

Министерство сельского хозяйства РФ
федеральное государственное бюджетное образовательное учреждение
Высшего образования «Вологодская государственная
молочнохозяйственная академия им. Н.В. Верещагина»

Экономический факультет

Кафедра экономики и управления в АПК

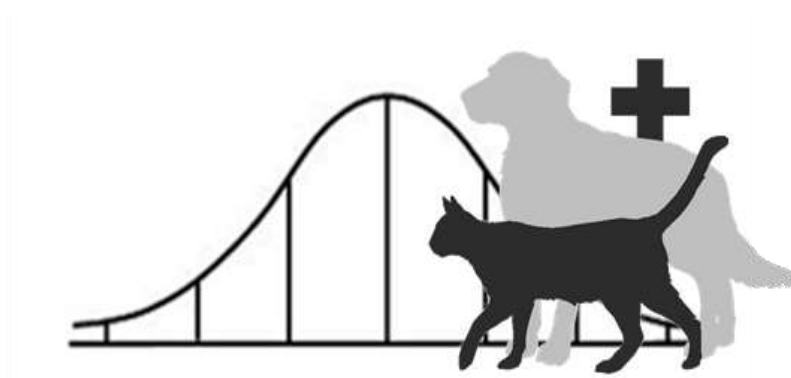
РАБОЧАЯ ТЕТРАДЬ

*для практических занятий по дисциплине
«Математическая биостатистика»
для студентов специальности 36.05.01 Ветеринария,
направления подготовки 36.03.01 Ветеринарно-санитарная экспертиза*

ФИО студента _____

Группа _____

Учебный год _____



Вологда–Молочное
2021

Рабочую тетрадь по курсу «Математическая биостатистика» для практических занятий студентов специальности 36.05.01 Ветеринария, направления подготовки 36.03.01 Ветеринарно-санитарная экспертиза составила:

канд. экон. наук, доцент кафедры экономики и управления в АПК
О.А. Шихова.

Рецензенты:

канд. экон. наук, доцент кафедры экономики и управления в АПК
И.Н. Шилова
канд. биолог. наук, доцент кафедры внутренних незаразных болезней,
хирургии и акушерства,
Л.Л. Фомина

Рабочая тетрадь по Математической биостатистике предназначена для студентов факультета ветеринарной медицины и биотехнологий специальности 36.05.01 Ветеринария, направления подготовки 36.03.01 Ветеринарно-санитарная экспертиза.

Рабочая тетрадь позволяет организовать аудиторную и самостоятельную работу студентов по закреплению теоретического материала и выработке практических навыков применения методов математической статистики к анализу медико-биологических данных. Содержит общие методические указания и инструктаж по работе с тетрадью, цель, нормы времени и методическое обеспечение для изучения каждой темы курса «Математическая биостатистика», задания для аудиторной работы и задания для самостоятельного применения базовых методов статистики в практике расчетов и анализа медико-биологических данных. Составлена в соответствии с учебным пособием по дисциплине «Математическая биостатистика» для студентов специальности для студентов специальности 36.05.01 Ветеринария, направления подготовки 36.03.01 Ветеринарно-санитарная экспертиза.

Рабочая тетрадь рассмотрена на заседании методической комиссии экономического факультета и рекомендована к изданию (протокол № 5 от 20.01.2021 г.).

© Шихова О.А., 2021

© ФГБОУ ВО Вологодская ГМХА, 2021

Общие методические указания

Дисциплина «Математическая биостатистика» - важная составляющая базовой части дисциплин учебного плана в системе подготовки студентов по специальности 36.05.01 Ветеринария, направления подготовки 36.03.01 Ветеринарно-санитарная экспертиза. Во время ее изучения рассматриваются основные принципы, показатели и методы математической статистики применительно к анализу медико-биологических данных; приемы осуществления статистических расчетов в MS Excel, табличный и графический методы представления результатов анализа. Данная дисциплина является основой для дальнейшего использования информационных технологий в процессе изучения других дисциплин, выполнения научной работы, а также использования в профессиональной деятельности.

Настоящая рабочая тетрадь разработана для студентов факультета ветеринарной медицины и биотехнологий специальности 36.05.01 Ветеринария, направления подготовки 36.03.01 Ветеринарно-санитарная экспертиза, чтобы организовать и методически обеспечить аудиторную и самостоятельную работу формированию и закреплению практических умений и навыков проведения статистического анализа в рамках изучения курса «Математическая биостатистика». По каждой теме курса сформулированы цель работы, приобретаемые обучающимися навыки и умения в результате выполнения представленных заданий для аудиторной и самостоятельной работы. Рабочая тетрадь предполагает использование другой учебно-методической литературы, в частности учебного пособия по дисциплине «Математическая биостатистика» [1].

Цели и задачи дисциплины, ее место в учебном процессе

Цель изучения дисциплины «Математическая биостатистика» состоит в реализации следующих направлений: подготовить выпускника, умеющего владеть средствами математико-статистического анализа медико-биологической информации; формирование основ статистического мышления; получение базовых знаний и формирование основных навыков по математической биостатистике, необходимых для решения задач, возникающих в практической деятельности врача-ветеринара и ветеринарно-санитарного эксперта.

Задачи дисциплины:

- 1) изучение теоретических основ изучаемой дисциплины, в частности иметь представление об основных методах статистического анализа экспериментальных данных;
- 2) изучение основ статистических методов представления, группировки и обработки материалов (результатов) медико-биологических исследований;
- 3) приобретение практических навыков по методам статистических исследований в биологии, вычислений важнейших статистических показателей и закономерностей, характеризующих совокупности биологических объектов для их эффективного применения в профессиональной деятельности.

В результате изучения дисциплины студенты:

- овладеют знаниями о возможностях и инструментах MS Excel для систематизации и статистического анализа опытных данных;
- научатся применять в практике статистических расчетов статистические и математические функции в MS Excel, пакет «Анализ данных»;
- освоят способы построения различных типов статистических диаграмм, используемых в анализе;
- овладеют навыками выполнения комплексного статистического анализа опытных данных в MS Excel, включающего систематизацию данных, их графическое представление, расчет обобщающих статистических показателей с помощью статистических и математических функций, проведение исследования взаимосвязей между явлениями с помощью инструмента «Регрессия» в пакете «Анализ данных».

Вводный инструктаж

Предлагаемая Рабочая тетрадь предназначена для того, чтобы помочь студентам факультета ветеринарной медицины и биотехнологий специальности 36.05.01 Ветеринария, направления подготовки 36.03.01 Ветеринарно-санитарная экспертиза изучить дисциплину «Математическая биостатистика».

Структура тетради соответствует структуре учебного пособия «Математическая биостатистика», а уровень заданий – требованиям, предъявляемым Государственным образовательным стандартом.

При работе с тетрадью по каждой теме следует:

1) внимательно ознакомиться с целью работы, требованиями в отношении приобретаемых навыков и умений, нормами времени на выполнение заданий, требованиями по оснащению рабочего места, предлагаемой литературой;

2) ознакомиться с заданием и исходными данными к нему;

3) все расчеты в соответствии с заданиями необходимо выполнять в MS Excel, для чего при изучении каждой темы создается предварительно перед практическим занятием расчетный файл в табличном процессоре MS Excel, при этом каждое задание выполняется на отдельном листе рабочей книги данного файла (название листа переименовывается в соответствии с номером выполняемого задания);

4) в файле необходимо накануне аудиторного занятия подготовить макеты расчетных таблиц в соответствии с их изображением и нумерацией в рабочей тетради к заданиям, предусмотренным к выполнению на аудиторном занятии по изучаемой теме;

5) для задний выполняемых в рамках самостоятельной (домашней) работы расчетные таблицы также необходимо выполнять в файле MS Excel;

6) все результаты расчетов из файла MS Excel переписываются в соответствующие таблицы рабочей тетради;

7) при записи решения задания следует заполнить предусмотренные в тетради «пустые» строки в тексте, в формулах, на рисунках или в таблицах;

8) при выполнении заданий и заполнении рабочей тетради рекомендуется ориентироваться на формулы и методические указания по выполнению расчетов, представленные в разделах «Основные понятия и формулы» и «Типовые примеры» соответствующих тем учебного пособия «Математическая биостатистика» [1];

9) в каждом задании, где это требуется, необходимо сформулировать выводы по полученным результатам расчетов;

10) аудиторные задания выполняются в Рабочей тетради;

11) самостоятельная работа предусматривает выполнение заданий как в Рабочей тетради, так и в виде текстовых документов в форме индивидуальной работы;

12) по требованию преподавателя студент должен предоставлять данную тетрадь на проверку, что позволяет осуществлять текущий контроль усвоения практического материала по изучаемой теме курса;

13) студент должен приносить тетрадь на каждое практическое аудиторное занятие, при этом иметь заранее подготовленный расчетный файл MS Excel с макетами расчетных таблиц.

Подготовка к практическим занятиям по математической биостатистике с использованием данной Рабочей тетради предполагает предварительное изучение студентами теоретического лекционного материала по каждой теме, поскольку использование тетради нацелено на закрепление практических навыков применения методов статистических расчетов.

Тема 1 Анализ медико-биологических данных на основе их группировки и графического представления

Наименование работы: проведение систематизации и группировки опытных данных, построение вариационных рядов, статистических диаграмм.

Цель работы: научиться выполнять систематизацию и группировку данных по одному признаку, строить дискретные и интервальные ряды распределения, представлять графически распределение наблюдений по исследуемому признаку, используя графики полигона, гистограммы и кумулятивной кривой.

Приобретаемые навыки и умения: проводить ранжирование и систематизацию исходных данных для дискретных и непрерывных количественных признаков, использовать инструмент MS Excel «Сортировка», «Мастер диаграмм», функции СУММ, МИН, МАКС, ЧАСТОТА.

Норма времени: аудиторная работа - 90 мин.; самостоятельная работа – 270 мин.

Оснащение рабочего места: рабочая тетрадь, ПК с табличным процессором MS Excel.

Литература: [1].

Аудиторная работа

Задание 1.1 (Построение дискретного вариационного ряда)

По результатам проведенного эксперимента (результаты выборочного обследования 24 овец мясошерстной породы «Горьковская» по росту в холке (см), представленные в таблице 1.1), требуется для изучаемого признака выполнить группировку его значений, построив дискретный вариационный ряд, рассчитать основные компоненты ряда, изобразить графически полигон и кумуляту распределения.

Таблица 1.1 – Исходные данные

Номер наблюдения	1	2	3	4	5	6	7	8	9	10	11	12
Рост овцы, см	66	68	67	69	66	71	70	65	67	70	68	70
Номер наблюдения	13	14	15	16	17	18	19	20	21	22	23	24
Рост овцы, см	69	71	69	66	72	70	72	68	65	69	70	71

Решение:

1) создайте файл MS Excel с именем «Группа_ФИО_Тема 1», в котором Лист 1 переименуйте «Задание 1.1»;

2) на данном листе задайте необходимые параметры форматирования текста в ячейках электронной таблицы: выделите лист целиком, используя сочетание клавиш **CTRL+A**, установите тип шрифта Times New Roman, 14 пт, формат отображения в ячейках – выравнивание «по центру» по вертикали и по горизонтали с переносом по словам;

3) подготовьте таблицу 1.1 в вертикальном варианте для удобства выполнения сортировки (ранжирования) наблюдений по значениям исследуемого признака «Рост овцы» (рис. 1);

Таблица 1.1 - Исходные данные	
Номер наблюдения	Рост овцы, см
1	66
2	68
3	67
4	69
5	66
6	71
7	70
8	65
9	67
10	70
11	68
12	70
13	69
14	71
15	69
16	66
17	72
18	70
19	72
20	68
21	65
22	69
23	70
24	71

Рисунок 1 – Пример вертикального оформления таблицы 1.1 с исходными данными в MS Excel

4) скопируйте таблицу 1.1 и вставьте справа от нее; переименуйте вставленную таблицу «Таблица 1.2 – Ранжированные данные» и в ней, используя инструмент «Сортировка», выполните ранжирование наблюдений по возрастанию значений роста овцы;

5) сгруппируйте наблюдения по повторяемости одинаковых значений роста овец (по возрастанию), определив абсолютные частоты – количество повторов одинаковых значений. Результаты группировки оформите в виде таблицы 1.3, заполнив вторую строчку таблицы, вычислив итоговый показатель – сумму абсолютных частот, используя функцию СУММ(Массив) в соответствии с формулой (1.1) из учебного пособия [1];

Таблица 1.3 – Распределение овец по росту в холке

Рост овцы, см (x_i)	65	66	67	68	69	70	71	72	Итого
Количество овец (абсолютная частота), голов (f_i)									
Доля овец (относительная частота) (d_i)									
Количество овец с ростом не более значения x_i (накопленная частота), голов (S_i)									

б) рассчитайте относительные и накопленные частоты для каждого значения признака (x_i), используя формулы (1.2) и (1.3) учебного пособия [1] (методика выполнения расчета показателей представлена в Примере 1.1 учебного пособия):

$$S_i = \quad ;$$

$$d_i = \text{————} = \text{————} ;$$

Результаты расчетов запишите в рабочую тетрадь в таблицу 1.3;

7) постройте две диаграммы: полигон распределения для абсолютных частот («**Рисунок 1.1 – Распределение овец по росту в холке**») и кумуляту распределения (кумулятивную кривую, «**Рисунок 1.2 – Кумулята распределения овец по росту в холке**») для накопленных частот, используя инструмент **Диаграмма** в разделе меню **Вставка**, выбирая тип диаграммы «Точечная с прямыми отрезками и маркерами» (примеры диаграмм представлены на рисунках 1.1 и 1.2 в Примере 1.1 учебного пособия).

