
Частота (f_i)	12	21	25	22	13	7
Частость (ω_i)	0,12	0,21	0,25	0,22	0,13	0,07
плотности распределения	12	14	17	7	13	7
МОДА	3,36					
МЕДИАНА	3,52					
СРЕДНЕЕ ВЗВЕШЕННОЕ	4,02					

Рисунок 21 – Пример расчета МЕДИАНЫ в MS Excel для переменных, значения которых измеряются по интервальной шкале

3. Расчет значений среднего арифметического и средневзвешенного

- Вычислите среднее арифметическое значение распределения студентов по степени их удовлетворенности учебой.
- Выберите лист «Удовлетворенность учебой».
- Выделите ячейку B10 и вызовите *Мастер функций* MS Excel с помощью команды f_x – *Вставка функции*.
- В диалоговом окне выберите категорию *Статистические*. В выпадающем списке выберите название функции *СРЗНАЧ*. Нажмите ОК. Появится окно *Аргументы функции* (рисунок 22).
- В строке *Число1* введите: *матрица!C2:C101*. Нажмите ОК.
- Рассчитайте средневзвешенное для третьего признака (переменной) – продолжительность затрат времени на учебу. Поскольку значения данной переменной (признака) измерялись по интервальной шкале, использовать функцию *СРЗНАЧ* мастера функций редактора электронных таблиц для данного признака невозможно. Для этого необходимо воспользоваться формулой (4). Для эмпирических данных, приведенных в методических указаниях, данная формула в редакторе электронных таблиц запишется следующим образом: $=СУММ(B4*((C3-B3)/2+B3);D4*((E3-D3)/2+D3);F4*((G3-F3)/2+F3);H4*((I3-H3)/2+H3);J4*((K3-J3)/2+J3);L4*((M3-L3)/2+L3))N4$ (рисунок 23).

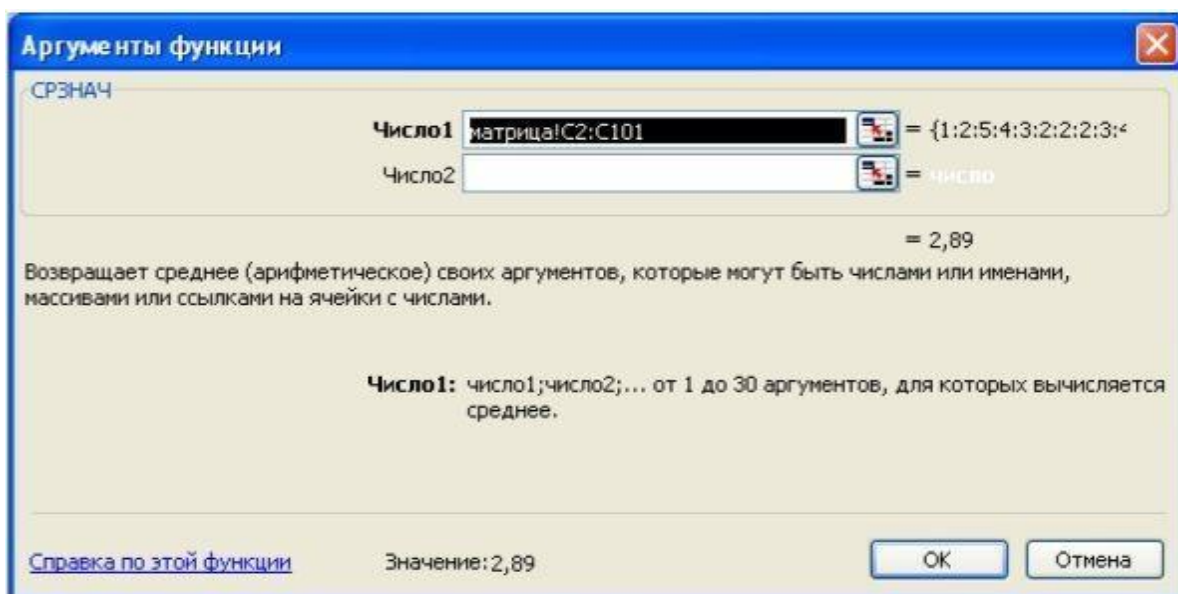


Рисунок 22 – Аргументы функции СРЗНАЧ

Рисунок 23 – Пример расчета средневзвешенного в MS Excel

- Вычислите дисперсию распределения студентов по степени их удовлетворенности учебой.
- Выберите лист «Удовлетворенность учебой».
- Выделите ячейку B11 и вызовите *Мастер функций* MS Excel с помощью команды f_x – *Вставка функции*.
- В диалоговом окне выберите категорию *Статистические*. В выпадающем списке выберите название функции *ДИСПР*. Нажмите ОК. Появится окно *Аргументы функции* (рисунок 24).

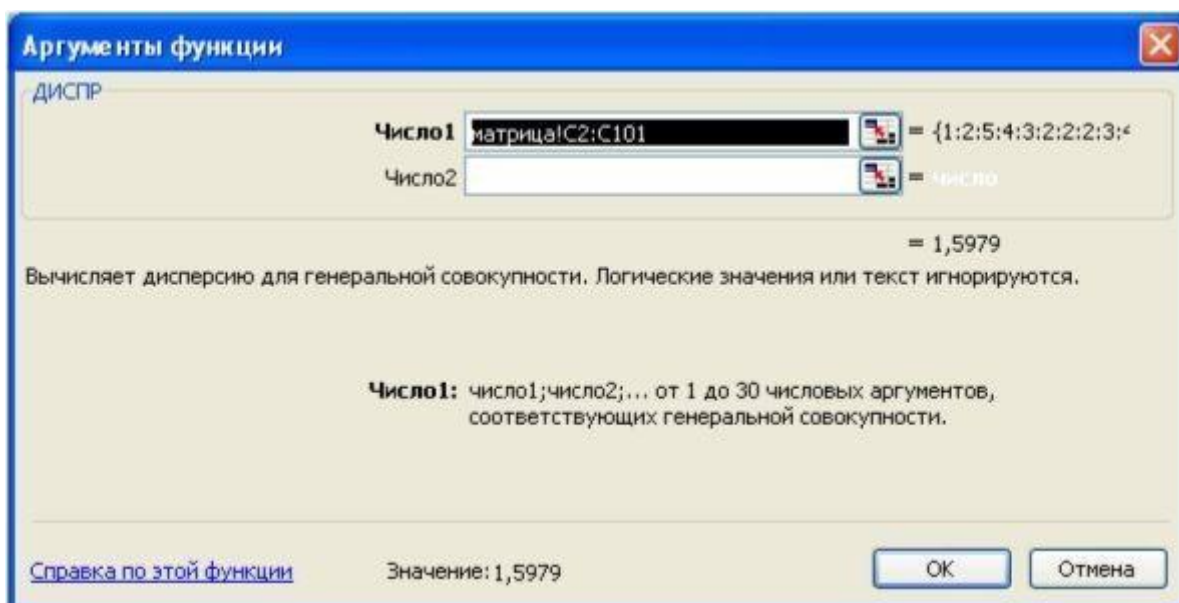


Рисунок 24 – Аргументы функции ДИСПР

- В строке *Число1* введите: *матрица!C2:C101*. Нажмите ОК.
- Вычислите среднеквадратическое (стандартное) отклонение распределения студентов по степени их удовлетворенности учебой.
- Выделите ячейку B12 и вызовите *Мастер функций* MS Excel с помощью команды f_x – *Вставка функции*.
- В диалоговом окне выберите категорию *Статистические*. В выпадающем списке выберите название функции *СТАНДОТКЛОНП*. Нажмите ОК. Появится окно *Аргументы функции* (рисунок 25).

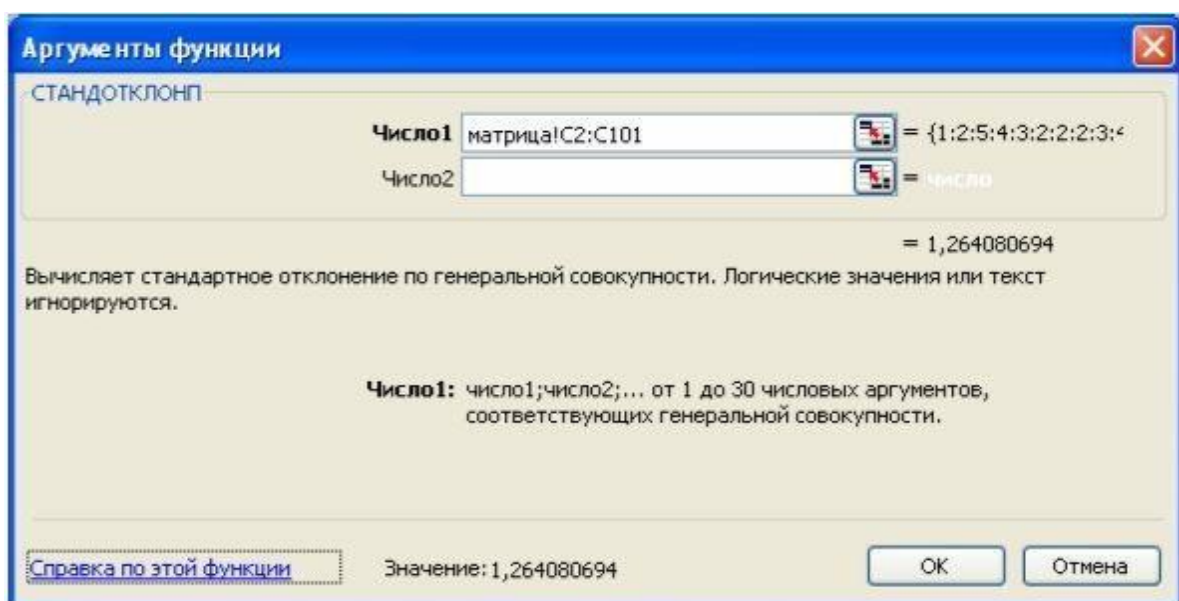


Рисунок 25 – Аргументы функции СТАНДОТКЛОНП

- В строке *Число1* введите: *матрица!C2:C101*. Нажмите ОК.
- После проведения всех вышеперечисленных расчетов лист «Удовлетворенность учебой» должен принять вид, аналогичный изображенному на рисунке 26.