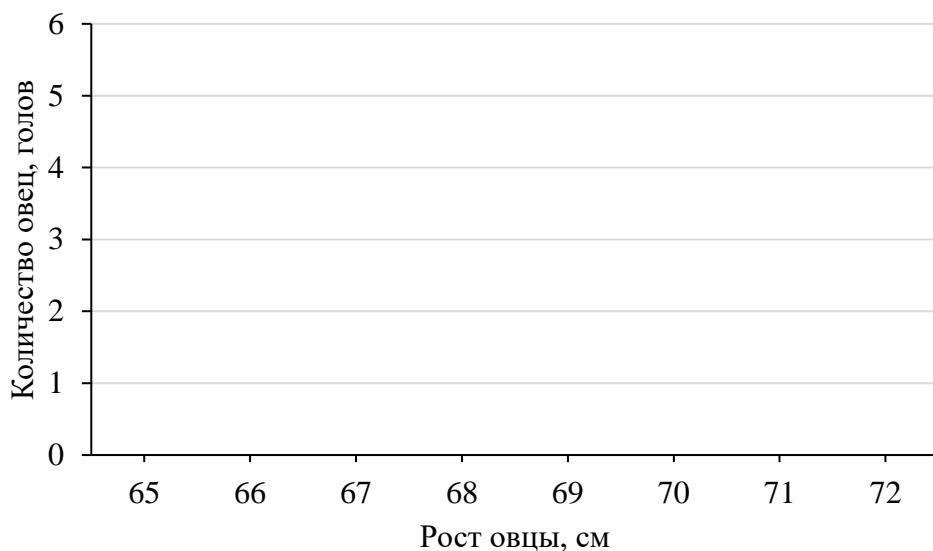


Рисунок 1.1 – Распределение овец по росту в холке

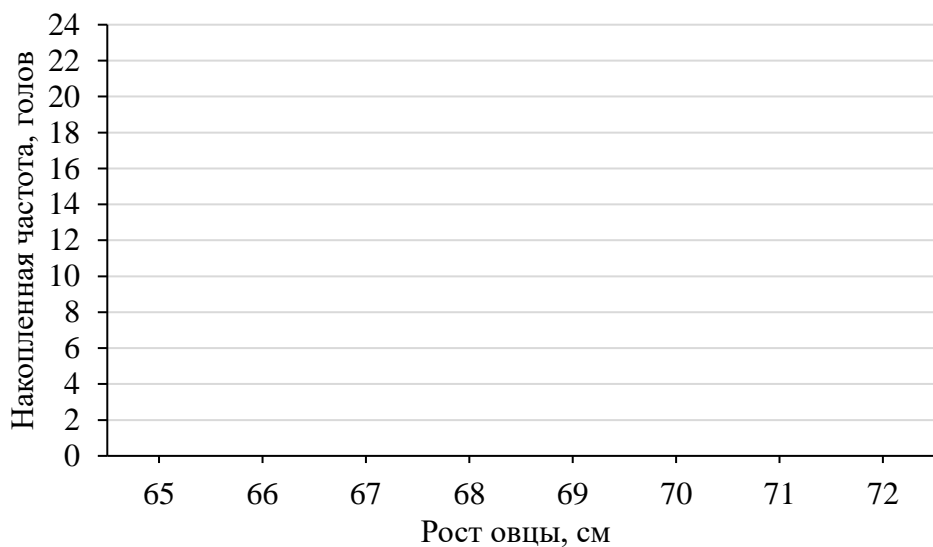


Рисунок 1.2 – Кумулята распределения овец по росту в холке

Задание 1.2 (Построение интервального вариационного ряда)

По результатам проведенного эксперимента (результаты выборочного обследования прироста объемов надоя молока в результате использования кормовой витаминной добавки у контрольной группы из 50 коров, представленные в таблице 1.4), требуется для изучаемого признака выполнить группировку его значений, построив интервальный вариационный ряд, рассчитать основные компоненты ряда, изобразить графически гистограмму, полигон и кумуляту распределения.

Таблица 1.4 – Прирост суточного удоя в результате использования витаминной добавки (литров)

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

Решение:

1) в файле MS Excel с именем «Группа_ФИО_Тема 1» Лист 2 переименуйте «Задание 1.2»;

2) на данном листе задайте необходимые параметры форматирования текста в ячейках электронной таблицы: выделите лист целиком, используя сочетание клавиш **CTRL+A**, установите тип шрифта Times New Roman, 14 пт, формат отображения в ячейках – выравнивание «по центру» по вертикали и по горизонтали с переносом по словам;

3) подготовьте таблицу 1.4 в MS Excel на листе «Задание 2» (рис. 2);

	A	B	C	D	E	F	G	H
1	Таблица 1.4 – Прирост суточного удоя в результате использования витаминной добавки (литров)						Таблица 1.5 - Вспомогательные характеристики	
2	2,3	3,4	2,8	3,6	1,8		Показатель	Значение
3	2,5	1,8	2,5	2,9	3,4		Минимум	
4	3,8	2,2	2,3	2,4	1,3		Максимум	
5	3,3	3,8	3,0	3,4	2,9		Число групп	
6	3,6	3,9	1,9	2,6	2,1		Размах вариации	
7	2,8	3,4	2,6	2,7	3,8		Длина интервала	
8	1,9	3,0	3,7	2,7	3,9			
9	2,9	1,5	3,7	2,2	2,0			
10	3,4	1,8	2,8	3,7	3,6			
11	2,1	3,3	2,6	2,2	3,4			
12								

Рисунок 2 - Пример оформления таблиц 1.4 и 1.5 в MS Excel

4) постройте таблицу 1.5 для расчета вспомогательных характеристик для выполнения равноинтервальной группировки и построения интервального ряда распределения (рис. 1.4);

5) используя статистические функции **МИН(Массив)** и **МАКС(Массив)** вычислите минимальное и максимальное значения прироста среднесуточного удоя в выборке (в качестве аргумента функций «Массив» выберите блок ячеек из таблицы 1.4 со значениями признака (A2:E11));

б) число групп рассчитайте по формуле Стерджесса (формула (1.4) из учебного пособия):

$$k = 1 + 3,322 \cdot \lg(n) = 1 + 3,322 \cdot \lg(50) = \dots \approx \dots;$$

При записи формулы в MS Excel для вычисления десятичного логарифма числа используйте математическую функцию **LOG10(Число)**. Полученное значение округлите до целого числа;

7) рассчитайте размах вариации по формуле:

$$R = x_{max} - x_{min} =$$

При записи формулы в MS Excel используйте адресные ссылки на ячейки с вычисленными ранее характеристиками **МИН(Массив)** и **МАКС(Массив)**;

8) вычислите длину интервала группировки, используя формулу (1.5) из учебного пособия:

$$h = \text{-----} = \text{-----} = \text{-----} ;$$

При записи формулы в MS Excel используйте адресные ссылки на ячейки с вычисленными ранее показателями R и k ;

9) определите границы интервалов, формируя их по возрастанию значений признака, начиная с наименьшего, прибавляя каждый раз значение длины интервала группировки h :

1 интервал: от ____ до ____ + ____ = ____;

2 интервал: от ____ до ____ + ____ = ____;

3 интервал: от ____ до ____ + ____ = ____;

4 интервал: от ____ до ____ + ____ = ____;

5 интервал: от ____ до ____ + ____ = ____;

6 интервал: от ____ до ____ + ____ = ____;

7 интервал: от ____ до ____ + ____ = ____;

10) вычислите для каждого интервала его абсолютную частоту (f_i) (количество коров) (табл. 1.6), используя статистическую функцию **ЧАСТОТА(Массив_данных;Массив_интервалов)**:

- подготовьте в MS Excel таблицу 1.6 согласно образцу;

Таблица 1.6 – Распределение коров по уровню прироста суточного удоя

Интервал значений прироста суточного удоя, литров		Количество коров, голов	Доля коров, %	Накопленная частота, голов
<i>Нижняя граница</i>	<i>Верхняя граница</i>	f_i	d_i	S_i
Итого				-

- впишите результаты расчета нижних и верхних границ интервалов (расчет этих значений можно и удобно выполнить непосредственно в этой таблице);

- выделите блок ячеек в столбце «Количество коров», в которые необходимо будет вывести значения абсолютных частот интервалов;

- нажимаем на клавиатуре « \Rightarrow » (оно отобразится в верхней ячейке блока ячеек, но в результате использования функции заполнятся значениями все выделенные ячейки), выбираем, используя «Вставить функцию», статистическую функцию **ЧАСТОТА**;

- в окне функции **ЧАСТОТА** в первой строке «**Массив данных**» делаем ссылку на блок ячеек с исходными данными в таблице 1.4; во второй строке «**Массив интервалов**» делаем ссылку на блок ячеек с верхними границами интервалов в таблице 1.6;

- **прежде, чем нажать «ОК» или Enter, следует обязательно нажать и удерживать две клавиши на клавиатуре Ctrl+Shift!** Только при этих нажатых клавишах нажимаем «ОК». Тогда заполнятся результатом применения функции все предварительно выделенные ячейки столбца «Количество коров». Если комбинацию клавиш не использовать, заполнится только одна ячейка (первая, в которой высвечивалось « \Rightarrow »);

11) вычислите для каждого интервала его относительную частоту (d_i) в процентах (долю коров с уровнем прироста удоя в пределах интервала в процентах от общего объема выборки) и накопленную частоту (S_i) (показывает количество коров в выборке, с уровнем прироста удоя не превышающем верхнюю границу интервала). Все результаты вычислений перепишите в таблицу 1.6 в рабочей тетради.

12) постройте две диаграммы: гистограмму распределения для абсолютных частот («**Рисунок 1.3 – Распределение коров по уровню прироста суточного удоя**») и кумуляту распределения (кумулятивную кривую, «**Рисунок 1.4 – Кумулята распределения коров по уровню прироста суточного удоя**») для накопленных частот, используя инструмент **Диаграмма** в разделе меню **Вставка**, выбирая тип диаграмм «Гистограмма» (для гистограммы) и «Точечная с прямыми отрезками и маркерами» (для кумуляты, в качестве абсцисс точек следует выбрать верхние границы интервалов, в качестве ординат – накопленные частоты соответствующих интервалов, начальная точка кумуляты имеет координаты: абсцисса – нижняя граница первого интервала, ордината равна нулю) (примеры диаграмм представлены на рисунках 1.3 и 1.4 в Примере 1.2 учебного пособия).

13) перенесите изображения диаграмм в рабочую тетрадь (рис. 1.3 и 1.4).