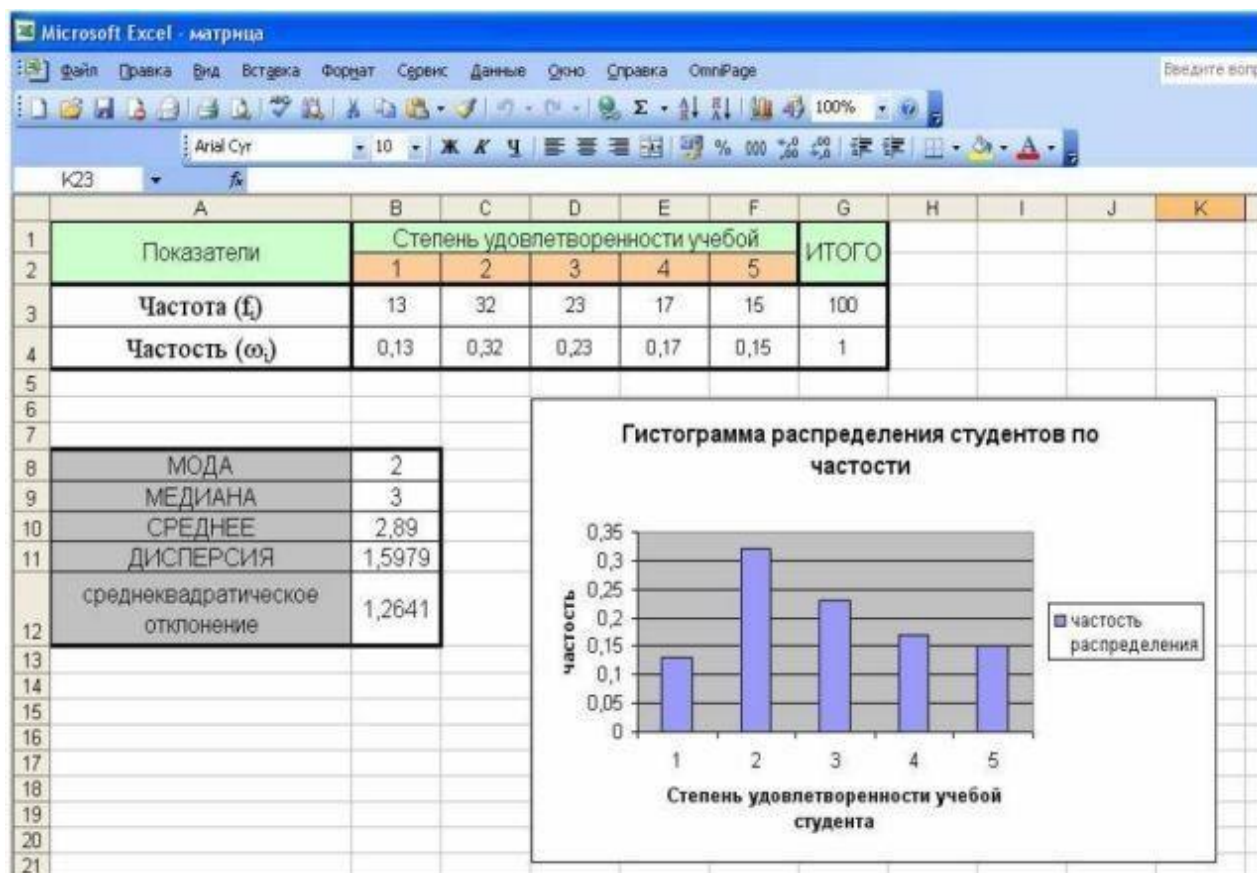


Рисунок 26 – Пример расчета мер нейтральной тенденции и мер вариации в MS Excel

- Сохраните изменения в документе «матрица».

Практическое занятие №3 АНАЛИЗ ВЗАИМОСВЯЗИ ПРИЗНАКОВ

Цель работы: выработать умения и навыки проведения анализа взаимосвязи признаков на основе критерия хи-квадрат и коэффициента взаимной сопряженности Пирсона, используя возможности редактора электронных таблиц.

Общие сведения

1. Критерий хи-квадрат

Одним из возможных критериев оценки взаимосвязи признаков является критерий хи-квадрат. Общий алгоритм использования критерия хи-квадрат для анализа взаимосвязи признаков выглядит следующим образом:

- 1) выбрать уровень значимости исследования (р-уровень);
- 2) построить таблицу сопряженности двух признаков;
- 3) рассчитать теоретические частоты признаков (см. формулу (9));
- 4) вычислить значение критерия хи-квадрат (см. формулу (10));
- 5) рассчитать количество степеней свободы (см. формулу (11));
- 6) сравнить полученное значение критерия хи-квадрат с табличным для выбранного р-уровня и соответствующего количества степеней свободы.

Значения теоретических частот вычисляются по формуле

$$n_{ij}^t = \frac{n_{0j} \times n_{i0}}{n_{00}}, \quad (9)$$

где n_{0j} – маргинальные частоты по столбцам (сумма по строкам);

n_{i0} – маргинальные частоты по строкам (сумма по столбцам);

n_{00} – общее число респондентов (или, в общем случае, испытаний).

Значения частоты в i -й строке и j -м столбце (n_{ij}) – это количество респондентов (или, в общем случае, частота испытаний), удовлетворяющих двум рассматриваемым признакам. Наиболее распространенным способом представления маргинальных и теоретических частот является построение таблицы сопряженности (корреляционной таблицы) по двум признакам.

Общая формула для вычисления критерия хи-квадрат выглядит следующим образом:

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^s \frac{(n_{ij} - n_{ij}^t)^2}{n_{ij}^t}, \quad (10)$$

где n_{ij}^t – значения теоретических частот;

n_{ij} – значение частоты в i -й строке и j -м столбце;

r – число столбцов;

s – число строк.

Для того чтобы на основе критерия хи-квадрат сделать вывод о том, имеется ли связь между двумя признаками, необходимо рассчитать количество степеней свободы и сравнить полученный результат с табличным для заданного p -уровня. Общая формула для вычисления количества степеней свободы критерия хи-квадрат выглядит следующим образом:

$$V = (r - 1) \times (s - 1), \quad (11)$$

где r – число столбцов таблицы сопряженности;

s – число строк таблицы сопряженности.

2. Коэффициент взаимной сопряженности Пирсона

С помощью вычисляемой величины χ^2 сравнивать силу связи затруднительно, потому что она зависит от объема выборки. К тому же величина χ^2 может принимать значения от 0 до бесконечности. Необходимо ввести нормировку. Одним из наиболее часто используемых коэффициентов корреляции является коэффициент взаимной сопряженности Пирсона, основанный на критерии хи-квадрат. Общая формула для вычисления данного коэффициента выглядит следующим образом:

$$C = \sqrt{\frac{\chi^2}{\chi^2 + N}}, \quad (12)$$

где N – объем выборки.

Значение коэффициента взаимной сопряженности Пирсона колеблется в диапазоне от 0 до +1. Коэффициент взаимной сопряженности Пирсона может быть применим для выявления силы связи между двумя признаками, измеряемыми по разным шкалам.

Корреляцию принято считать:

- 1) слабой, если коэффициент принимает значение от 0 до 0,3;
- 2) умеренной при коэффициенте от 0,3 до 0,5;
- 3) заметной при коэффициенте от 0,5 до 0,7;
- 4) высокой при коэффициенте свыше 0,7.

Порядок выполнения работы

Используя эмпирические данные, отраженные в матрице типа объект-признак (практическое занятие №1), постройте таблицу сопряженности по двум переменным (признакам): будущая профессия студентов и степень их удовлетворенности учебой. Выберите для данного исследования значение p -уровня. На основе критерия хи-квадрат установите наличие (или отсутствие) взаимосвязи между признаками; если взаимосвязь существует, то с помощью коэффициента взаимной сопряженности Пирсона оцените силу связи признаков. Выделите типологические группы по уровню удовлетворенности учебой и определите, какие из профессиональных групп наиболее однородны по степени удовлетворенности учебой.

1. Построение таблицы сопряженности (корреляционной таблицы)

- Выберите для данного исследования 5% уровень значимости ($p=0,05$).
- Откройте книгу «матрица».
- В данной книге создайте ещё один лист, измените его название на имя «Таблица сопряженности» и разместите этот лист в конец книги.
- На листе «Таблица сопряженности» создайте и оформите таблицу, как показано на рисунке 27. Ячейки для расчета маргинальных и теоретических частот пока оставьте пустыми.
- Используя эмпирические данные, занесенные в матрицу типа объект-признак, рассчитайте частоты для данной таблицы, то есть количество студентов определенной будущей профессии с определенной степенью удовлетворенности учебой (например, количество политологов с удовлетворенностью учебой, равной 1, количество историков с удовлетворенностью учебой, равной 4 и так далее). Для этого:

- 1) выделите ячейку B3;

Microsoft Excel - матрица								
Файл Правка Вид Вставка Формат Сервис Данные Окно Справка OmniPage								
Arial Cyr 10 Ж К Ч [бул] [бул] [бул] [бул] [бул] [бул] [бул] [бул] 100%								
D18 fx								
	A	B	C	D	E	F	G	H
1	Будущая профессия студента	Степени удовлетворенности учебой					маргинальные частоты по строкам	
2		1	2	3	4	5		
3	1. Политолог							
4								
5	2. Социолог							
6								
7	3. Культуролог							
8								
9	4. Филолог							
10								
11	5. Психолог							
12								
13	6. Историк							
14								
15	маргинальные частоты по столбцам Π_{0j}							
16								
17								

Рисунок 27 – Шаблон для построения таблицы сопряженности

2) в ячейку B3 введите следующую формулу: $=СУММ(ЕСЛИ(матрица!B2:B101=1;ЕСЛИ(матрица!C2:C101=1;1;0)))$. После набора формулы нажмите клавиши Ctrl+Shift+Enter, после этого в строке формул будет выражение в фигурных скобках. Таким образом зададим данную формулу как массив данных (см. приложение А). Вышеприведенной формулой для редактора электронных таблиц задали команду среди студентов всех профессий выбрать только политологов – профессиональная группа 1 (ЕСЛИ(матрица!B2:B101=1), у которых степень удовлетворенности учебой равна 1 (ЕСЛИ(матрица!C2:C101=1), и посчитать их количество (если выполняются оба этих условия (студент и политолог и степень удовлетворенности учебой равна 1), то MS Excel обозначает его как 1, если нет, то, как 0; после этого Excel суммирует все единицы, тем самым подсчитывает количество студентов, удовлетворяющих обоим условиям);

3) аналогичным образом посчитайте все остальные частоты (например, формула для расчета количества культурологов со степенью удовлетворенностью учебой равной 3 запишется в ячейку D7 следующим образом:

$\{=СУММ(ЕСЛИ(матрица!B2:B101=3;ЕСЛИ(матрица!C2:C101=3;1;0)))\}$.

• Рассчитайте маргинальные частоты по строкам n_{i0} . Для этого:

1) выделите ячейку G3;

2) введите в ячейку G3 следующую формулу: $=СУММ(B3:F3)$;

3) аналогичным образом рассчитайте оставшиеся маргинальные частоты по строкам. В данном случае не рекомендуется использовать автозаполнение, так как в ячейках 4-й, 6-й, 8-й, 12-й и 14-й строки будут рассчитаны не маргинальные, а теоретические частоты.

• Рассчитайте маргинальные частоты по столбцам n_{0j} . Для этого:

1) выделите ячейку B15;

2) введите в ячейку B15 следующую формулу: $=B3+B5+B7+B9+B11+B13$;

3) используя функцию автозаполнения (см. приложение А), рассчитайте оставшиеся маргинальные частоты по столбцам.