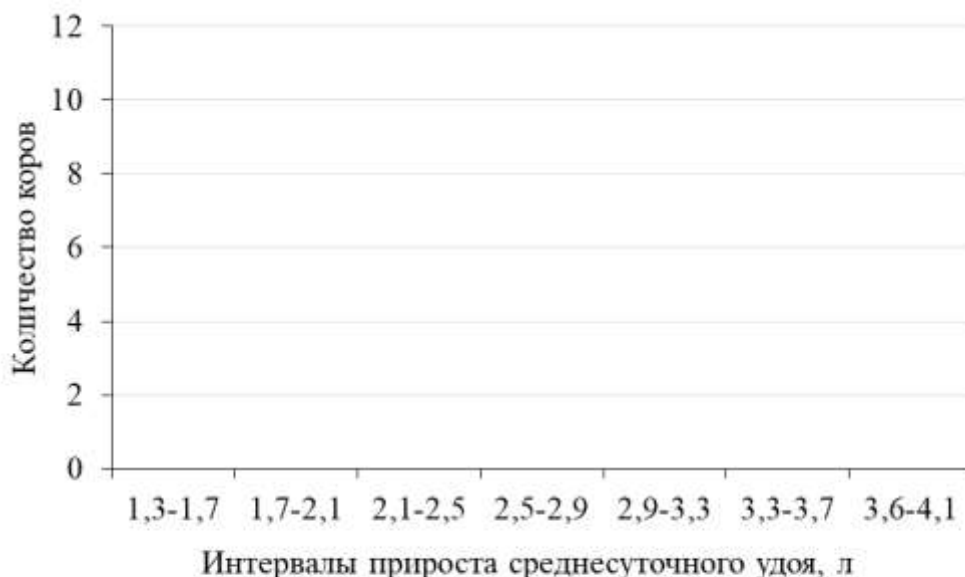


Рисунок 1.3 - Распределение коров по приросту среднесуточного удоя

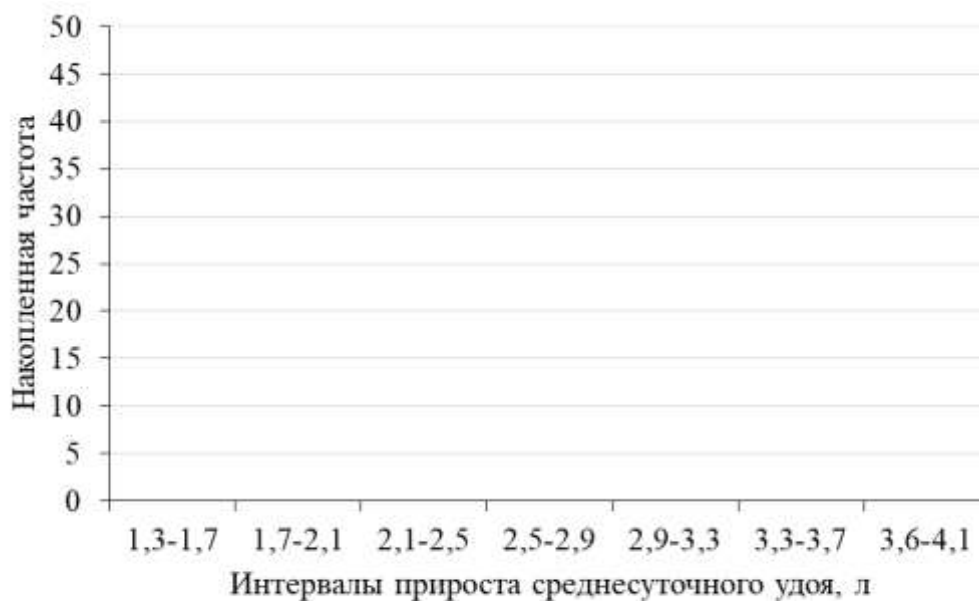


Рисунок 1.4 - Кумулята распределения коров по приросту среднесуточного удоя

Самостоятельная работа

Задание 1.3

По результатам проведенного эксперимента (результаты выборочного обследования 30 крольчих по продолжительности их беременности, представленные в таблице 1.7), требуется для изучаемого признака выполнить группировку его значений, построив дискретный вариационный ряд, рассчитать основные компоненты ряда, изобразить графически полигон и кумуляту распределения.

Таблица 1.7 – Исходные данные

Номер наблюдения	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Период беременности, дни	28	30	29	30	29	31	31	28	29	34	33	31	30	32	33
Номер наблюдения	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Период беременности, дни	29	32	31	32	30	28	30	31	32	30	29	33	33	34	31

Решение:

1) в файле MS Excel с именем «Группа_ФИО_Тема 1» добавьте Лист и переименуйте его в «Задание 1.3»;

2) на данном листе задайте необходимые параметры форматирования текста в ячейках электронной таблицы;

3) подготовьте таблицу 1.7 в вертикальном варианте для удобства выполнения сортировки (ранжирования) наблюдений по значениям исследуемого признака «Период беременности», переименуйте ее в «Таблица 1.7 – Исходные данные»;

4) непосредственно в таблице 1.7, используя инструмент «Сортировка», выполните ранжирование наблюдений по возрастанию значений признака «Период беременности»;

5) сгруппируйте наблюдения по повторяемости одинаковых значений признака (по возрастанию), определив абсолютные частоты – количество повторов одинаковых значений. Результаты группировки оформите в виде таблицы 1.8;

Таблица 1.8 – Распределение крольчих по продолжительности периода беременности

Период беременности, дней (x_i)									Итого
Количество крольчих (абсолютная частота), голов (f_i)									
Доля крольчих (относительная частота) (d_i)									
Количество крольчих с продолжительностью периода беременности не более значения x_i (накопленная частота), голов (S_i)									

1) рассчитайте относительные (d_i) и накопленные (S_i) частоты для каждого значения признака (x_i), используя формулы:

$$S_i = \quad ;$$

$$d_i = \frac{\quad}{\quad} = \frac{\quad}{\quad} ;$$

Результаты расчетов запишите в рабочую тетрадь в таблицу 1.8;

2) постройте две диаграммы: полигон распределения для абсолютных частот (рис. 1.5) и кумуляту распределения (кумулятивную кривую, рис.1.6) для накопленных частот.

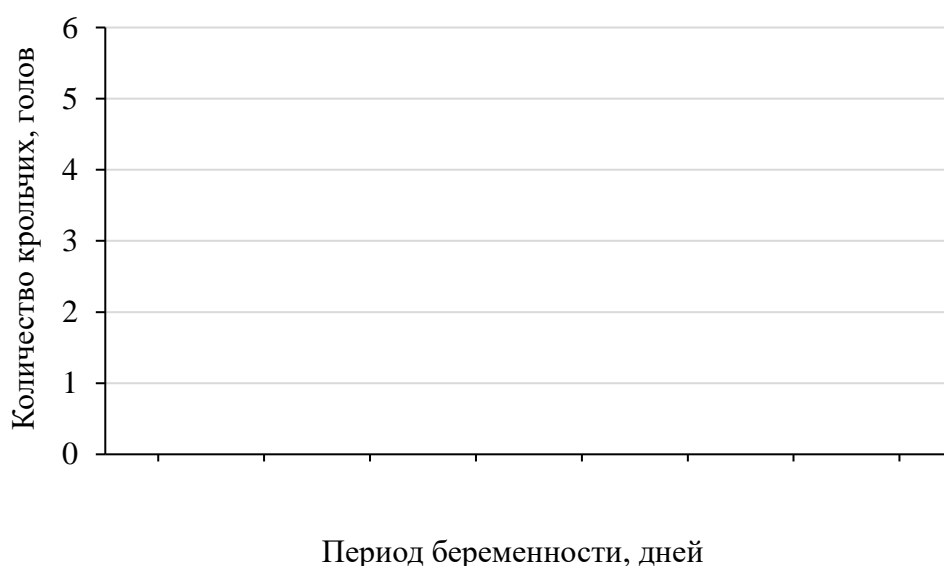


Рисунок 1.6 – Распределение крольчих по продолжительности периода беременности

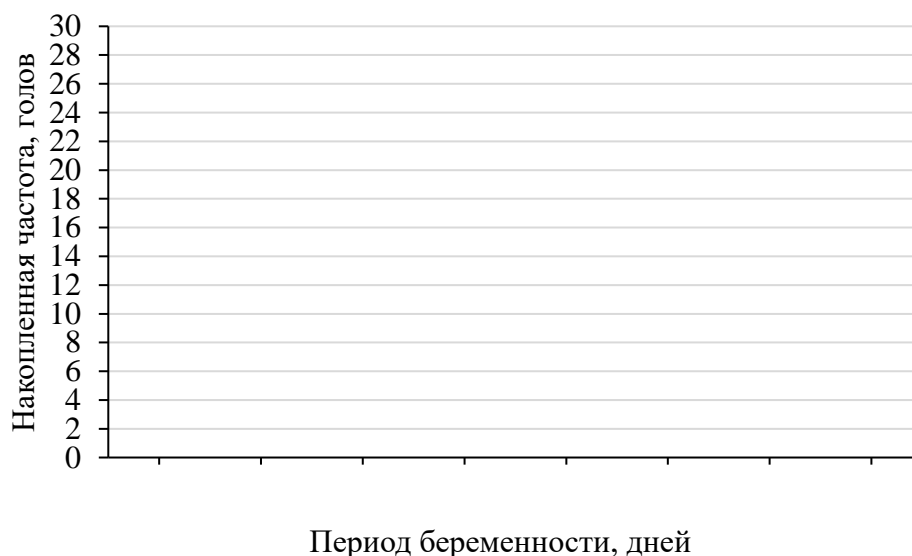


Рисунок 1.7 – Кумулята распределения крольчих по продолжительности периода беременности

Задание 1.4

По результатам проведенного исследования качества надоенного молока (40 проб молока на содержание жира, представленные в таблице 1.9), требуется для изучаемого признака выполнить группировку его значений, построив интервальный вариационный ряд, рассчитать основные компоненты ряда, изобразить графически гистограмму, полигон и кумуляту распределения.

Таблица 1.9 – Содержание жира в молоке (%)

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

Решение:

- 1) в файле MS Excel с именем «Группа_ФИО_Тема 1» добавьте Лист и переименуйте «Задание 1.4»;
- 2) на данном листе задайте необходимые параметры форматирования текста в ячейках электронной таблицы;
- 3) подготовьте таблицу 1.9 в MS Excel на листе «Задание 1.4»;
- 4) аналогично заданию 1.2 постройте таблицу 1.10 для расчета вспомогательных характеристик для выполнения равноинтервальной группировки и построения интервального ряда распределения;

Таблица 1.10 – Вспомогательные характеристики

Показатель	Значение
Минимум	
Максимум	
Число групп	
Размах вариации	
Длина интервала	

- 5) вычислите минимальное и максимальное значения процента жира в молоке в выборке;

б) число групп рассчитайте по формуле Стерджесса (формула (1.4) из учебного пособия):

$$k = 1 + 3,322 \cdot \lg(n) = 1 + 3,322 \cdot \lg(50) = \dots \approx \dots;$$

7) рассчитайте размах вариации по формуле:

$$R = x_{max} - x_{min} =$$

8) вычислите длину интервала группировки, используя формулу (1.5) из учебного пособия:

$$h = \text{-----} = \text{-----} = \quad ;$$

9) определите границы интервалов:

1 интервал: от ____ до ____ + ____ = ____;

2 интервал: от ____ до ____ + ____ = ____;

3 интервал: от ____ до ____ + ____ = ____;

4 интервал: от ____ до ____ + ____ = ____;

5 интервал: от ____ до ____ + ____ = ____;

6 интервал: от ____ до ____ + ____ = ____;

10) аналогично заданию 1.2 вычислите для каждого интервала его абсолютную частоту (f_i) (количество проб) (табл. 1.11), используя статистическую функцию **ЧАСТОТА(Массив_данных;Массив_интервалов)**;

Таблица 1.11 – Распределение проб по уровню процента жира в молоке

Интервал значений уровня жира в молоке, %		Количество проб	Доля проб, %	Накопленная частота
<i>Нижняя граница</i>	<i>Верхняя граница</i>	f_i	d_i	S_i
Итого				-

11) вычислите для каждого интервала его относительную частоту (d_i) в процентах и накопленную частоту (S_i). Все результаты вычислений перепишите в таблицу 1.11 в рабочей тетради.