• Пересчитайте общее число объектов. Для этого:

1) выделите ячейку G15;

2) в ячейку G15 введите следующую формулу: $=СУММ(G3:G14)$.

• Используя формулу (9), рассчитайте теоретические частоты. Для этого:

1) выделите ячейку B4;

2) введите в ячейку B4 следующую формулу: $=B15*G3/G15$;

3) аналогично рассчитайте оставшиеся теоретические частоты.

• В итоге получите таблицу сопряженности, аналогичную таблице, изображенной на рисунке 28. В ячейках B3:F3, B5:F5, B7:F7, B9:F9, B11:F11 рассчитаны реальные частоты; в ячейках B4:F4, B6:F6, B8:F8, B10:F10, B12:F12 рассчитаны теоретические частоты.

Microsoft Excel - матрица

Файл Правка Вид Вставка Формат Сервис Данные Окно Справка OmniPage

Arial Cyr 10 Ж К Ч

D7 $\{=СУММ(ЕСЛИ(матрица!$B$2:$B$101=3;ЕСЛИ(матрица!$C$2:$C$101=3;1;0)))\}$

	A	B	C	D	E	F	G
1	Будущая профессия студента	Степени удовлетворенности учебой					маргинальные частоты по строкам
2		1	2	3	4	5	
3	1. Политолог	1	5	3	3	3	15
4		2	5	3	3	2	
5	2. Социолог	5	5	5	1	0	16
6		2	5	4	3	2	
7	3. Культуролог	1	6	5	5	2	19
8		2	6	4	3	3	
9	4. Филолог	4	7	4	4	3	22
10		3	7	5	4	3	
11	5. Психолог	1	5	2	1	3	12
12		2	4	3	2	2	
13	6. Историк	1	4	4	3	4	16
14		2	5	4	3	2	
15	маргинальные частоты по столбцам Π_{0j}	13	32	23	17	15	100

Рисунок 28 – Таблица распределения студентов по их будущей профессии и степени их удовлетворенности учебой

2. Расчет критерия хи-квадрат и коэффициента Пирсона

- Рассчитайте критерий хи-квадрат для данной таблицы используя, формулу (10). Для этого выделите ячейку B19 и введите в неё следующую формулу: $=СУММ(((B3-B4)^2)/B4;((C3-C4)^2)/C4;((D3-D4)^2)/D4;((E3-E4)^2)/E4;((F3-F4)^2)/F4;((B5-B6)^2)/B6;((C5-C6)^2)/C6;((D5-D6)^2)/D6;((E5-E6)^2)/E6;((F5-F6)^2)/F6;((B7-B8)^2)/B8;((C7-C8)^2)/C8;((D7-D8)^2)/D8;((E7-E8)^2)/E8;((F7-F8)^2)/F8;((B9-B10)^2)/B10;((C9-C10)^2)/C10;((D9-D10)^2)/D10;((E9-E10)^2)/E10;((F9-F10)^2)/F10;((B11-B12)^2)/B12;((C11-C12)^2)/C12;((D11-D12)^2)/D12;((E11-E12)^2)/E12;((F11-F12)^2)/F12;((B13-B14)^2)/B14;((C13-C14)^2)/C14;((D13-D14)^2)/D14;((E13-E14)^2)/E14;((F13-F14)^2)/F14)$. Данная формула, записанная в редакторе электронных таблиц, справедлива для таблицы сопряженности, представленной на рисунке 28.

- Используя формулу (11), рассчитайте количество степеней свободы.

- Сравните найденное вами значение критерия хи-квадрат с табличным значением критерия для данного уровня значимости исследования ($p=0,05$) и данного количества степеней свободы. Сделайте вывод о наличии (или отсутствии) связи между будущей профессией студентов и степенью их удовлетворенности учебой. Табличное значение критерия хи-квадрат для данного p -уровня и данного количества степеней свободы можно получить в редакторе электронных таблиц, используя функцию *ХИ2ОБР*. При этом в поле *Вероятность* аргумента функции необходимо ввести 95% (рисунок 29), так как в данном исследовании $p=0,05$, а в поле *Степени_свободы* – количество степеней свободы (см. формулу (11)). Кроме того, табличное значение критерия хи-квадрат можно узнать из таблицы в приложении Б.

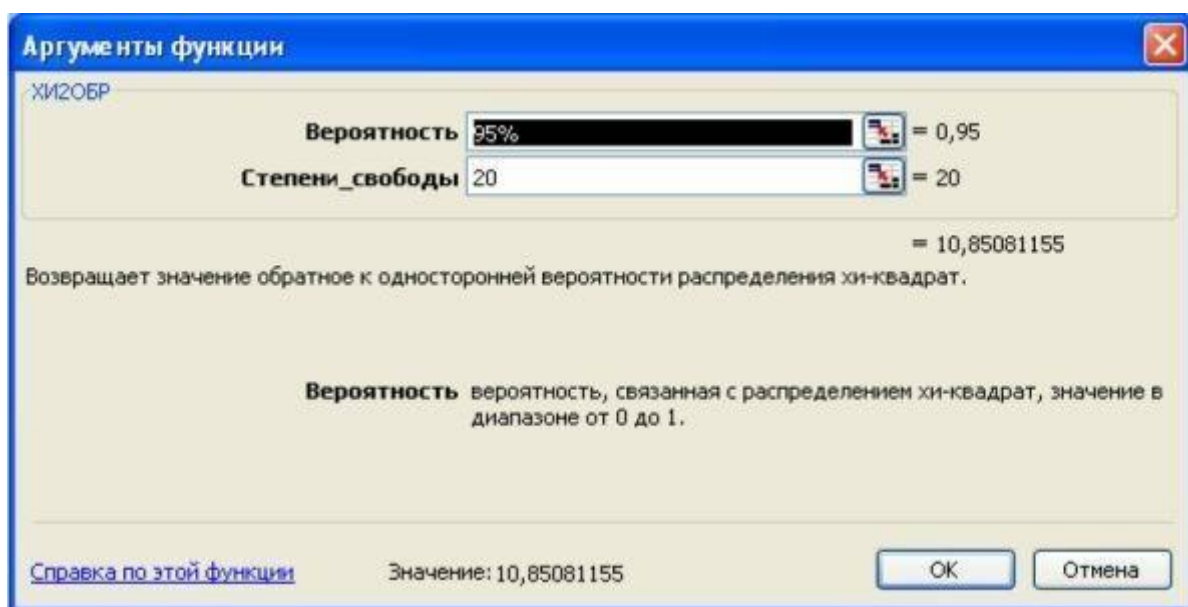


Рисунок 29 – Аргументы функции ХИ2ОБР

- Используя формулу (12) найдите коэффициент взаимной сопряженности Пирсона. Для этого выделите ячейку B20 и введите в нее формулу $=КОРЕНЬ(B19/(B19+G15))$. В итоге получите таблицу MS Excel аналогичную, представленной на рисунке 30. Сделайте вывод о силе связи между будущей профессией студентов и степенью их удовлетворенности учебой.

	A	B	C	D	E	F	G
1	Будущая профессия студента	Степени удовлетворенности учебой					маргинальные частоты по строкам
2		1	2	3	4	5	
3	1. Политолог	1	5	3	3	3	15
4		2	5	5	3	2	
5	2. Социолог	5	5	5	1	0	16
6		2	5	4	3	2	
7	3. Культуролог	1	5	5	5	2	19
8		2	6	4	3	3	
9	4. Филолог	4	7	4	4	3	22
10		3	7	5	4	3	
11	5. Психолог	1	5	2	1	3	12
12		2	4	3	2	2	
13	6. Историк	1	4	4	3	4	16
14		2	5	4	3	2	
15	маргинальные частоты по столбцам P_{0j}	13	32	23	17	15	100
16							
17							
19	хи-квадрат	15,85					
20	КОЭФФИЦИЕНТ ПИРСОНА	0,37					
21							
22	хи-квадрат табличное	10,85					
23							

Рисунок 30 – Пример расчета критерия хи-квалрат и коэффициента Пирсона в Excel

3. Выделение типологических групп. Определение однородности распределения

- Для визуального сравнения структуры удовлетворенности учебой в разных группах студентов постройте в одной системе координат полигоны частот, отражающие распределение студентов всех шести профессиональных групп по степени их удовлетворенности учебой. В итоге получите систему графиков, аналогичную системе, представленной на рисунке 31. Сравните полученные полигоны частот и выделите типологические группы, то есть выделите профессиональные группы, наиболее похожие по степени удовлетворенности учебой.

- Определите студенческую группу с наибольшей степенью удовлетворенности учебой. Для этого, используя функцию *СРЗНАЧ* мастера функций редактора электронных таблиц, рассчитайте среднее арифметическое значение степени удовлетворенности учебой в каждой профессиональной группе студентов.

- Найдите наиболее однородную группу студентов по степени удовлетворенности учебой. Для этого, используя функцию

СТАНДОТКЛОНП, вычислите стандартное отклонение степени удовлетворенности учебой в каждой профессиональной группе. В итоге получите таблицу аналогичную таблице, представленной на рисунке 32.

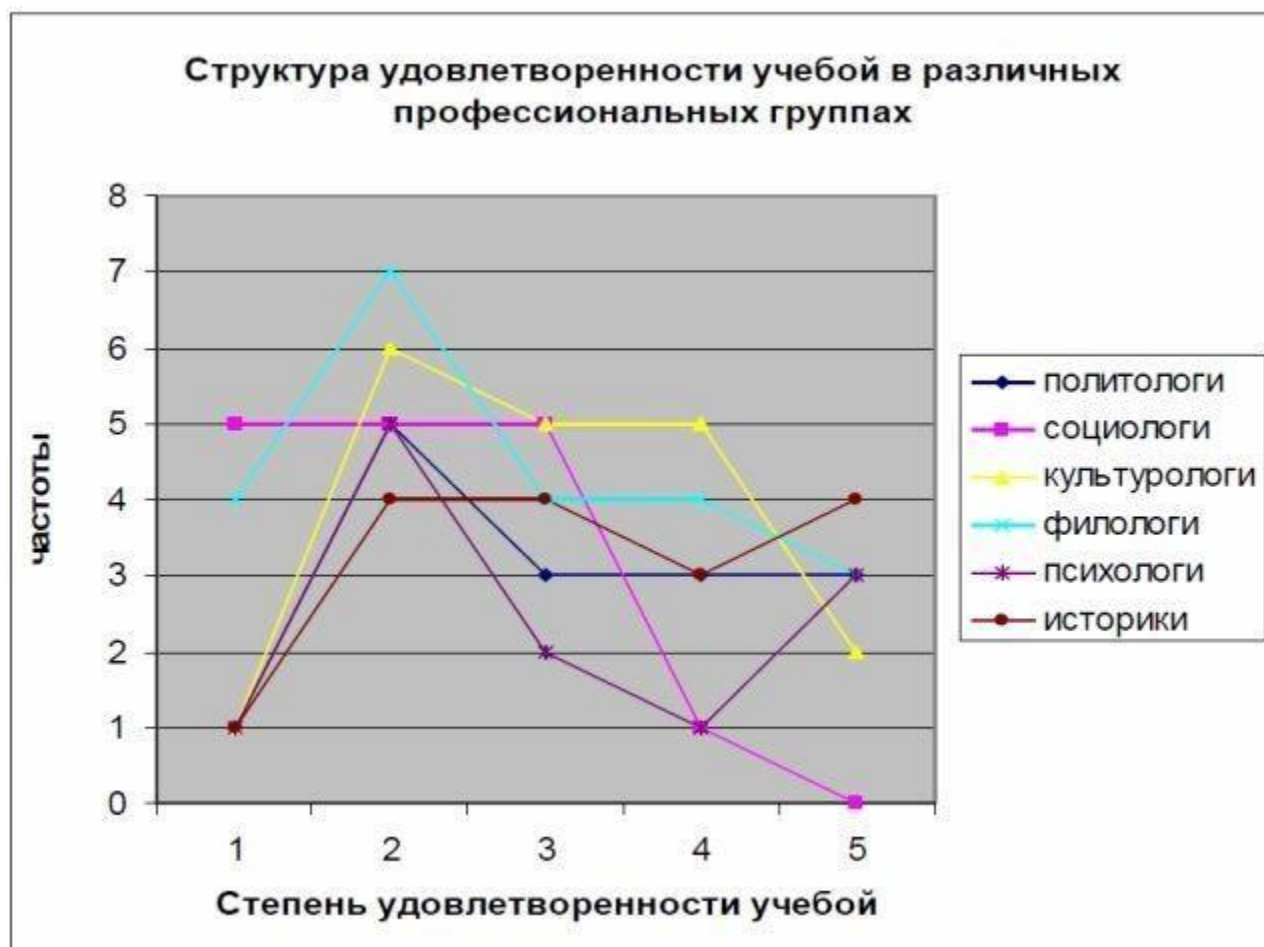


Рисунок 31 – Структура удовлетворенности учебой в различных профессиональных группах

- На основании полученных данных, выделите наиболее схожие профессиональные группы по степени удовлетворенности учебой. Сравните эти выводы со сделанными вами после выделения типологических групп.

Самостоятельные задания

1. В книге «матрица» создайте новый лист и назовите его «Таблица сопряженности 2». На листе «Таблица сопряженности 2» постройте корреляционную таблицу распределения студентов по двум признакам: будущая профессия студента и время,

Рисунок 32 – Пример расчета среднего арифметического значения и стандартного отклонения степени удовлетворенности учебой в различных профессиональных группах

2. В книге «матрица» создайте новый лист и назовите его «Таблица сопряженности 3». На листе «Таблица сопряженности 3» постройте корреляционную таблицу распределения студентов по двум признакам: время, затрачиваемое студентами на учебу и степень их удовлетворенности учебой. На основе критерия хи-квадрат сделайте вывод о наличии связи между этими признаками. Используя коэффициент взаимной сопряженности Пирсона, установите силу связи между этими признаками.

Практическое занятие №4
**РАСЧЕТ КРИТЕРИЯ СТЬЮДЕНТА. ВЫЧИСЛЕНИЕ
 КОЭФФИЦИЕНТА ЛИНЕЙНОЙ КОРРЕЛЯЦИИ ПИРСОНА.
 ЗАПИСЬ УРАВНЕНИЯ ЛИНЕЙНОЙ РЕГРЕССИИ**

Цели работы:

- 1) научиться проводить проверку истинности научных гипотез с помощью критерия Стьюдента;
- 2) научиться устанавливать степень взаимосвязи между двумя признаками, используя коэффициент линейной корреляции Пирсона;
- 3) научиться выводить функциональную зависимость между двумя признаками путем записи уравнения линейной регрессии.

Общие сведения

Критерий Стьюдента направлен на оценку различий величин средних значений двух выборок, которые распределены по нормальному закону. Одним из главных достоинств критерия является широта его применения. Он может быть использован для сопоставления средних значений у связанных и несвязанных выборок, причем выборки могут быть не равны по величине.

Общая формула для вычисления t-критерия для равных выборок (либо незначительно отличающихся по объему) выглядит следующим образом:

$$t = \frac{|M_1 - M_2|}{\sqrt{\frac{\sigma_1^2}{N_1} + \frac{\sigma_2^2}{N_2}}}, \quad (13)$$

где M_1, M_2 – средние арифметические;
 σ_1, σ_2 – среднеквадратические (стандартные) отклонения
 (следовательно, σ_1^2, σ_2^2 – дисперсии);

N_1, N_2 – объемы выборок.

Коэффициент линейной корреляции Пирсона (R) служит для измерения силы связи двух признаков, каждый из которых измеряется по интервальной шкале. Диапазон его изменения – от –1 до 1, притом значение ~ 0 означает отсутствие связи между ответами на изучаемые вопросы (т.е. то, что признаки независимы). Значение $R \sim 1$ означает, что чем больше у респондента значение 1-го фактора,

тем больше у него ожидаемое значение 2-го фактора. Наоборот, $R \sim -1$ означает, что чем больше значение 1-го фактора, тем меньше ожидаемое значение 2-го фактора. Общая формула для вычисления коэффициента линейной корреляции Пирсона выглядит следующим образом:

$$R = \frac{\sum (x_i \times y_i) - N \times \bar{x} \times \bar{y}}{N \times \sigma_x \times \sigma_y}, \quad (14)$$

где x_i и y_i – значение i -го признака в 1-й и 2-й «группе исследований» соответственно;

\bar{x}, \bar{y} – средние значения;

σ_x и σ_y – стандартные отклонения;

N – объем выборки.

Уравнение линейной регрессии – это уравнение прямой линии $y = b_1 \cdot x + b_0$, (15)

описывающее зависимость ожидаемых значений одной случайной величины (Y) от значений другой случайной величины (X).

Предполагается, что связь между величинами линейная. Если мы знаем уравнение линейной регрессии, то по ответу человека на вопрос X мы можем спрогнозировать (с некоторой точностью) его ответ на вопрос Y .

Коэффициент b_1 в уравнении регрессии с математической точки зрения имеет смысл тангенса угла наклона прямой $Y(x)$, а с другой стороны, имеет смысл коэффициента линейной корреляции Пирсона (R), нормированного на величины стандартных отклонений (σ_x, σ_y):

$$b_1 = R \cdot \frac{\sigma_x}{\sigma_y}. \quad (16)$$

Коэффициент b_0 имеет смысл значения y в точке пересечения оси y (при $x=0$). Его можно найти, если подставить в уравнение регрессии значение b_1 и средние значения (\bar{x}, \bar{y}):

$$\bar{y} = b_1 \cdot \bar{x} + b_0. \quad (17)$$

Порядок выполнения работы

1. Вычисление критерия Стьюдента

С помощью критерия Стьюдента проверьте справедливость гипотезы о том, что средний возраст женатых жителей города N равен среднему возрасту холостых жителей города N.

- Откройте новый документ Excel и введите следующие данные: возраст женатых людей: 30, 28, 19, 19, 23, 19, 27, 24, 26, 20, 26 и 21 лет; возраст холостых: 29, 30, 21, 41, 32, 33, 18, 21, 18, 31, 27, 20 и 19 лет.

- Используя формулу (13), рассчитайте критерий Стьюдента. Для этого в любую свободную ячейку введите формулу $=ABS(CPЗНАЧ(B3:B14)-CPЗНАЧ(C3:C15))/(КОРЕНЬ((ДИСПР(B3:B14)/СЧЁТ(B3:B14))+(ДИСПР(C3:C15)/СЧЕТ(C3:C15))))$. В данной формуле адреса ячеек соответствуют адресам ячеек в таблице, изображенной на рисунке 33. В данной формуле: 1) функция ABS выводит абсолютное значение (модуль) числа; 2) функция CPЗНАЧ вычисляет среднее значение; 3) функция КОРЕНЬ вычисляет корень квадратный; 4) функция ДИСПР вычисляет дисперсию; 5) функция СЧЕТ вычисляет количество ячеек в диапазоне (в данном случае – объем выборки).

- Рассчитайте количество степеней свободы, используя формулу $V=N_1+N_2-2$, (18) где N_1, N_2 – объемы выборок.

Для этого в любую свободную ячейку введите формулу $=СЧЁТ(B3:B14)+СЧЁТ(C3:C15)-2$. В данной формуле адреса ячеек соответствуют адресам ячеек в таблице, изображенной на рисунке 33.

- Выберите значение уровня значимости исследования $p=0,05$ и рассчитайте для данного p -уровня и данного количества степеней свободы табличное значение критерия Стьюдента. Для этого в любую свободную ячейку введите формулу $=СТЮДРАСПОБР(0,05;G4)$ или воспользуйтесь таблицей приложения В. В данной формуле адреса ячеек соответствуют адресам ячеек в таблице, изображенной на рисунке 33.

- Сравните полученное вами значение критерия Стьюдента с табличным и сделайте вывод о верности (или не верности) выдвинутой гипотезы.

Microsoft Excel - Книга1

Файл Правка Вид Вставка Формат Сервис Данные Окно Справка ОптимPage

Введите вопрос

Г3 $=ABS(CPЗНАЧ(B3:B14)-CPЗНАЧ(C3:C15))/(КОРЕНЬ((ДИСПР(B3:B14)/СЧЁТ(B3:B14))+ДИСПР(C3:C15)/СЧЁТ(C3:C15))))$

	A	B	C	D	E	F	G	H	I
1		средний возраст женатых	средний возраст холостых						
2									
3		30	29		Коэффициент Стьюдента		1,21		
4		28	30		Степени свободы		23		
5		19	21		Коэффициент Стьюдента табл		2,07		
6		19	41						
7		23	32						
8		19	33						
9		27	18						
10		24	21						
11		26	18						
12		20	31						
13		26	27						
14		21	20						
15			19						

Рисунок 33 – Пример расчета критерия Стьюдента в MS Excel

- Сохраните данный документ Excel под именем «критерий Стьюдента».

2. Вычисление коэффициента линейной корреляции Пирсона

С помощью коэффициента линейной корреляции Пирсона установите, насколько сильно успеваемость учеников по математике связана с успеваемостью учеников по информатике.

- Откройте новый документ Excel и введите следующие данные:
 - оценка по математике 5 4 2 4 5 3 2 4 4;
 - оценка по информатике 4 4 3 5 5 3 3 5 4.
- Используя формулу (15), рассчитайте коэффициент линейной корреляции Пирсона. Для этого в любой свободной ячейке вызовите *Мастер функций* MS Excel с помощью команды f_x – *Вставка функции*.
- В диалоговом окне выберите категорию *Статистические*. В выпадающем списке выберите название функции *ПИРСОН*. Нажмите ОК. Появится окно *Аргументы функции* (рисунок 34).