12) постройте две диаграммы: гистограмму распределения для абсолютных частот и кумуляту распределения для накопленных частот, перенесите изображения диаграмм в рабочую тетрадь (рис. 1.5 и 1.6).

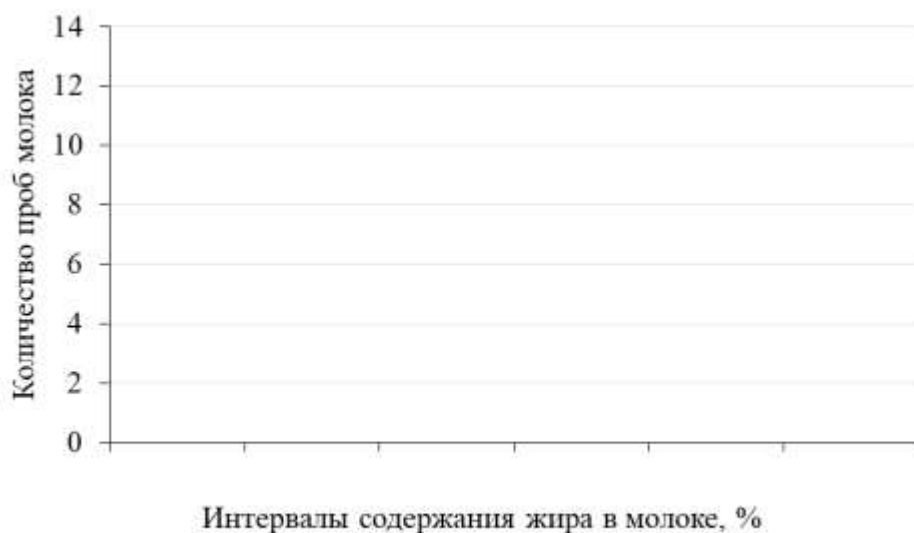


Рисунок 1.5 - Распределение проб по уровню жира в молоке

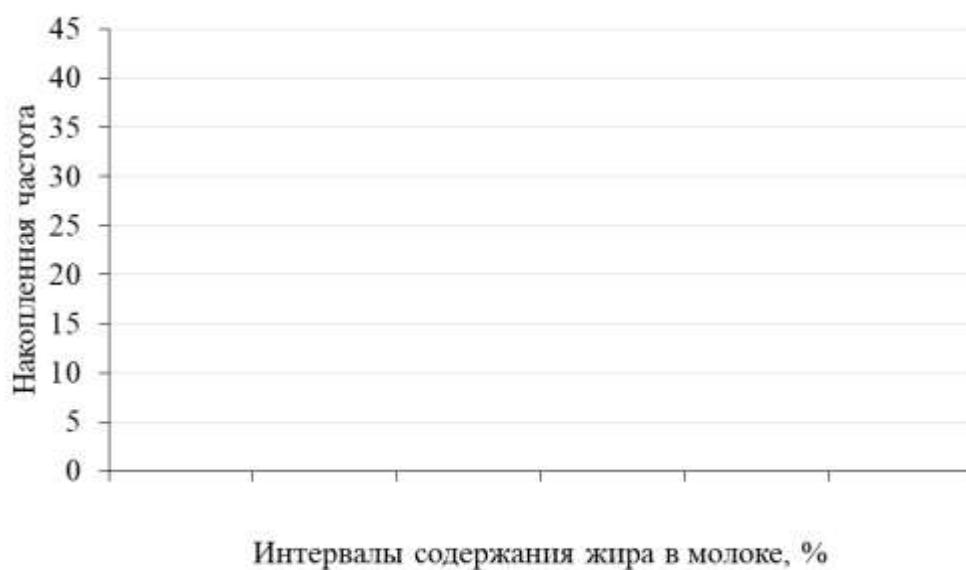


Рисунок 1.6 - Кумулята распределения проб молока по проценту жира

Тема 2 Оценка характеристик генеральной совокупности по выборке

Наименование работы: оценка величины основных характеристик генеральной совокупности по данным выборки: генеральной средней и генеральной доли единиц, обладающих изучаемым признаком.

Цель работы: научиться вычислять по не сгруппированным и сгруппированным данным числовые характеристики положения (среднее значение признака, показатели вариации), выполнять по результатам анализа выборки точечную и интервальную оценку генеральных значений средней и доли для случаев повторной и бесповторной случайных выборок.

Приобретаемые навыки и умения: расчет обобщающих числовых характеристик совокупности значений признака по не сгруппированным данным с использованием статистических формул и функций MS Excel СРЗНАЧ, ДИСПР, СТАНДОТКЛОНП; расчет обобщающих числовых характеристик совокупности значений признака по сгруппированным данным (дискретный и интервальный вариационный ряд); применение выборочного метода для оценки значения генеральной средней и доли (случаи повторного и бесповторного отбора).

Норма времени: аудиторная работа - 180 мин.; самостоятельная работа – 270 мин.

Оснащение рабочего места: рабочая тетрадь, ПК с табличным процессором MS Excel.

Литература: [1], [2].

Аудиторная работа

Задание 2.1

По данным, представленным в таблице 2.1, вычислите по совокупности обследованных коров среднее значение прироста среднесуточного удоя, охарактеризуйте степень вариации значений этого признака, вычислив показатели вариации. Для расчетов используйте в MS Excel статистические функции СРЗНАЧ, ДИСПР, СТАНДОТКЛОНП. Сделайте выводы.

Таблица 2.1 – Прирост суточного удоя в результате использования витаминной добавки (литров)

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

Решение:

1) создайте файл MS Excel с именем «Группа_ФИО_Тема 2», в котором Лист 1 переименуйте «Задание 2.1»;

2) на данном листе задайте необходимые параметры форматирования текста в ячейках электронной таблицы: выделите лист целиком, используя сочетание клавиш **CTRL+A**, установите тип шрифта Times New Roman, 14 пт, формат отображения в ячейках – выравнивание «по центру» по вертикали и по горизонтали с переносом по словам;

3) подготовьте на этом листе таблицу 2.1 с исходными данными;

4) для выполнения расчетов числовых характеристик прироста суточного удоя с применением статистических функций подготовьте таблицу 2.2;

Таблица 2.2 – Числовые характеристики прироста суточного удоя коров

Показатель, единицы измерения	Символ	Значение
Среднее значение, литров	x_{cp}	
Размах вариации, литров	R	
Дисперсия	S^2	
Стандартное отклонение, литров	S	
Коэффициент вариации, %	V	

- вычислите с точностью до 0,01 среднее значение признака (x_{cp}), используя функцию **СРЗНАЧ(Массив)**, при этом в качестве **Массива** выделите блок ячеек со значениями признака в исходных данных (табл. 2.1);

- вычислите размах вариации (R) по формуле, используя при записи формулы функции **МИН(Массив)** и **МАКС(Массив)**;

- вычислите с точностью до 0,001 показатель дисперсии (S^2), используя статистическую функцию, возвращающую значение неисправленной выборочной дисперсии, т.е. дисперсии по генеральной совокупности (**ДИСПР(Массив)**), при этом в качестве **Массива** выделите блок ячеек со значениями признака в исходных данных (табл. 2.1);

- вычислите с точностью до 0,001 показатель стандартного отклонения (S), используя статистическую функцию, возвращающую значение неисправленного выборочного стандартного отклонения, т.е. стандартного отклонения по генеральной совокупности (**СТАНДОТКЛОНП (Массив)**), при этом в качестве **Массива** выделите блок ячеек со значениями признака в исходных данных (табл. 2.1);

- вычислите значение коэффициента вариации, по формуле:

$$V = \frac{S}{x_{cp}} \cdot 100 = \frac{S}{x_{cp}} \cdot 100 = \quad \%;$$

При выполнении расчетов используйте адресные ссылки на ячейки с ранее вычисленными значениями показателей в этой таблице.

5) по результатам расчетов сформулируйте вывод в соответствии со следующим шаблоном (заполните пропуски):


Данные таблицы 2.2 показывают, что амплитуда изменения значений величины суточного прироста удоя коров в результате применения витаминной добавки составляет _____ литров. Величина прироста суточного удоя в выборке из 50 коров составила в среднем _____ ± _____ литров молока. Значение коэффициента вариации, равное _____ %, указывает на _____ степень вариации значений прироста удоя в исследуемой выборке и _____ совокупности наблюдений в отношении изучаемого признака.

Задание 2.2

По данным о приросте среднесуточного удоя коров вычислите по совокупности обследованных коров среднее значение прироста удоя, охарактеризуйте степень вариации значений этого признака, вычислив показатели вариации для интервального ряда распределения значений этого признака, полученного при выполнении задания 1.2, используя для расчетов сгруппированные данные и расчетную таблицу (табл. 2.3). Сделайте выводы. Сравните полученные результаты расчетов числовых характеристик изучаемого признака с результатами задания 2.1. Сделайте вывод о том, как влияет форма представления исходных данных о признаке на точность результатов оценки его основных статистических характеристик.

Таблица 2.3 – Расчет числовых характеристик величины прироста среднесуточного удоя

Интервал значений прироста среднесуточного удоя, литров	Количество коров (f_i)	Середина интервала (x_i)	$x_i \cdot f_i$	$(x_i - \tilde{x})^2 \cdot f_i$
Итого				

 **Рекомендация:** при записи символьных выражений в MS Excel вместо значка *выборочной средней* \tilde{x} используйте символ x_{cp} .

Решение:

- 1) в файле MS Excel «Группа_ФИО_Тема 2» Лист 2 переименуйте «Задание 2.2»;
- 2) на данном листе задайте необходимые параметры форматирования текста в ячейках электронной таблицы;
- 3) подготовьте на этом листе таблицу 2.2, заполнив первый и второй столбцы таблицы в соответствии с результатами группировки, полученными в задании 1.2 (табл. 1.6);
- 4) для выполнения расчетов числовых характеристик прироста суточного удоя с применением статистических формул подготовьте таблицу 2.4;

Таблица 2.4 – Числовые характеристики прироста суточного удоя коров

Показатель, единицы измерения	Символ	Значение
Среднее значение, литров	x_{cp}	
Размах вариации, литров	R	
Дисперсия	S^2	
Стандартное отклонение, литров	S	
Коэффициент вариации, %	V	

- предварительно запишите формулы для расчета перечисленных числовых характеристик:

$\tilde{x} = \text{-----} = \text{-----} =$

$R =$

$S^2 = \text{-----} = \text{-----} =$

$S =$

$V = \text{-----} \cdot 100 = \text{-----} \cdot 100 = \quad \%$

- в таблице 2.4 вычислите середину каждого интервала как среднее между значениями нижней и верхней границами этого интервала;
- вычислите результаты произведений ($x_i \cdot f_i$) в каждой строке таблицы 2.4, в итоговой строке, используя функцию **СУММ (Массив)** определите сумму таких произведений в этом столбце, подставьте полученное значение в формулу для расчета *среднего значения* (\bar{x}), а также вычислите и подставьте значение суммы частот интервалов (f_i); в файле MS Excel в таблице 2.4 вычислите среднее значение (точность до 0,01), при этом в записи функции используйте адресные ссылки на соответствующие ячейки из итоговой строки таблицы 2.3, результат расчета запишите в рабочую тетрадь в таблицу 2.4 и в конце соответствующей формулы (укажите единицы измерения);
- вычислите по формуле *размах вариации* (R), взяв в качестве минимального значения признака нижнюю границу первого интервала, в качестве максимального значения признака – верхнюю границу последнего интервала в таблице 2.3, результат расчета запишите в рабочую тетрадь в виде выражения (укажите единицы измерения) и в таблице 2.4;
- вычислите результаты выражений $((x_i - \bar{x})^2 \cdot f_i)$ в каждой строке таблицы 2.4, в итоговой строке, используя функцию **СУММ (Массив)** определите сумму результатов таких выражений в этом столбце, подставьте полученное значение в формулу для расчета *дисперсии* (S^2), а также вычислите и подставьте значение суммы частот интервалов (f_i); в файле MS Excel в таблице 2.4 вычислите дисперсию (точность до 0,001), при этом в записи функции используйте адресные ссылки на соответствующие ячейки из итоговой строки таблицы 2.3, результат расчета запишите в рабочую тетрадь в таблицу 2.4 и в конце соответствующей формулы (у показателя дисперсии единицы измерения не пишутся);
- вычислите с точностью до 0,001 показатель стандартного отклонения (S), используя соответствующую формулу (в MS Excel для вычисления арифметического квадратного корня используйте функцию **КОРЕНЬ(Число)**), результат расчета запишите в рабочую тетрадь в виде выражения (укажите единицы измерения) и в таблице 2.4;
- вычислите значение коэффициента вариации, по соответствующей формуле (при выполнении расчетов в файле используйте адресные ссылки на ячейки с ранее вычисленными значениями показателей в таблице 2.4), результат расчета запишите в рабочую тетрадь в виде выражения и в таблице 2.4.