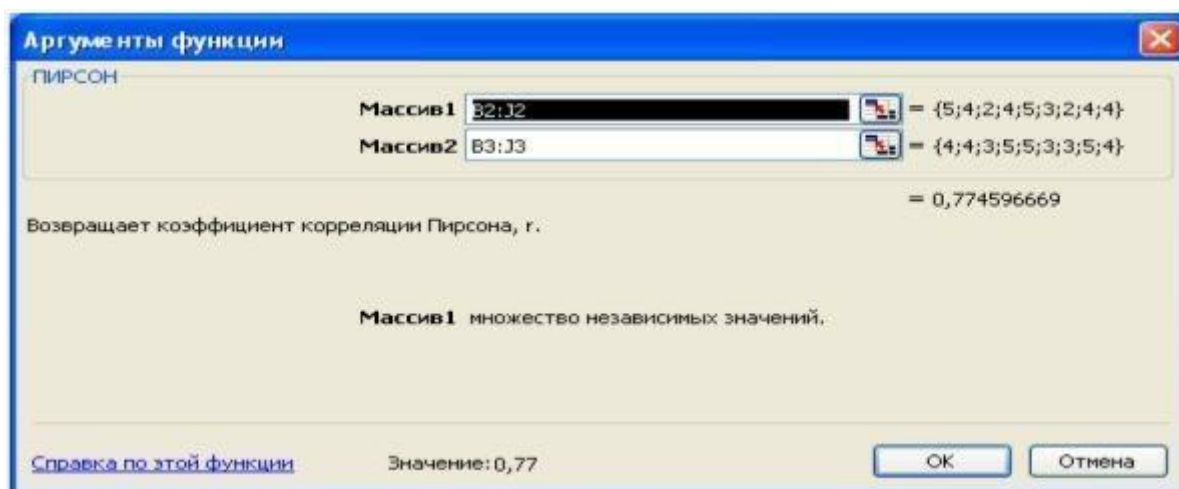


Рисунок 34 – Аргументы функции ПИРСОН

• В строку *Массив1* введите диапазон ячеек B2:J2. В строку *Массив2* введите диапазон ячеек B3:J3. Нажмите ОК. В данной формуле адреса ячеек соответствуют адресам ячеек в таблице, изображенной на рисунке 35.

	A	B	C	D	E	F	G	H	I	J
1										
2	Оценки по математике	5	4	2	4	5	3	2	4	4
3	Оценки по информатике	4	4	3	5	5	3	3	5	4
4										
5										
6	Коэффициент Пирсона	0,77								
7										

Рисунок 35 – Пример вычисления коэффициента линейной корреляции Пирсона в MS Excel

• На основании вычисленного коэффициента линейной корреляции Пирсона сделайте вывод о силе связи между двумя признаками (переменными): успеваемость учеников по математике и успеваемость учеников по информатике.

- Сохраните данный документ Excel под именем «*коэффициент линейной корреляции Пирсона*».

3. Запись уравнения линейной регрессии

Запишите уравнение линейной регрессии для двух признаков: успеваемость учеников по математике и успеваемость учеников по информатике (см. пункт 2 порядка выполнения работы).

- Откройте документ Excel «*коэффициент линейной корреляции Пирсона*».

- Используя формулу (17), рассчитайте коэффициент b_1 для записи уравнения линейной регрессии. Для этого в ячейку B10 введите следующую формулу:

$$=B6*(СТАНДОТКЛОНП(B3:J3)/СТАНДОТКЛОНП(B2:J2)).$$
С помощью функции $СТАНДОТКЛОНП(B3:J3)$ вычисляем стандартное отклонение успеваемости по информатике (σ_y), с помощью функции $СТАНДОТКЛОНП(B2:J2)$ вычисляем стандартное отклонение успеваемости по математике (σ_x). В ячейке B6 записано значение коэффициента линейной корреляции Пирсона, вычисленное в пункте 2 данной работы.

- Используя формулу (18), рассчитайте коэффициент b_0 для записи уравнения линейной регрессии. Для этого в ячейку B11 введите следующую формулу:

$$=CPЗНАЧ(B3:J3)-B10*CPЗНАЧ(B2:J2).$$
С помощью функции $CPЗНАЧ(B3:J3)$ вычисляем среднее арифметическое значений успеваемости по информатике, с помощью функции $CPЗНАЧ(B2:J2)$ вычисляем среднее арифметическое значений успеваемости по математике. В ячейке B10 записано значение коэффициента b_1 .

- В итоге, используя формулу (16) и значения коэффициентов b_1 и b_0 , получите уравнение линейной регрессии. Таблица редактора электронных таблиц будет аналогична представленной на рисунке 35.

Самостоятельные задания

1. В книге «матрица» создайте новый лист и назовите его «критерий Стьюдента». На данном листе, используя критерий Стьюдента, подтвердите (или опровергните) гипотезу о том, что

средняя удовлетворенность учебой политологов и историков одинакова.

2. С помощью критерия Стьюдента определите, повлиял ли внедренный в коллектив психологический метод, повысивший уровень развития межличностных отношений группы, на эффективность деятельности группы (таблица 1). Для этого сравните показатели эффективности групповой деятельности экспериментальной группы (то есть коллектива, в который был внедрен метод, повышающий уровень межличностных отношений) и контрольной (то есть коллектив развивающийся «стандартно», без внедрения новой психологической методики, повышающей уровень межличностных отношений). Оценка уровня эффективности деятельности происходит по 7 различным показателям, каждый из которых оценивается по 10-балльной шкале. Данные эксперимента приведены в таблице 1.

Таблица 1 – Оценка эффективности групповой деятельности

№ показателя эффективности групповой деятельности	Экспериментальная группа	Контрольная группа
	Значения эффективности деятельности	Значения эффективности деятельности
1	5	6
2	6	3
3	7	4
4	10	5
5	6	5
6	8	3
7	7	2

3. В книге «коэффициент линейной корреляции Пирсона» перейдите на «Лист2». На данном листе, используя коэффициент линейной корреляции Пирсона, установите силу взаимосвязи между двумя признаками (переменными): успеваемость учеников по литературе и успеваемость учеников по истории. Для этого используйте следующие данные:

- оценки по литературе: 4 4 5 3 3 4 3 5 5 5;
- оценки по истории: 3 4 5 3 4 5 4 4 5 4.

4. В книге «*коэффициент линейной корреляции Пирсона*» перейдите на «Лист2». На данном листе вычислите коэффициенты и запишите уравнение линейной регрессии для двух признаков: успеваемость учеников по литературе и успеваемость учеников по истории (данные смотри в пункте 3 самостоятельного задания работы №4).

Практическое занятие №5

УСТАНОВЛЕНИЕ ФУНКЦИОНАЛЬНОЙ ЗАВИСИМОСТИ ДЛЯ ИССЛЕДУЕМЫХ ПАРАМЕТРОВ

Цели работы.

1) научиться выводить функциональную зависимость между двумя признаками графическим способом, используя возможности редактора электронных таблиц:

2) научиться устанавливать степень взаимосвязи между двумя признаками, значения которых получены с помощью различных шкал и измеряются в различных единицах измерения.

Общие сведения

В научных исследованиях часто важно не просто установить наличие взаимосвязи и силу связи между двумя признаками, но и определить, если это возможно, какова функциональная зависимость между этими признаками. Если исследователю удалось установить функциональную зависимость между двумя признаками, то по значению первого признака он может спрогнозировать значение второго. Так, например, в работе №4 с помощью уравнения линейной регрессии установили, что между рассматриваемыми признаками (успеваемость по математике и успеваемость по информатике) существует следующая функциональная зависимость: $y = 0,6 \cdot x + 1,8$. Во многих случаях функциональная зависимость между рассматриваемыми признаками носит более сложный характер, чем линейная зависимость. Наиболее просто установить, какова функциональная зависимость между взаимосвязанными признаками, графическим способом, используя возможности добавления линии тренда на графике в редакторе электронных таблиц.

Кроме того, в научных исследованиях бывают ситуации, когда необходимо устанавливать силу связи между двумя признаками (переменными), значения которых получены в ходе проведения одного научного эксперимента, но измеряются с помощью различных шкал и в разных единицах измерения. В этом случае формула для вычисления коэффициента корреляции будет выглядеть следующим образом:

$$r_{x,y} = \frac{\text{cov}(x,y)}{\sigma_x \sigma_y}. \quad (19)$$

Наиболее просто установить силу связи между признаками в этом случае можно с помощью функции *КОРРЕЛ* мастера функций редактора электронных таблиц.

Порядок выполнения работы

Установите какова функциональная зависимость и какова сила связи между двумя признаками, полученными в ходе выполнения научного эксперимента по вычислению энергии кислотообразования в процессе сквашивания: временем сквашивания продукта и кислотностью. Значения времени сквашивания продукта: 1, 2, 3, 5, 6, 7, 8, 9; соответственно значения кислотности: 18, 21, 26, 34, 44, 58, 68, 82.

- Откройте новый документ Excel.
- Внесите в ячейки приведенные выше значения времени сквашивания продукта и значения кислотности.
- Используя мастер диаграмм редактора электронных таблиц, постройте график зависимости между данными признаками. *По оси X откладывать данные для времени сквашивания; по оси Y – данные для кислотности (то есть в качестве ДИАПАЗОНА данных указать значения кислотности, в качестве ПОДПИСЕЙ ОСИ X – данные для времени сквашивания).*
- Щелкните правой кнопкой мыши по графику. В появившемся контекстном меню выберите пункт **ДОБАВИТЬ ЛИНИЮ ТРЕНДА**
- На вкладке **ТИП** окна **ЛИНИЯ ТРЕНДА** (рисунок 36) выберите тип функции, которая наилучшим образом описывает исследуемый процесс. В данном случае это *Полиномиальная*.

На вкладке **ПАРАМЕТРЫ** окна **ЛИНИЯ ТРЕНДА** (рисунок 37) для вывода уравнения функции отметьте позиции:

- 1) показать уравнение на диаграмме;
- 2) поместить на диаграмму величину достоверности аппроксимации (R^2).