5) по результатам расчетов сформулируйте вывод в соответствии со следующим шаблоном (заполните пропуски):

Данные таблицы 2.4 показывают, что амплитуда изменения значений величины суточного прироста удоя коров в результате применения витаминной добавки составляет ____ литров. Величина прироста суточного удоя в выборке из 50 коров составила в среднем _____ ± _____ литров молока. Значение коэффициента вариации, равное _____ %, указывает на _____ степень вариации значений прироста удоя в исследуемой выборке и _____ совокупности наблюдений в отношении изучаемого признака.

Сделайте вывод, сравнив результаты расчетов с заданием 2.1.

При сравнении результатов расчета числовых характеристик величины прироста удоя коров, полученных в таблицах 2.2 и 2.4, можно сделать вывод, что значения показателей, вычисленные по не сгруппированным и сгруппированным данным _____ (отличаются/не отличаются). Это объясняется тем, что _____.

Задание 2.3

По результатам расчетов, полученным в задании 2.1, с вероятностью 0,997 выполните оценку среднего ожидаемого прироста среднесуточного удоя у коров в результате применения витаминной добавки и доли случаев прироста удоя на величину свыше 3 литров.

Решение:

1) выпишите из задания 2.1 исходные данные, полученные по результатам анализа:

Исходные данные по выборке:

- объем выборки: $n =$ _____ ;
- выборочная средняя: $\tilde{x} =$ _____ ;
- выборочная дисперсия: $S^2 =$ _____ ;
- количество коров в выборке с приростом удоя свыше 3 литров: $m =$ _____ ;
- доля коров в выборке с приростом удоя свыше 3 литров: $W = \frac{\quad}{\quad} =$ _____ ;
- доверительная вероятность (надежность) оценки: $P =$ _____ ;
- коэффициент доверия: $t =$ _____ .

2) запишите формулы, по которым необходимо будет выполнить вычисления (в учебном пособии [1] таблица 2.2):

Формулы:

- стандартная ошибка выборки для оценки генеральной средней значения признака:

$$\mu_{\bar{x}} = \quad ;$$

- предельная ошибка выборки для оценки генеральной средней значения признака:

$$\Delta_{\bar{x}} = \quad ;$$

- доверительный интервал оценки генеральной средней значения признака:

$$\underline{\quad} \leq \bar{x} \leq \underline{\quad} ;$$

- стандартная ошибка выборки для оценки генеральной доли:

$$\mu_p = \quad ;$$

- предельная ошибка выборки для оценки генеральной доли:

$$\Delta_p = \quad ;$$

- доверительный интервал оценки генеральной доли:

$$\underline{\quad} \leq p \leq \underline{\quad} .$$

3) выполните расчеты по формулам, подставляя исходные данные:

Расчеты:

1) оценка генеральной средней по выборке:

- стандартная ошибка выборки для оценки генеральной средней значения признака:

$$\mu_{\tilde{x}} =$$

- предельная ошибка выборки для оценки генеральной средней значения признака:

$$\Delta_{\bar{x}} =$$

- доверительный интервал оценки генеральной средней значения признака:

$$\underline{\hspace{2cm}} \leq \bar{x} \leq \underline{\hspace{2cm}} ;$$
$$\underline{\hspace{2cm}} \leq \bar{x} \leq \underline{\hspace{2cm}} .$$

2) оценка генеральной доли по выборке:

- стандартная ошибка выборки для оценки генеральной доли:

$$\mu_p =$$

- предельная ошибка выборки для оценки генеральной доли:

$$\Delta_p =$$

- доверительный интервал оценки генеральной доли:

$$\underline{\hspace{2cm}} \leq p \leq \underline{\hspace{2cm}} ;$$
$$\underline{\hspace{2cm}} \leq p \leq \underline{\hspace{2cm}} .$$

4) запишите вывод по результатам расчетов по шаблону (заполните пропуски):

С надежностью _____ можно утверждать, что в результате применения витаминной добавки величина среднего прироста удоя коров во всем стаде составит от _____ до _____ литров. При этом доля коров в стаде с величиной прироста удоя свыше 3 литров составит от _____% до _____%.

Задание 2.4

По исходным данным задания 1.4 (табл. 1.9), с вероятностью 0,954 дайте оценку среднего процента жира в молоке и доли коров в дойном стаде, дающих жирное молоко (4,5% жира и более), зная, что пробы молока были взяты у 15% коров этого стада.

Решение:

1) по исходным данным задания 1.4 (табл. 1.9), используя статистические функции, вычислите в MS Excel показатели выборочной средней и дисперсии процента жира в молоке, заполните раздел с исходными данными по выборке:

Исходные данные по выборке:

- объем выборки: $n =$ _____ ;
- выборочная средняя: $\bar{x} =$ _____ ;
- выборочная дисперсия: $S^2 =$ _____ ;
- доля отбора: $q =$ _____ ;
- количество проб молока в выборке с жирностью 4,5% и выше: $m =$ _____ ;
- доля проб молока в выборке с жирностью 4,5% и выше: $W = \frac{\hspace{1cm}}{\hspace{1cm}} =$ _____ ;
- доверительная вероятность (надежность) оценки: $P =$ _____ ;
- коэффициент доверия: $t =$ _____ .

2) запишите формулы, по которым необходимо будет выполнить вычисления (в учебном пособии [1] таблица 2.2):

Формулы:

- стандартная ошибка выборки для оценки генеральной средней значения признака:

$$\mu_{\bar{x}} = \quad ;$$

- предельная ошибка выборки для оценки генеральной средней значения признака:

$$\Delta_{\bar{x}} = \quad ;$$

- доверительный интервал оценки генеральной средней значения признака:

$$\underline{\hspace{2cm}} \leq \bar{x} \leq \underline{\hspace{2cm}} ;$$

- стандартная ошибка выборки для оценки генеральной доли:

$$\mu_p = \quad ;$$

- предельная ошибка выборки для оценки генеральной доли:

$$\Delta_p = \quad ;$$

- доверительный интервал оценки генеральной доли:

$$\underline{\hspace{2cm}} \leq p \leq \underline{\hspace{2cm}} .$$

3) выполните расчеты по формулам, подставляя исходные данные:

Расчеты:

1) оценка генеральной средней по выборке:

- стандартная ошибка выборки для оценки генеральной средней значения признака:

$$\mu_{\bar{x}} =$$

- предельная ошибка выборки для оценки генеральной средней значения признака:

$$\Delta_{\bar{x}} =$$

- доверительный интервал оценки генеральной средней значения признака:

$$\underline{\hspace{2cm}} \leq \bar{x} \leq \underline{\hspace{2cm}} ;$$
$$\underline{\hspace{2cm}} \leq \bar{x} \leq \underline{\hspace{2cm}} .$$

2) оценка генеральной доли по выборке:

- стандартная ошибка выборки для оценки генеральной доли:

$$\mu_p =$$

- предельная ошибка выборки для оценки генеральной доли:

$$\Delta_p =$$

- доверительный интервал оценки генеральной доли:

$$\underline{\hspace{2cm}} \leq p \leq \underline{\hspace{2cm}} ;$$
$$\underline{\hspace{2cm}} \leq p \leq \underline{\hspace{2cm}} .$$

4) запишите вывод по результатам расчетов по шаблону (заполните пропуски):

С надежностью _____ можно утверждать, что в дойном стаде средняя жирность молока у коров составляет от _____ до _____ процентов. При этом доля коров в стаде, дающих высокожирное молоко (с жирностью 4,5% и выше) составляет от _____% до _____%.

Задание 2.5

По результатам выборочной статистической разработки 25% карточек профилактического осмотра собак, наблюдавшихся в ветеринарной клинике в течение года, выявлено 18 случаев клещевой инвазии. С вероятностью 0,954 оцените долю случаев выявления данного заболевания у собак, зарегистрированных в клинике, учитывая, что на учете числится 328 собак.

Решение:

Исходные данные по выборке:

- объем выборки: $n =$ _____ ;
- доля отбора: $q =$ _____ ;
- объем генеральной совокупности: $N =$ _____ ;
- количество случаев заболевания в выборке: $m =$ _____ ;
- доля случаев заболевания в выборке: $W = \frac{\quad}{\quad} =$ _____ ;
- доверительная вероятность (надежность) оценки: $P =$ _____ ;
- коэффициент доверия: $t =$ _____ .

Формулы:

- стандартная ошибка выборки для оценки генеральной доли:

$$\mu_p = \quad ;$$

- предельная ошибка выборки для оценки генеральной доли:

$$\Delta_p = \quad ;$$

- доверительный интервал оценки генеральной доли:

$$\underline{\hspace{2cm}} \leq p \leq \underline{\hspace{2cm}} .$$

Расчеты:

- стандартная ошибка выборки для оценки генеральной доли:

$$\mu_p =$$

- предельная ошибка выборки для оценки генеральной доли:

$$\Delta_p =$$

- доверительный интервал оценки генеральной доли:

$$\underline{\hspace{2cm}} \leq p \leq \underline{\hspace{2cm}} ;$$
$$\underline{\hspace{2cm}} \leq p \leq \underline{\hspace{2cm}} .$$

запишите вывод по результатам расчетов по шаблону (заполните пропуски):

С надежностью _____ можно утверждать, что доля случаев выявления клещевой инвазии у всех собак, зарегистрированных в клинике, составляет от _____ до _____ процентов.

Задание 2.6

По исходным данным задания 1.1 (результаты выборочного обследования 24 овец мясошерстной породы «Горьковская» по росту в холке (см), представленные в таблице 1.1), дайте интервальную оценку среднего роста овец данной породы с надежностью 0,997.

Решение:

по исходным данным задания 1.1 (табл. 1.1), используя статистические функции, вычислите в MS Excel показатели выборочной средней и дисперсии роста овцы в холке, заполните раздел с исходными данными по выборке:

Исходные данные по выборке:

- объем выборки: $n =$ _____ ;
- выборочная средняя: $\tilde{x} =$ _____ ;
- выборочная дисперсия: $S^2 =$ _____ ;
- доверительная вероятность (надежность) оценки: $P =$ _____ ;
- коэффициент доверия: $t =$ _____ .

Формулы:

- стандартная ошибка выборки для оценки генеральной средней значения признака:

$$\mu_{\bar{x}} = \underline{\hspace{2cm}} ;$$

- предельная ошибка выборки для оценки генеральной средней значения признака:

$$\Delta_{\bar{x}} = \underline{\hspace{2cm}} ;$$

- доверительный интервал оценки генеральной средней значения признака:

$$\underline{\hspace{2cm}} \leq \bar{x} \leq \underline{\hspace{2cm}} ;$$

Расчеты:

- стандартная ошибка выборки для оценки генеральной средней значения признака:

$$\mu_{\bar{x}} =$$

- предельная ошибка выборки для оценки генеральной средней значения признака:

$$\Delta_{\bar{x}} =$$

- доверительный интервал оценки генеральной средней значения признака:

$$\underline{\hspace{2cm}} \leq \bar{x} \leq \underline{\hspace{2cm}} ;$$

$$\underline{\hspace{2cm}} \leq \bar{x} \leq \underline{\hspace{2cm}} .$$

запишите вывод по результатам расчетов по шаблону (заполните пропуски):

С надежностью _____ можно утверждать, что средняя величина роста в холке у овец всего стада составляет от _____ до _____ сантиметров.

Задание 2.7

По исходным данным задания 1.3 (результаты выборочного обследования 30 крольчих по продолжительности их беременности, представленные в таблице 1.7), требуется с вероятностью 0,997 оценить среднюю продолжительность беременности крольчих и долю самок с беременностью менее 30 дней.

Решение:

по исходным данным задания 1.3 (табл. 1.7), используя статистические функции, вычислите в MS Excel показатели выборочной средней и дисперсии величины продолжительности беременности крольчих, заполните раздел с исходными данными по выборке:

Исходные данные по выборке:

- объем выборки: $n =$ _____ ;
- выборочная средняя: $\bar{x} =$ _____ ;
- выборочная дисперсия: $S^2 =$ _____ ;
- количество крольчих в выборке с беременностью менее 30 дней: $m =$ _____ ;
- доля крольчих в выборке с беременностью менее 30 дней: $W = \frac{\quad}{\quad} =$ _____ ;
- доверительная вероятность (надежность) оценки: $P =$ _____ ;
- коэффициент доверия: $t =$ _____ .

Формулы:

- стандартная ошибка выборки для оценки генеральной средней значения признака:

$$\mu_{\bar{x}} = \quad ;$$

- предельная ошибка выборки для оценки генеральной средней значения признака:

$$\Delta_{\bar{x}} = \quad ;$$

- доверительный интервал оценки генеральной средней значения признака:

$$\underline{\hspace{2cm}} \leq \bar{x} \leq \underline{\hspace{2cm}} ;$$

- стандартная ошибка выборки для оценки генеральной доли:

$$\mu_p = \quad ;$$

- предельная ошибка выборки для оценки генеральной доли:

$$\Delta_p = \quad ;$$

- доверительный интервал оценки генеральной доли:

$$\underline{\hspace{2cm}} \leq p \leq \underline{\hspace{2cm}} .$$

Расчеты:

1) оценка генеральной средней по выборке:

- стандартная ошибка выборки для оценки генеральной средней значения признака:

$$\mu_{\bar{x}} =$$

- предельная ошибка выборки для оценки генеральной средней значения признака:

$$\Delta_{\bar{x}} =$$

- доверительный интервал оценки генеральной средней значения признака:

$$\underline{\hspace{2cm}} \leq \bar{x} \leq \underline{\hspace{2cm}} ;$$

$$\underline{\hspace{2cm}} \leq \bar{x} \leq \underline{\hspace{2cm}} .$$

2) оценка генеральной доли по выборке:

- стандартная ошибка выборки для оценки генеральной доли:

$$\mu_p =$$

- предельная ошибка выборки для оценки генеральной доли:

$$\Delta_p =$$

- доверительный интервал оценки генеральной доли:

$$\underline{\hspace{2cm}} \leq p \leq \underline{\hspace{2cm}} ;$$

$$\underline{\hspace{2cm}} \leq p \leq \underline{\hspace{2cm}} .$$

запишите вывод по результатам расчетов по шаблону (заполните пропуски):

С надежностью _____ можно утверждать, что в средняя продолжительность беременности крольчих составляет от _____ до _____ дней. При этом доля крольчих с беременностью менее 30 дней составляет от _____% до _____%.

Самостоятельная работа

Задание 2.8

По данным, представленным в таблице 2.8, вычислите по совокупности результатов анализа проб молока среднее значение процента содержания жира в молоке, охарактери-

зуйте степень вариации значений этого признака, вычислив показатели вариации. Для расчетов используйте табличный метод. Проверьте правильность расчетов, используя в MS Excel статистические функции СРЗНАЧ, ДИСПР, СТАНДОТКЛОНП. Сделайте выводы.

Таблица 2.8 – Содержание жира в молоке (%)

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

Решение:

- 1) в файле MS Excel с именем «Группа_ФИО_Тема 2» Лист 3 переименуйте «Задание 2.8»;
- 2) на данном листе задайте необходимые параметры форматирования текста в ячейках электронной таблицы;
- 3) подготовьте на этом листе таблицу 2.8 с исходными данными;
- 4) для выполнения расчетов числовых характеристик признака «содержание жира в молоке» с применением статистических функций подготовьте таблицу 2.9;

Таблица 2.9 – Числовые характеристики уровня жирности молока в выборке

Показатель, единицы измерения	Символ	Значение
Среднее значение, %	x_{cp}	
Размах вариации, %	R	
Дисперсия	S^2	
Стандартное отклонение, %	S	
Коэффициент вариации, %	V	

5) вычислите с точностью до 0,01 среднее значение признака (x_{cp}), размах вариации (R); с точностью до 0,001 показатели дисперсии (S^2) и стандартного отклонения (S), значение коэффициента вариации, по формуле:

$$V = \frac{S}{x_{cp}} \cdot 100 = \frac{S}{x_{cp}} \cdot 100 = \quad \%;$$

При выполнении расчетов используйте адресные ссылки на ячейки с ранее вычисленными значениями показателей в этой таблице.

5) по результатам расчетов сформулируйте вывод:

Данные таблицы 2.9 показывают, что

Задание 2.9

По данным о содержании жира в молоке вычислите по совокупности результатов проб среднее значение процента жира, охарактеризуйте степень вариации значений этого признака, вычислив показатели вариации для интервального ряда распределения значений этого признака, используя для расчетов исходные сгруппированные данные и расчетную

таблицу (таблица 2.10). Сделайте выводы. Сравните полученные результаты расчетов с результатами задания 2.8.

Таблица 2.10 – Расчет числовых характеристик величины уровня процента жира в молоке

Интервал значений процента жирности молока, %	Количество проб (f_i)	Середина интервала (x_i)	$x_i \cdot f_i$	$(x_i - \tilde{x})^2 \cdot f_i$
Итого				

Решение:

- 1) в файле MS Excel «Группа_ФМО_Тема 2» Лист 4 переименуйте «Задание 2.9»;
- 2) на данном листе задайте необходимые параметры форматирования текста в ячейках электронной таблицы;
- 3) подготовьте на этом листе таблицу 2.10, заполнив первый и второй столбцы таблицы в соответствии с результатами группировки, полученными в задании 1.4 (табл. 1.11);
- 4) для выполнения расчетов числовых характеристик уровня жирности молока с применением статистических формул подготовьте таблицу 2.11;

Таблица 2.11 – Числовые характеристики уровня жирности молока в выборке

Показатель, единицы измерения	Символ	Значение
Среднее значение, %	x_{cp}	
Размах вариации, %	R	
Дисперсия	S^2	
Стандартное отклонение, %	S	
Коэффициент вариации, %	V	

- предварительно запишите формулы для расчета перечисленных числовых характеристик:

$$\tilde{x} = \text{-----} = \text{-----} =$$

$$R =$$

$$S^2 = \text{-----} = \text{-----} =$$

$$S =$$

$$V = \text{-----} \cdot 100 = \text{-----} \cdot 100 = \quad \%$$

- в таблице 2.10 вычислите середину каждого интервала, а также все необходимые промежуточные расчеты и итоговые показатели в последней строке таблицы;

- в файле MS Excel в таблице 2.11 вычислите с точностью до 0,01 среднее значение, размах вариации (R), с точностью до 0,001 дисперсию (S^2), стандартное отклонение (S), коэффициент вариации, результаты расчетов запишите в рабочую тетрадь в виде выражений и в таблицу 2.11.

5) по результатам расчетов сформулируйте выводы:

Данные таблицы 2.11 показывают, что

При сравнении результатов расчета числовых характеристик уровня жирности молока в отобранных пробах, полученных в таблицах 2.9 и 2.11, можно сделать вывод, что значения показателей, вычисленные по не сгруппированным и сгруппированным данным _____ (отличаются/не отличаются). Это объясняется тем, что

Задание 2.10

В соответствии с номером варианта, указанным преподавателем, выполните **Индивидуальную работу №1** по теме «Анализ медико-биологических данных на основе их группировки и графического представления. Оценка характеристик генеральной совокупности по выборке», представленную в методичке [2] на стр. 6. Для выполнения расчетов в MS Excel создайте файл с именем «Группа_ФИО_ИР №1_Вариант». Рукопись работы можно выполнить в тетради или в виде текстового файла MS Word (на титульном листе обязательно укажите ФИО, номер группы и номер варианта работы). Работа может быть сдана в печатном (рукописном) виде или в электронном виде (текстовый файл формата PDF) в комплекте с расчетным файлом MS Excel. Срок выполнения работы две недели после выдачи задания.

Тема 3 Нормальный закон распределения

Наименование работы: оценка величины характеристик положения: показателей асимметричности и крутости эмпирического распределения по данным выборки; практическое использование свойств нормального распределения (распределения Гаусса) для статистического оценивания по опытным данным.

Цель работы: научиться вычислять показатели центральной тенденции – модальное и медианное значения признака, коэффициенты асимметрии и эксцесса, применять свойства нормального распределения для статистического оценивания характеристик и формы эмпирического распределения.

Приобретаемые навыки и умения: расчет параметрических и не параметрических характеристик положения для выборки с использованием инструмента «Описательная статистика» в наборе средств анализа данных «Пакет анализа» в MS Excel; проверка условий близости эмпирического распределения к нормальному закону; применение свойств нормального закона для статистического оценивания свойств и характеристик генеральной совокупности.

Норма времени: аудиторная работа - 90 мин.; самостоятельная работа – 180 мин.

Оснащение рабочего места: рабочая тетрадь, ПК с табличным процессором MS Excel.

Литература: [1].

Аудиторная работа

Задание 3.1

По исходным данным задания 1.2 (табл. 1.4) и результатам расчетов, полученным в задании 2.1 (табл. 2.2), определите:

1) непараметрические характеристики положения (моду и медиану), коэффициенты асимметрии и эксцесса для распределения значений величины прироста суточного удоя коров в результате использования витаминной добавки, применив инструмент «Описательная статистика» в Пакете анализа MS Excel;

2) сравните значения параметрических характеристик положения, полученные в задании 2.1 (табл. 2.2) с их значениями, полученными с помощью *Пакета анализа*, сделайте выводы;

3) выполните оценку близости данного распределения к нормальному закону на основе рассчитанных показателей асимметрии и эксцесса, сделайте выводы;

4) запишите выражение функции плотности вероятности, описывающей нормальный закон распределения величины прироста суточного удоя;

5) вычислите вероятность того, что прирост суточного удоя в результате использования витаминной добавки в изучаемой генеральной совокупности составит от 2,7 до 3,0 литров;

6) вычислите вероятность того, что прирост суточного удоя в результате использования витаминной добавки будет отличаться от среднего по выборке (опытным данным) не более, чем на 1 литр;

7) вычислите границы интервала, в пределах которого достоверно можно ожидать размер прироста суточного удоя в результате использования витаминной добавки.

Решение:

1) создайте файл MS Excel с именем «Группа_ФИО_Тема 3», в котором Лист 1 переименуйте «Задание 3.1»;

2) на данном листе задайте необходимые параметры форматирования текста в ячейках электронной таблицы: выделите лист целиком, используя сочетание клавиш **CTRL+A**,

установите тип шрифта Times New Roman, 14 пт, формат отображения в ячейках – выравнивание «по центру» по вертикали и по горизонтали с переносом по словам;

3) из файла «Группа_ФИО_Тема 1» скопируйте таблицу 1.4 с исходными данными;

4) для корректного применения Пакета анализа преобразуйте таблицу 1.4 так, чтобы значения величины прироста удоя были расположены в один столбец (переименуйте таблицу в «Таблица 3.1 – Прирост суточного удоя ...»);

5) активируйте надстройку **Пакет анализа** в MS Excel (алгоритм активации будет разный в зависимости от версии MS Excel);

Рекомендация:



1. Чтобы вывести список доступных инструментов анализа, выберите команду **Анализ данных** в меню **Сервис**. Если команда **Анализ данных** отсутствует, то можно ее активировать по следующему алгоритму:

- в меню **Сервис** выбрать команду **Надстройки**;
- в диалоговом окне **Надстройки** выбрать пункт **Пакет анализа** (поставить) и нажать **ОК** (начнется загрузка пакета **Анализ данных**);
- проверить факт появления в меню **Сервис** строки **Анализ данных** (обычно последняя строка в диалоговом окне **Сервис**);
- при выборе строки **Анализ данных** появится окно с инструментами **Пакета анализа**.