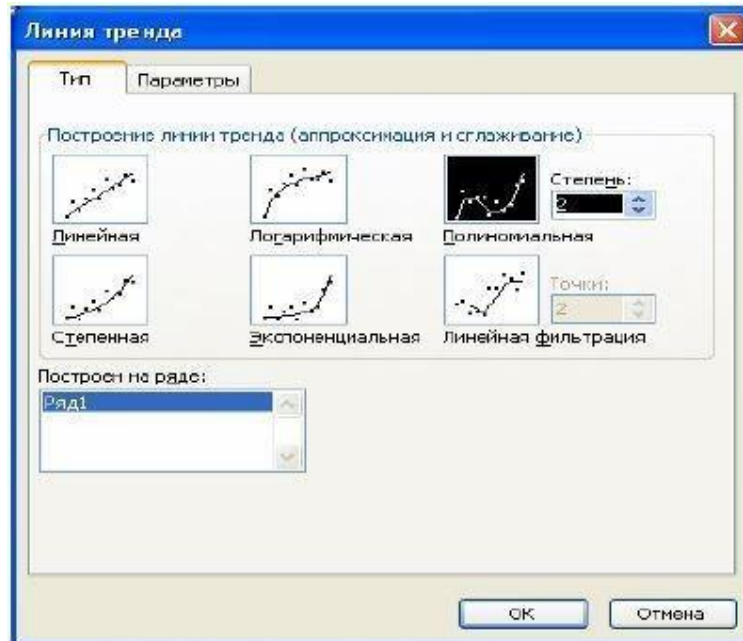


Рисунок 36 – Вкладка Тип окна Линия тренда

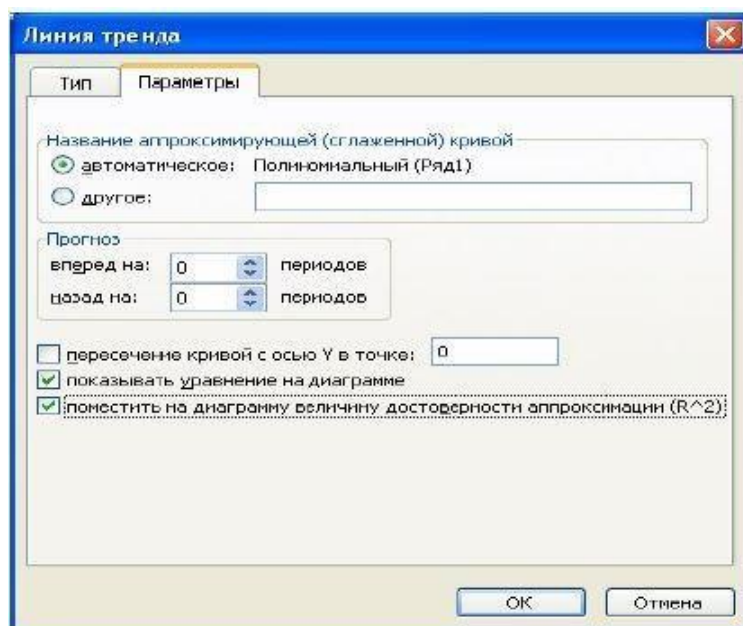


Рисунок 37 – Вкладка Параметры окна Линия тренда

В итоге получите график, аналогичный изображенному на рисунке 38, а функциональная зависимость между рассматриваемыми признаками примет вид: $y = 0,875 \cdot x^2 + 1,5179 \cdot x + 14,732$. R^2 должно быть не менее 0,88, если величина достоверности аппроксимации меньше, чем 0,88, необходимо выбрать другой ТИП, приближающий R^2 к единице.

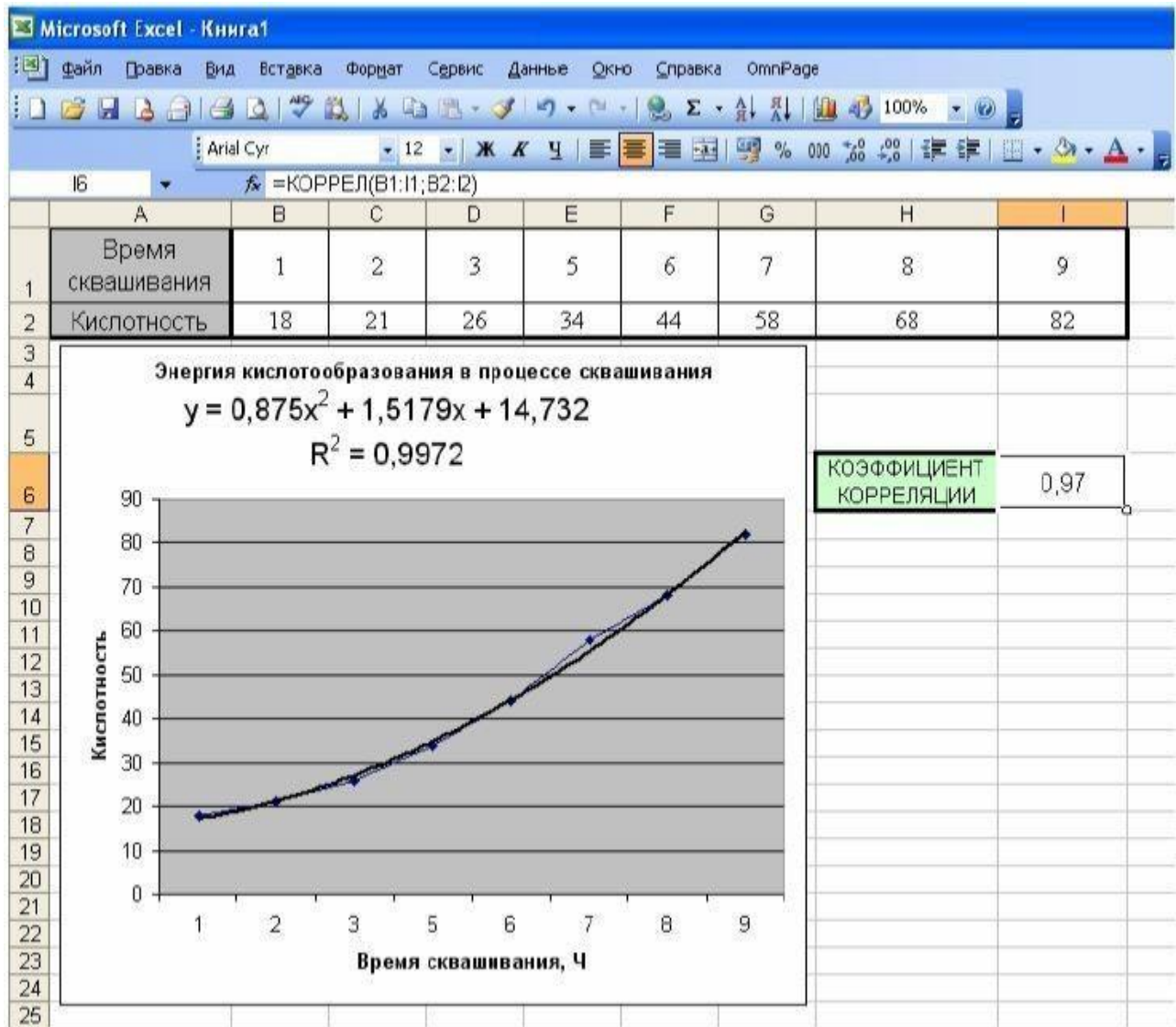


Рисунок 38 –Пример определения функциональной зависимости и силы связи между двумя признаками в MS Excel

Рассчитайте силу связи между временем сквашивания продукта и кислотностью. Для этого:

- 1) выделите любую пустую ячейку:
- 2) щелкните левой кнопкой мыши, по команде fx – *Вставка функции* на стандартной панели инструментов и вызовите *Мастер функций* редактора электронных таблиц:
- 3) в диалоговом окне выберите категорию *Статистические*. В выпадающем списке выберите название функции *KORREL*. Нажмите ОК. Появится окно *Аргументы функции* (рисунок 39);

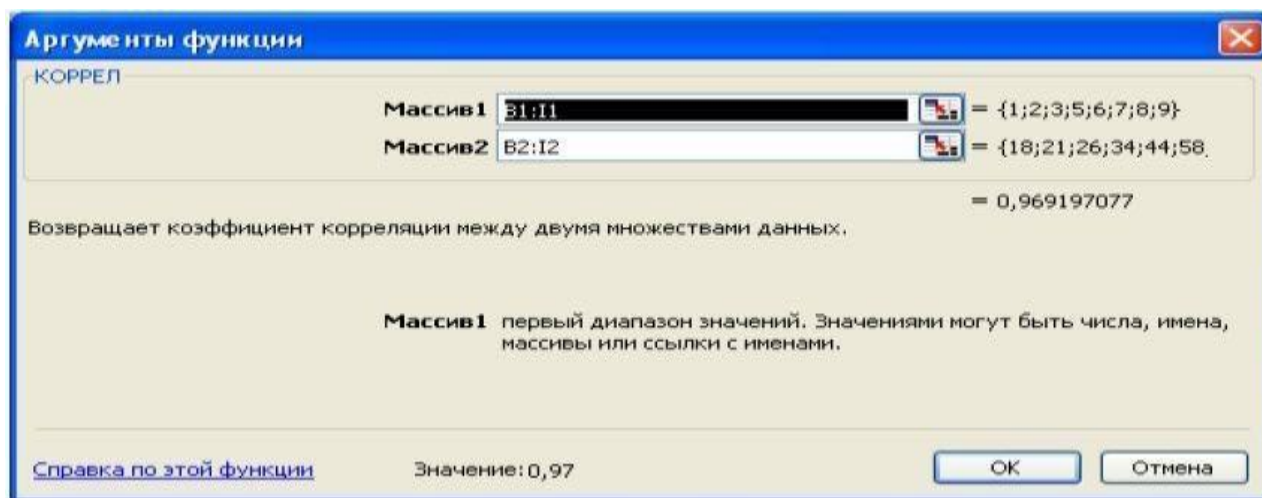


Рисунок 39 – Аргументы функции КОРРЕЛ

4) в поле *Массив1* введите B1:I1 (значения времени сквашивания продукта), в поле *Массив2* введите B2:I2 (значения кислотности). Нажмите ОК.

- Сделайте вывод о том, какова сила связи между данными признаками.

Самостоятельные задания

1. Провести статистическую обработку результатов анализа жирности молока в пробе с доверительной вероятностью $P=0,95$, если в опыте было проведено 8 анализов и получены следующие результаты:

№ анализа	1	2	3	4	5	6	7	8
Жирность молока, %	3,80	3,75	3,82	3,83	3,78	3,90	3,92	3,84

2. Установить корреляционную и функциональную зависимости между концентрацией соли в сыре (C_c) и температурой его замерзания (T_z), если в 8 опытах были получены следующие результаты:

C_c	X	1,0	1,3	2,2	2,6	5,5	3,6	5,3	6,0
T_z	Y	-1,3	-2,2	-2,8	-3,2	-3,8	-4,4	-5,6	5,9

3. Провести статистическую обработку результатов анализа содержания жира в масле с доверительной вероятностью $P=0,90$, если было проведено 8 анализов и получены следующие результаты:

№ анализа	1	2	3	4	5	6	7	8
Жирность молока, %	73,6	74,8	73,3	73,0	74,6	74,6	73,4	73,3

4. Установить корреляционную и функциональную зависимости между временем культивирования молочнокислых бактерий и нарастанием кислотности молока, если в опыте получены результаты:

X, ч	0	1	2	3	4	5	6	7	8
Y, Т	19	21	23	28	33	46	60	63	78

5. Провести статистическую обработку результатов анализа кислотности кефира с доверительной вероятностью $P=0,95$, если в опыте были получены следующие результаты:

№ анализа	1	2	3	4	5	6	7	8
Кк, Т°	86	87	83	88	87	86	88	89

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

- 1 Айвазян С.А. Прикладная статистика в задачах и упражнениях [Текст] / С.А. Айвазян, В.С. Мхитарян. – М.: Юнити-Дана, 2001. 270 с.
- 2 Айвазян С.А. Прикладная статистика и основы эконометрики [Текст] / С.А. Айвазян, В.С. Мхитарян. – М.: Юнити, 2005. 1022 с.
- 3 Борзов Д.Б. Информатика [Текст]: учебное пособие / Д.Б. Борзов, И.Е. Чернецкая. - Курск: ЮЗГУ, 2010. - 305 с.
- 4 Вуколов Э.А. Основы статистического анализа: Практикум по статистическим методам и исследованию операций с использованием пакетов STATISTICA и EXCEL [Текст] / Э.А. Вуколов. – М.: Форум, 2008. 464 с.
- 5 Геронимус Б.А. Экономико-математические методы в планировании на автомобильном транспорте [Текст] / Б.А. Геронимус. – М.: Транспорт, 1982.
- 6 Дегтярев Ю.И. Исследование операций [Текст] / Ю.И. Дегтярев. – М.: Высшая школа, 1986.
- 7 Дик В.В. Информационные системы в экономике [Текст] / В.В. Дик. М.: Финансы и статистика, 1996. 250 с.
- 8 Зайцев Е.И. Информационные технологии в управлении эксплуатационной эффективностью автотранспорта [Текст] / Е.И. Зайцев. СПб.: СПбГИАЭ, 1998. 227 с.
- 9 Кузнецов Ю.Н. Математическое программирование [Текст] / Ю.Н. Кузнецов, В.И. Кузубов, А.Б. Волощенко. – М.: Высшая школа, 1980.
- 10 Луенбергер Д.Д. Информатика [Текст]: учебно-методическое пособие / Д.Д. Луенбергер ; пер. с англ. Ю.Л. Цвирко; под ред. К.К. Колина. - М.: Техносфера, 2008. 448 с.
- 11 Смородинский С.С. Анализ и оптимизация систем на основе аналитических моделей [Текст] / С.С. Смородинский, Н.В. Батин. – Минск: БГУИР, 1997.

Приложение А

ТЕОРЕТИЧЕСКИЕ СВЕДЕНИЯ ПО ОСОБЕННОСТЯМ РАБОТЫ В РЕДАКТОРЕ ЭЛЕКТРОННЫХ ТАБЛИЦ MS EXCEL

Табличный процессор MS Excel используют для хранения и автоматизации расчета данных, представленных в табличной форме.

Документы Excel называют рабочими книгами. Они сохраняются в файлах с расширением XLS. Внешний вид экрана табличного процессора приведен на рисунке 40.

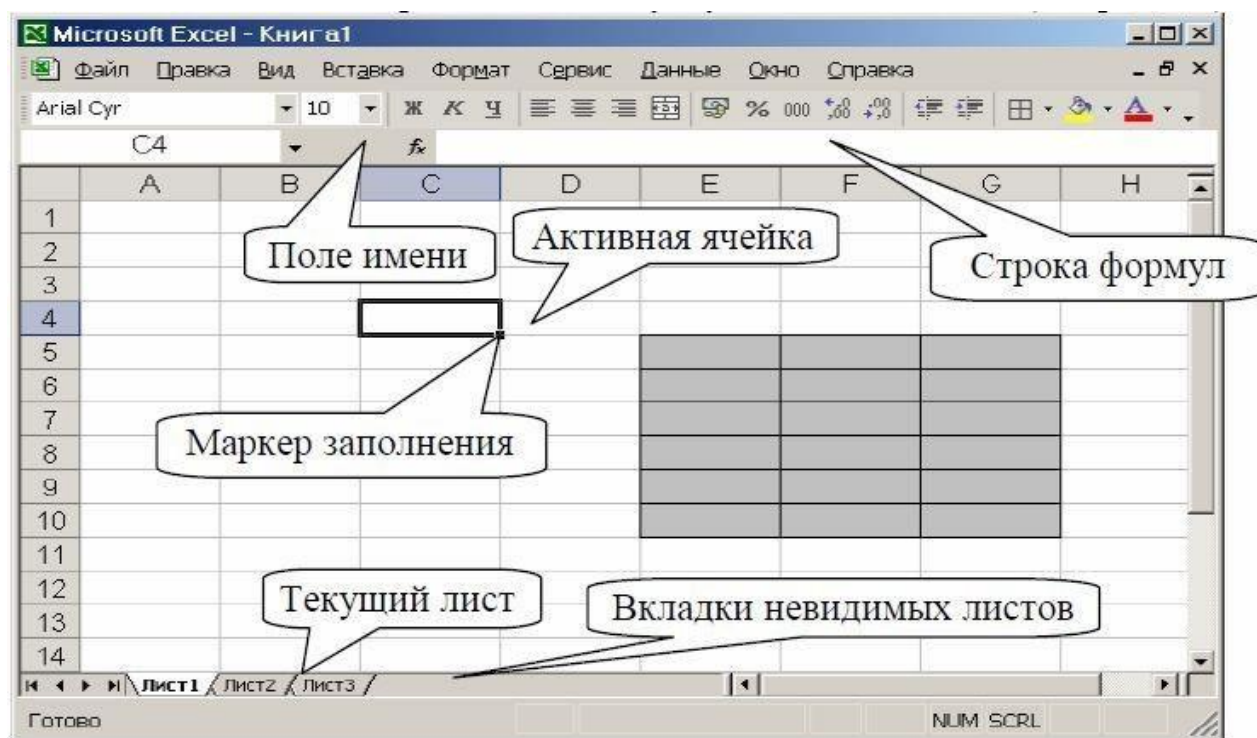


Рисунок 40 – Рабочая книга MS Excel

Рабочая книга состоит из рабочих листов (количество листов <255). Рабочий лист служит для формирования электронных таблиц, имеет прямоугольную форму и состоит из строк, столбцов и ячеек на их пересечении. Ячейка – элементарная единица данных. Элементом листа может быть также группа ячеек, называемых блоком. Ячейки предназначены для хранения различных значений. Ячейка похожа на переменную в математике: она имеет обозначение (адрес) и может иметь и менять значение. Количество ячеек во всех строках одинаково, и количество ячеек во все столбцах также одинаково. Ячейку однозначно определяет адрес ячейки при ссылке на ячейку.

Стандартные имена листов: Лист1, Лист2, ...

Имена столбцов – латинские буквы или их двухбуквенные сочетания (от A до IV).

Номера строк – от 1 до 65536.

Адрес ячейки состоит из имени столбца и номера строки – A5, B1, D16, ... (на рисунке 40 – C4).

Адрес блока – диапазон ячеек, заданный адресами конечных ячеек любой диагонали блока, разделенных двоеточием (E5:G10 или G10:E5 или G5:E10 или E10:G5 – для выделенного цветом блока на рисунке 40): перечень адресов ячеек, разделенных точкой с запятой (E5;F5;G5;E6;F6: ...).

Значением ячейки могут быть текст или число. Их можно набрать с клавиатуры в виде константы или вычислить по формуле, которая набирается в ячейке. Формула может содержать математические, логические, статистические и другие операции и функции. Внешний вид значения ячейки задается форматом ячейки.

Основная идея электронных таблиц: одни ячейки рабочего листа являются независимыми переменными, в которые нужно вводить константы с клавиатуры, а другие ячейки используются как зависимые переменные, в которые вводятся формулы, использующие значения первых ячеек. При изменении значения хотя бы одной независимой ячейки происходит автоматический пересчет значений по формулам, в которые эта ячейка входит прямо или косвенно через другие формулы. Любая ячейка может быть как зависимой, так и независимой. Это определяется наличием или отсутствием формулы в ней.

Правила работы в Excel: пользователь вводит исходные данные в независимые ячейки и формулы в зависимые. При необходимости устанавливается в ячейке нужный формат значения. Расчетные значения появляются в таблице в ячейках с формулами. Затем нужно проверить правильность расчетов, изменяя значения независимых ячеек и, при необходимости, исправить ошибки в формулах. Далее можно строить диаграмму по любым значениям таблицы (исходным и вычисленным). Подготовить таблицу к печати, используя форматирование – задание границ, цвета фона, параметров шрифта и т.д., напечатать или сохранить ее в виде файла на диске.

Ввод значения (константы или формулы) выполняется в активную ячейку, которая выделена курсором – рамочкой (на рисунке 40 – С4). Ее адрес помещается в поле имени. Вводимая информация одновременно отображается и в ячейке, и в строке формул (см. рисунок 40). Ее можно активизировать щелчком в ней левой кнопки мыши. Завершение ввода значения – нажатие клавиши Enter для запоминания значения или клавиши Esc – для отмены ввода.

Некоторые значения можно не набирать, а заполнять копированием или переносом значений других ячеек, а также пользуясь автозаполнением, например, значениями числовой арифметической прогрессии или значениями последовательных дат, месяцев, элементов автосписка и т.д. Для копирования формул, констант и автозаполнения используется маркер заполнения – точка в правом нижнем углу активной ячейки (см. рисунок 40).

Excel может работать в режимах:

Готово – можно перемещаться по таблице и использовать меню и пиктограммы.

Ввод – режим ввода данных в ячейку. Большинство команд меню недоступно.

Правка – режим редактирования введенных ранее данных.

Указать – режим выбора объекта (ячейки или области ячеек) для включения его в формулу. Объект выделяется бегущей пунктирной линией.

Текущий режим отображен в строке состояния окна Excel.

Основное преимущество электронных таблиц заключается в возможности помещать в них не только данные, но и формулы.

Правила создания формул:

1. Любая формула должна начинаться со знака «=». Вводимая формула отображается в строке формул, находящейся под панелями инструментов Excel (см. рисунок 40).

2. Формула может включать константы, знаки операций, функции, адреса ячеек.

3. В Excel допустимы следующие операторы:

- – смена знака;
- % – операция процента (применима к отдельному числу);
- ^ – операция возведения в степень;

*, / – умножение, деление;
 +, – – сложение, вычитание.

Операции выполняются слева направо в порядке их приоритетов, которые могут быть изменены круглыми скобками.

Преимущество формул заключается в том, что изменение значения ячейки, адрес которой используется в формуле, ведет к автоматическому пересчету этой формулы, что влечет пересчет формул других ячеек, использующих данную, и так далее. В итоге может обновиться вся таблица.

Для ссылки на ячейки используется абсолютная и относительная адресация. Оба способа указывают на одни и те же объекты – на ячейки. Их особенности проявляются при копировании.

Относительная адресация. Пусть в ячейке C3 (рисунок 41) записана формула=A1+B2. Если выполнить ее копирование в разные стороны, мы увидим, что она изменяется. При копировании по горизонтали изменяется имя столбца, при копировании по вертикали – номер строки, при копировании по диагонали – и то и другое. Так, при копировании в направлении стрелки формула перемещается на два столбца вправо. Поэтому адреса столбцов получили приращение 2 – вместо А – С, вместо В – D. Также формула перемещается на две строки вниз. Поэтому адреса строк также увеличились на 2. Это правило изменения формул при копировании достаточно часто удовлетворяет пользователя. Ему не приходится вручную создавать множество однотипных формул, которые должны отличаться только номерами ячеек. Но так бывает не всегда.