2. В меню перейдите в раздел **ДАнные** и найдите раздел **Анализ данных**. Если раздел **Анализ данных** отсутствует, то можно его активировать (установить) по следующему алгоритму:

- в разделе **ДАнные** по щелчку правой кнопкой мыши в диалоговом окне выберите строку **Настройка ленты**;
- в появившемся окне слева выберите строку **Надстройки**;
- в открывшемся окне категории **Надстройки** в самом низу в строке **Управление – Надстройки Excel** нажмите кнопку «**Перейти...**»;
- в диалоговом окне «**Надстройки**» в списке **Доступные надстройки** выберите пункт **Пакет анализа** (поставить) и нажать **ОК** (начнется загрузка пакета **Анализ данных**);
- справа на ленте раздела **ДАнные** появиться раздел (кнопка перехода) **Анализ данных**.

б) откройте окно **Анализ данных** и выберите инструмент **Описательная статистика**;

7) в диалоговом окне **Описательная статистика** укажите **Входной интервал** (блок ячеек со значениями прироста удоя из таблицы 3.1), в **Параметрах вывода** – **Выходной интервал** (любое количество ячеек, расположенных правее таблицы 3.1), поставьте в строке **Итоговая статистика** (рис. 3), нажмите **ОК**;

8) в результате будет представлена таблица с основными числовыми характеристиками исследуемой выборки (опытных данных), которую оформите как таблицу 3.2, добавив название «Таблица 3.2 – Числовые характеристики величины прироста удоя коров», заголовки столбцов, границы, отформатировав тип и размер шрифта (табл. 3.2);

9) перепишите в рабочую тетрадь (табл. 3.2) значения характеристик из файла, полученные с помощью **Пакета анализа**, а также значения тех характеристик, которые были получены для данной выборки в заданиях 1.2, 2.1 (табл. 2.2) и 2.3;

10) сделайте вывод о соотношении значений этих показателей (равны или если не равны, то почему):

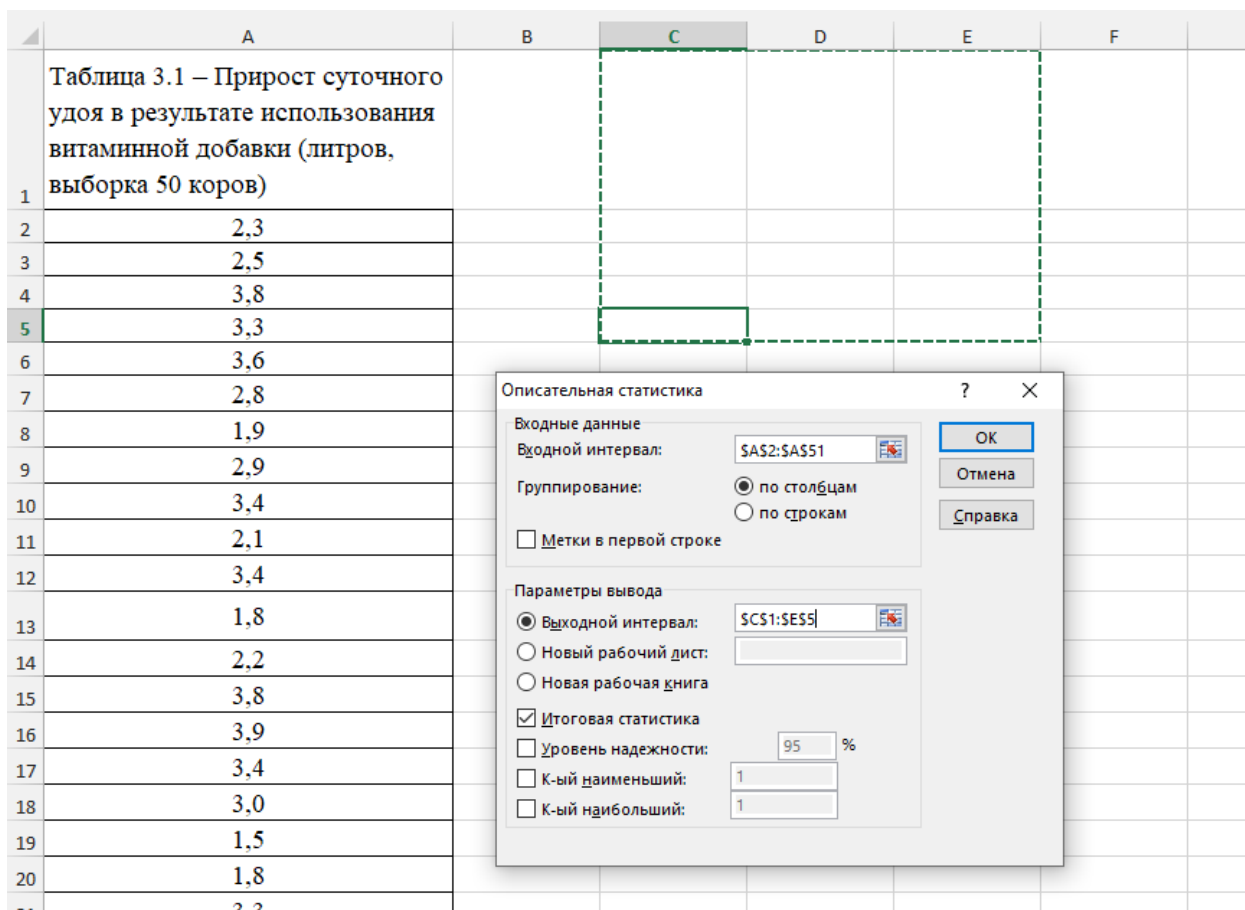


Рисунок 3 - Пример оформления таблицы 3.1 и установки параметров в диалоговом окне **Описательная статистика** в MS Excel

Таблица 3.2 – Числовые характеристики прироста суточного удоя коров

Показатель	Значение (Пакет анализа)	Значение из таблицы 2.2
Среднее		
Стандартная ошибка		
Медиана		
Мода		
Стандартное отклонение		
Дисперсия выборки		
Экссесс		
Асимметричность		
Интервал		
Минимум		
Максимум		
Сумма		
Счет		

Данные таблицы 3.2 показывают, что

11) для оценки близости данного распределения к нормальному закону на основе рассчитанных показателей асимметрии и эксцесса запишите формулы предельных неравенств (формулы (3.12) и (3.13) в учебном пособии [1]) и проверьте их выполнимость:

$|As| < \quad$; расчет:

$|Ex| < \quad$; расчет:

12) сформулируйте выводы по предельным неравенствам, а также по непараметрическим характеристикам положения (мода и медиана), коэффициентам асимметрии и эксцесса в соответствии с шаблоном:

<p><i>Значение медианы распределения, равное $Me = \quad$ литров, показывает, что</i></p> <hr/> <hr/>
<p><i>Значение моды распределения, равное $Mo = \quad$ литров, показывает, что</i></p> <hr/> <hr/>
<p><i>Соотношение значений среднего, моды и медианы имеет вид \quad, а значит распределение величины прироста суточного уdoa обладает \quad асимметрией.</i></p>
<p><i>Значение коэффициента асимметрии, равное $As = \quad$, указывает на то, что опытное распределение имеет \quad асимметрию.</i></p>
<p><i>Значение коэффициента эксцесса, равное $Ex = \quad$, указывает на то, что опытное распределение является \quad по сравнению с нормальным законом распределения.</i></p>
<p><i>Таким образом, можно предположить, что опытное распределение \quad близким к нормальному закону распределения с параметрами: среднее значение $\bar{x} = \quad$ и стандартное отклонение $\sigma = \quad$.</i></p>
<p><i>Предельные неравенства в отношении значений коэффициентов асимметрии и эксцесса (выполняются / не выполняются) \quad.</i></p>
<p><i>Выдвинутое предположение о близости распределения величины прироста суточного уdoa к нормальному закону \quad.</i></p>

13) формула функции плотности вероятности, описывающей нормальный закон распределения имеет выражение (формула 3.8 из учебного пособия [1]):

$$f(x) = \quad ,$$

тогда функция плотности вероятности, описывающей нормальный закон распределения величины прироста суточного уdoa с параметрами $a = \quad$ и $\sigma = \quad$ будет иметь выражение:

$$f(x) = \quad ,$$

14) по первому свойству нормального закона распределения (формула (3.9) из

учебного пособия [1]): *вероятность того, что случайная величина X, распределенная нормально, примет значение в интервале от x₁ до x₂ равна*

$$P(x_1 < X < x_2) = \Phi\left(\frac{x_2 - a}{\sigma}\right) - \Phi\left(\frac{x_1 - a}{\sigma}\right),$$

где $\Phi(x)$ – интегральная функция стандартного нормального распределения, называемая *функцией Лапласа*, значения которой для разных значений аргумента x рассчитаны и приводятся в специальных статистических таблицах (Приложение 1).

Тогда вероятность того, что прирост суточного удоя (X) в результате использования витаминной добавки в изучаемой генеральной совокупности составит от 2,7 до 3,0 литров, будет равна:

$$P(2,7 < X < 3,0) = \Phi\left(\frac{3,0 - a}{\sigma}\right) - \Phi\left(\frac{2,7 - a}{\sigma}\right) = \Phi\left(\frac{3,0 - 2,7}{\sigma}\right) - \Phi\left(\frac{2,7 - 2,7}{\sigma}\right) = \Phi\left(\frac{0,3}{\sigma}\right) - \Phi(0) = \Phi\left(\frac{0,3}{\sigma}\right) - 0,5$$

15) по второму свойству нормального закона распределения (формула (3.1) из учебного пособия [1]): *вероятность того, что значение случайной величины X, распределенной нормально, полученное в результате наблюдения будет отличаться от среднего ее значения по абсолютной величине не более, чем на некоторую малую положительную величину Δ равна*

$$P(|X - a| \leq \Delta) = 2 \cdot \Phi\left(\frac{\Delta}{\sigma}\right) - 1$$

Тогда вероятность того, что прирост суточного удоя (X) в результате использования витаминной добавки будет отличаться от среднего по выборке (опытным данным) не более, чем на 1 литр, будет равна:

$$P(|X - 2,7| \leq 1) = 2 \cdot \Phi\left(\frac{1}{\sigma}\right) - 1 = 2 \cdot \Phi\left(\frac{1}{\sigma}\right) - 1$$

16) по третьему свойству нормального распределения («правило трех сигм»): *вероятность того, что отклонение по абсолютной величине значения случайной величины X, распределенной нормально, от своего среднего значения не превысит утроенного среднего квадратического отклонения, равна 0,9973, т.е. является практически достоверным.*

Тогда границы интервала, в пределах которого достоверно можно ожидать размер прироста суточного удоя в результате использования витаминной добавки по «правилу трех сигм», будут равны:

формула: $a - 3\sigma \leq X \leq a + 3\sigma$

расчеты: $2,7 - 3 \cdot 0,9 \leq X \leq 2,7 + 3 \cdot 0,9$

$0,0 \leq X \leq 5,4$

Самостоятельная работа

Задание 3.2

По исходным данным задания 1.4 (табл. 1.9) и результатам расчетов, полученным в заданиях 2.8 (табл. 2.9) и 2.4, определите:

1) непараметрические характеристики положения (моду и медиану), коэффициенты асимметрии и эксцесса для распределения значений величины уровня жирности молока, применив инструмент «**Описательная статистика**» в **Пакете анализа MS Excel**;

2) сравните значения параметрических характеристик положения, полученные в задании 2.8 (табл. 2.9) с их значениями, полученными с помощью **Пакета анализа**, сделайте выводы;

3) выполните оценку близости данного распределения к нормальному закону на основе рассчитанных показателей асимметрии и эксцесса, сделайте выводы;

4) запишите выражение функции плотности вероятности, описывающей нормальный закон распределения величины уровня жирности молока;

5) вычислите вероятность того, что уровень жирности молока в изучаемой генеральной совокупности составит от 3,5 до 4,5 процентов;

6) вычислите вероятность того, что уровень жирности молока в любой пробе будет отличаться от среднего по выборке (опытным данным) не более, чем на 0,5 процента;

7) вычислите границы интервала, в пределах которого достоверно можно ожидать уровень жирности молока в любой пробе.