	A	B	C	D	E
1	5				
2	4	6			
3	5	8	=A1+B2	=B1+C2	=C1+D2
4			=A2+B3		
5			=A3+B4		=C3+D4
6					




Рисунок 41 – Особенности относительной адресации, проявляющиеся при копировании

Абсолютная адресация. Чтобы предотвратить автоматическое изменение адресов, перед «замораживаемой» координатой нужно поставить знак \$ и превратить ее в абсолютную ссылку. Знак \$ может быть установлен и только перед одной координатой (только перед номером столбца или только перед номером строки) или перед обоими координатами сразу. Например, формула =\$A\$1 не будет изменяться ни при каком копировании. В формуле =A\$1 при копировании не будет изменяться адрес строки, а в формуле =\$A1 не будет изменяться номер столбца. Знак \$ можно вводить непосредственно с клавиатуры, а можно с помощью клавиши F4 в режиме ввода формулы для текущей ссылки. Последовательное нажатие этой клавиши влечет поочередный ввод знака \$ перед элементами адреса: A1 => \$A\$1 => A\$1 => \$A1 => A1. Поэтому для получения желаемого результата следует последовательно нажимать F4.

Книга – это документ Excel, который при сохранении на диск становится файлом с именем, указанным пользователем, и стандартным расширением .xls. Это имя – имя книги.

Книга Excel состоит из листов с именами Лист1, Лист2, ... или с именами, заданными пользователем. Количество листов в книге и их последовательность можно менять. В рабочей книге может быть 255 листов. Лист может иметь 256 колонок и 65536 строк.

Активный или текущий лист – это верхний (видимый) лист, в котором находится указатель активной ячейки. Для каждого листа в нижней части окна Excel имеется ярлычок (рисунок 42). Активизирует лист щелчок левой кнопкой мыши на его ярлычке. При этом лист перемещается вверх и становится видимым, а ярлычок ярким. Щелчок правой кнопкой мыши на ярлычке вызывает контекстное меню для выполнения с листом перемещения, удаления, переименования и т.д.

В формулах могут использоваться данные:

- 1) одного (текущего) листа;
- 2) нескольких листов одной книги;
- 3) данные из другой книги, т.е. из другого файла.

Например, нужно сослаться на данные блока A6:C8 на Листе5. Тогда: A6:C8 – ссылка на блок из текущего листа, Лист5!A6:C8 – из любого листа данной книги.



Рисунок 42 – Ярлыки рабочих листов Excel

Операции с листами рабочей книги:

Выделение текущего листа – щелчок левой кнопкой мыши в верхнем левом углу рабочего листа.

Выделение листа – щелчки левой кнопки мыши по вкладкам нужных листов при нажатой клавише CTRL или SHIFT для смежных листов.

Удаление текущего листа – команда Правка/Удалить лист.

Удаление любого листа – щелчок правой кнопкой мыши на вкладке листа и затем команда Удалить из появившегося контекстного меню.

Вставка листа – команда Вставка/лист добавляет новый лист перед текущим листом. Вставлять листы можно путем копирования – команда Правка/Переименовать Скопировать лист... выводит на экран окно Переместить или скопировать. Для копирования нужно установить флажок Создать копию.

Переименование листа – щелчок правой кнопкой мыши на вкладке листа и затем команда Переименовать из появившегося контекстного меню. Затем в ярлыке листа ввести новое название – набор символов длиной до 31-го символа, кроме знаков * : / \ ? [] .

Перемещение – щелчок правой кнопкой мыши на вкладке листа и затем команда Переместить/Скопировать. На экран выдается окно Переместить или скопировать, в котором в списке Перед листом выбирается лист, перед которым размещается перемещаемый.

Приложение Б

ТАБЛИЦА РАСПРЕДЕЛЕНИЯ КРИТЕРИЯ ХИ-КВАДРАТ

Число степеней свободы	Уровень значимости					
	0,01	0,025	0,05	0,95	0,975	0,99
1	6,6	5,0	3,8	0,0039	0,00098	0,00016
2	9,2	7,4	6,0	0,103	0,051	0,020
3	11,3	9,4	7,8	0,352	0,216	0,115
4	13,3	11,1	9,5	0,711	0,484	0,297
5	15,1	12,8	11,1	1,15	0,831	0,554
6	16,8	14,4	12,6	1,64	1,24	0,872
7	18,5	16,0	14,1	2,17	1,69	1,24
8	20,1	17,5	15,5	2,73	2,18	1,65
9	21,7	19,0	16,9	3,33	2,70	2,09
10	23,2	20,5	18,3	3,94	3,25	2,56
11	24,7	21,9	19,7	4,57	3,82	3,05
12	26,2	23,3	21,0	5,23	4,40	3,57
13	27,7	24,7	22,4	5,89	5,01	4,11
14	29,1	26,1	23,7	6,57	5,63	4,66
15	30,6	27,5	25,0	7,26	6,26	5,23
16	32,0	28,8	26,3	7,96	6,91	5,81
17	33,4	30,2	27,6	8,67	7,56	6,41
18	34,8	31,5	28,9	9,39	8,23	7,01
19	36,2	32,9	30,1	10,1	8,91	7,63
20	37,6	34,2	31,4	10,9	9,59	8,26
21	38,9	35,5	32,7	11,6	10,3	8,90
22	40,3	36,8	33,9	12,3	11,0	9,54
23	41,6	38,1	35,2	13,1	11,7	10,2
24	43,0	39,4	36,4	13,8	12,4	10,9
25	44,3	40,6	37,7	14,6	13,1	1,5
26	45,6	41,9	38,9	15,4	13,8	12,2
27	47,0	43,2	40,1	16,2	14,6	12,9
28	48,3	44,5	41,3	16,9	15,3	13,6
29	49,6	45,7	42,6	17,7	16,0	14,3
30	50,9	47,0	43,8	18,5	16,8	15,0

Приложение В

ТАБЛИЦА РАСПРЕДЕЛЕНИЯ КРИТЕРИЯ СТЬЮДЕНТА

Число степеней свободы	Уровень значимости					
	0,10	0,05	0,02	0,01	0,002	0,001
1	6,31	12,7	31,82	63,7	318,3	637,0
2	2,92	4,30	6,97	9,92	22,33	31,6
3	2,35	3,18	4,54	5,84	10,22	12,9
4	2,13	2,78	3,75	4,60	7,17	8,61
5	2,01	2,57	3,37	4,03	5,89	6,86
6	1,94	2,45	3,14	3,71	5,21	5,96
7	1,89	2,36	3,00	3,50	4,79	5,40
8	1,86	2,31	2,90	3,36	4,50	5,04
9	1,83	2,26	2,82	3,25	4,30	4,78
10	1,81	2,23	2,76	3,17	4,14	4,59
11	1,80	2,20	2,72	3,11	4,03	4,44
12	1,78	2,18	2,68	3,05	3,93	4,32
13	1,77	2,16	2,65	3,01	3,85	4,22
14	1,76	2,14	2,62	2,98	3,79	4,14
15	1,75	2,13	2,60	2,95	3,73	4,07
16	1,75	2,12	2,58	2,92	3,69	4,01
17	1,74	2,11	2,57	2,90	3,65	3,96
18	1,73	2,10	2,55	2,88	3,61	3,92
19	1,73	2,09	2,54	2,86	3,58	3,88
20	1,73	2,09	2,53	2,85	3,55	3,85
21	1,72	2,08	2,52	2,83	3,53	3,82
22	1,72	2,07	2,51	2,82	3,51	3,79
23	1,71	2,07	2,50	2,81	4,49	3,77
24	1,71	2,06	2,49	2,80	3,47	3,74
25	1,71	2,06	2,49	2,79	3,45	3,72
26	1,71	2,06	2,48	2,78	3,44	3,71
27	1,71	2,05	2,47	2,77	3,42	3,69
28	1,70	2,05	2,46	2,76	3,40	3,66
29	1,70	2,05	2,46	2,76	3,40	3,66
30	1,70	2,04	2,46	2,75	3,39	3,65
40	1,68	2,02	2,42	2,70	3,31	3,55
60	1,67	2,00	2,39	2,66	3,23	3,46
120	1,66	1,98	2,36	2,62	3,17	3,37
∞	1,64	1,96	2,33	2,58	3,09	3,29

Приложение Г

Перечень вопросов для подготовки к зачету

- 1) Объективные факты
- 2) Научный факт
- 3) Фактуальное описание
- 4) Организационно-процедурные этапы исследования
- 5) Обработка данных
- 6) Описание данных
- 7) Обобщение данных
- 8) Анализ данных
- 9) Объяснение (интерпретация) данных
- 10) Эмпирическая проверка гипотез
- 11) Теоретическая проверка гипотез
- 12) Количественная и качественная обработка данных
- 13) Средняя арифметическая величина
- 14) Медиана
- 15) Меры вариабельности (колеблемости)
- 16) Интерквартильный размах
- 17) Размах вариации в ряду
- 18) Среднее линейное отклонение
- 19) Дисперсия
- 20) Стандартное отклонение (или среднее квадратическое отклонение)
- 21) Относительный размах вариации
- 22) Относительное линейное отклонение
- 23) Коэффициент вариации
- 24) Коэффициент осцилляции
- 25) Общая дисперсия
- 26) Внутригрупповая (частная) дисперсия
- 27) Межгрупповая дисперсия
- 28) Коэффициент детерминации
- 29) Эмпирическое корреляционное отношение
- 30) Условное распределение.
- 31) Совместное «поведение» двух признаков.
- 32) Таблица сопряженности.
- 33) Показатели таблицы сопряженности.
- 34) Маргинальные частоты.
- 35) Сравнение структуры условных распределений.

- 36) Критерий t-Стьюдента для одной выборки
- 37) Т-критерий Стьюдента для независимых выборок
- 38) Непараметрический критерий U-Манна-Уитни
- 39) Эмпирического значения критерия t-Стьюдента
- 40) Коэффициент корреляции Пирсона
- 41) Графический анализ простой линейной регрессии
- 42) Теоретическое уравнение регрессии
- 43) Диаграмма рассеяния
- 44) Основная цель регрессионного анализа
- 45) Линия регрессии
- 46) Построение теоретической линии регрессии
- 47) Критерий наименьших квадратов для нахождения прямой линии
- 48) Квадратическая ошибка оценки, проверка их типичности
- 49) Криволинейная регрессия
- 50) Составление таблицы экспериментальных данных
- 51) Построения приблизительного графика экспериментальных данных
- 52) Сравнение экспериментальной кривой с известными графиками функций
- 53) Выбор функциональной зависимости, наиболее подходящей под экспериментальную кривую
- 54) Подбор вида формулы для полученной экспериментальной зависимости и значения констант, входящих в нее