Решение:

1) в файле MS Excel с именем «**Группа_ФИО_Тема 3**» Лист 2 переименуйте «**Задание 3.2**»;

2) на данном листе задайте необходимые параметры форматирования текста в ячейках электронной таблицы;

3) из файла «**Группа_ФИО_Тема 1**» скопируйте таблицу 1.9 с исходными данными;

4) для корректного применения Пакета анализа преобразуйте таблицу 1.9 так, чтобы значения величины уровня жира в молоке были расположены в один столбец (переименуйте таблицу в «**Таблица 3.3 – Содержание жира в молоке ...**»);

5) откройте окно **Анализ данных** и примените инструмент **Описательная статистика** (таблицу с основными числовыми характеристиками исследуемой выборки оформите как таблицу 3.4, добавив название «**Таблица 3.4 – Числовые характеристики величины уровня жирности молока**», заголовки столбцов, границы, отформатировав тип и размер шрифта (табл. 3.4);

Таблица 3.4 – Числовые характеристики величины уровня жирности молока

Показатель	Значение (Пакет анализа)	Значение из таблицы 2.9
Среднее		
Стандартная ошибка		
Медиана		
Мода		
Стандартное отклонение		
Дисперсия выборки		
Эксцесс		
Асимметричность		
Интервал		
Минимум		
Максимум		
Сумма		
Счет		

б) перепишите в рабочую тетрадь (табл. 3.4) значения характеристик из файла, полученные с помощью **Пакета анализа**, а также значения тех характеристик, которые были получены для данной выборки в заданиях 1.4, 2.8 (табл. 2.9) и 2.4;

7) сделайте вывод о соотношении значений этих показателей:

Данные таблицы 3.4 показывают, что

8) для оценки близости данного распределения к нормальному закону на основе рассчитанных показателей асимметрии и эксцесса запишите формулы предельных неравенств и проверьте их выполнимость:

$|As| < \quad ;$ расчет:

$|Ex| < \quad ;$ расчет:

9) сформулируйте выводы по предельным неравенствам, а также по непараметрическим характеристикам положения (мода и медиана), коэффициентам асимметрии и эксцесса в соответствии с шаблоном:

Значение медианы распределения, равное $Me = \quad \%$, показывает, что

Значение моды распределения, равное $Mo = \quad \%$, показывает, что

Соотношение значений среднего, моды и медианы имеет вид \quad , а значит распределение величины уровня жирности молока обладает \quad асимметрией.

Значение коэффициента асимметрии, равное $As = \quad$, указывает на то, что опытное распределение имеет \quad асимметрию.

Значение коэффициента эксцесса, равное $Ex = \quad$, указывает на то, что опытное распределение является \quad по сравнению с нормальным законом распределения.

Таким образом, можно предположить, что опытное распределение \quad близким к нормальному закону распределения с параметрами: среднее значение $\bar{x} = \quad$ и стандартное отклонение $\sigma = \quad$.

Предельные неравенства в отношении значений коэффициентов асимметрии и эксцесса (выполняются / не выполняются) \quad .

Выдвинутое предположение о близости распределения величины уровня жирности молока к нормальному закону \quad .

10) формула функции плотности вероятности, описывающей нормальный закон распределения имеет выражение (формула 3.8 из учебного пособия [1]):

$$f(x) = \quad ,$$

тогда функция плотности вероятности, описывающей нормальный закон распределения величины уровня жирности молока с параметрами $a = \underline{\hspace{2cm}}$ и $\sigma = \underline{\hspace{2cm}}$ будет иметь выражение:

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{(x-a)^2}{2\sigma^2}\right),$$

11) по первому свойству нормального закона распределения (формула (3.9) из учебного пособия [1]): *вероятность того, что случайная величина X , распределенная нормально, примет значение в интервале от x_1 до x_2 равна*

$$P(x_1 < X < x_2) = \Phi\left(\frac{x_2 - a}{\sigma}\right) - \Phi\left(\frac{x_1 - a}{\sigma}\right),$$

где $\Phi(x)$ – интегральная функция стандартного нормального распределения, называемая *функцией Лапласа*, значения которой для разных значений аргумента x рассчитаны и приводятся в специальных статистических таблицах (Приложение 1).

Тогда вероятность того, что уровень жирности молока (X) в любой пробе составит от 3,5 до 4,5 процентов, будет равна:

$$\begin{aligned} P(3,5 < X < 4,5) &= \Phi\left(\frac{4,5 - a}{\sigma}\right) - \Phi\left(\frac{3,5 - a}{\sigma}\right) = \\ &= \Phi(\underline{\hspace{2cm}}) - \Phi(\underline{\hspace{2cm}}) = \end{aligned}$$

12) по второму свойству нормального закона распределения (формула (3.1) из учебного пособия [1]): *вероятность того, что значение случайной величины X , распределенной нормально, полученное в результате наблюдения будет отличаться от среднего ее значения по абсолютной величине не более, чем на некоторую малую положительную величину Δ равна*

$$P(|X - a| \leq \Delta) = 2 \cdot \Phi\left(\frac{\Delta}{\sigma}\right),$$

Тогда вероятность того, что уровня жирности молока (X) в любой пробе будет отличаться от среднего по выборке (опытным данным) не более, чем на 0,5 процентов, будет равна:

$$P(|X - \underline{\hspace{2cm}}| \leq \underline{\hspace{2cm}}) = 2 \cdot \Phi\left(\frac{\underline{\hspace{2cm}}}{\sigma}\right) = 2 \cdot \Phi(\underline{\hspace{2cm}}) = \underline{\hspace{2cm}}.$$

13) по третьему свойству нормального распределения («правило трех сигм»): *вероятность того, что отклонение по абсолютной величине значения случайной величины X , распределенной нормально, от своего среднего значения не превысит утроенного среднего квадратического отклонения, равна 0,9973, т.е. является практически достоверным.*

Тогда границы интервала, в пределах которого достоверно можно ожидать уровень жирности молока в любой пробе, по «правилу трех сигм» будут равны:

формула: $\underline{\hspace{2cm}} \leq X \leq \underline{\hspace{2cm}},$

расчеты: $\underline{\hspace{2cm}} \leq X \leq \underline{\hspace{2cm}},$

$\underline{\hspace{2cm}} \leq X \leq \underline{\hspace{2cm}}.$

Тема 4 Статистические гипотезы и их проверка

Наименование работы: применение параметрических и непараметрических статистических критериев для проверки гипотез, сформулированных по результатам анализа опытных данных.

Цель работы: усвоить алгоритм проверки статистической гипотезы с помощью статистического критерия, научиться применять критерий Стьюдента для зависимых и независимых выборок, усвоить методику применения критерия Уилкоксона для малых выборок, научиться делать практически значимые выводы по результатам применения статистических критериев.

Приобретаемые навыки и умения: формирование навыка табличной формы выполнения статистических расчетов, умение применять критерий Стьюдента для зависимых и независимых выборок, критерия Уилкоксона для малых выборок, формирование и закрепление навыка использования статистических функций СРЗНАЧ, ДИСПР, СТЬЮДРАС-ПОБР, РАНГ, закрепление навыка графического представления результатов наблюдений в виде статистических диаграмм.

Норма времени: аудиторная работа - 180 мин.; самостоятельная работа – 270 мин.

Оснащение рабочего места: рабочая тетрадь, ПК с табличным процессором MS Excel.

Литература: [1], [2].

Аудиторная работа

Задание 4.1

Известны результаты измерения связанного холестерина крови (мг%) у здоровых собак (норма) и у собак, больных гипотиреозом (заболевание щитовидной железы), представленные в таблице 4.1.

Таблица 4.1 – Содержание холестерина (мг%)

Норма	5,9	5,1	6,1	5,3	6,0	7,0	5,3	6,1	8,3
Гипотиреоз	10,3	8,7	11,2	10,6	9,1	9,8	9,5	8,1	-

Проверить гипотезу о равенстве двух генеральных средних с использованием критерия Стьюдента. Сформулировать нулевую и альтернативную гипотезы для исследуемого случая опытных данных. Сделать выводы на уровне значимости $\alpha=0,05$. Представьте результаты исследования уровня холестерина в графическом виде.

Решение:

1) создайте файл MS Excel с именем «Группа_ФИО_Тема 4», в котором Лист 1 переименуйте «Задание 4.1»;

2) на данном листе задайте необходимые параметры форматирования текста в ячейках электронной таблицы: выделите лист целиком, используя сочетание клавиш **CTRL+A**, установите тип шрифта Times New Roman, 14 пт, формат отображения в ячейках – выравнивание «по центру» по вертикали и по горизонтали с переносом по словам;

3) в файле создайте таблицу 4.1 с исходными данными;

4) определите, с каким случаем выборок связан исследуемый эксперимент (зависимые или независимые выборки), сформулируйте нулевую (проверяемую) и альтернативную гипотезы для исследуемого случая опытных данных (заполните пропуски в шаблоне формулировок гипотез):

Эксперимент связан с исследованием содержания холестерина в крови у двух групп подопытных собак, которые представляют собой две _____ выборки, т.к. _____.

Проверим с помощью критерия Стьюдента (*t*-критерий) для _____ выборки справедливость нулевой гипотезы H_0 : _____ (различия между средними значениями содержания холестерина в крови у здоровых и больных гипотиреозом собак _____, а значит гипотиреоз _____ повышение уровня холестерина в крови больного животного).

Сформулируем альтернативную гипотезу H_1 : _____ (различия между средними значениями содержания холестерина в крови у здоровых и больных гипотиреозом собак _____, а значит гипотиреоз _____ повышение уровня холестерина в крови больного животного).

Проверим справедливость нулевой гипотезы на уровне значимости $\alpha = 0,05$, используя критерий Стьюдента, расчетное значение которого следует вычислять по формуле (формула 4.3 из учебного пособия [1]):

$$t_{\text{РАСЧ}} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s^2}{n_1} + \frac{s^2}{n_2}}},$$

где _____

5) для выполнения расчетов постройте расчетную таблицу 4.2 (скопируйте таблицу 4.1 и вставьте ее, используя специальную вставку с транспонированием данных (преобразование строк в столбцы таблицы), переименуйте таблицу в «Таблица 4.2 – Расчет средних значений, дисперсий и стандартных ошибок выборок», добавьте необходимые столбцы и строки (табл. 4.2);

Таблица 4.2 – Расчет средних значений, дисперсий и стандартных ошибок выборок

Номер наблюдения	Норма	Гипотиреоз
<i>символ</i>	$x_{1;i}$	$x_{2;i}$
1	5,9	10,3
2	5,1	8,7
3	6,1	11,2
4	5,3	10,6
5	6,0	9,1
6	7,0	9,8
7	5,3	9,5
8	6,1	8,1
9	8,3	-
Среднее значение		
Дисперсия		
Стандартная ошибка выборки		
Значение $t_{\text{РАСЧ}}$		
Значение $t_{\text{ТАБЛ}}$		

6) используя статистические функции **СРЗНАЧ** и **ДИСПР**, вычислите средние значения (с точностью до 0,01) и дисперсии (с точностью до 0,001) исследуемого признака в выборках, перепишите результаты расчетов в рабочую тетрадь в таблицу 4.2;

7) запишите формулы для расчета стандартных ошибок выборок (формулы (4.1) и (4.2) из учебного пособия [1]), подставив числовые значения показателей в эти формулы, расчеты (с точностью до 0,0001) выполните в файле в последней строке таблицы 4.2, результаты их перепишите в рабочую тетрадь:

$$m_1 = \quad = \quad = \quad ,$$

$$m_2 = \quad = \quad = \quad ,$$

8) вычислите (с точностью до 0,01) расчетное значение критерия Стьюдента в расчетном файле и запишите порядок и результат расчета в рабочую тетрадь:

$$t_{\text{РАСЧ}} = \frac{\quad}{\quad} = \frac{\quad}{\quad} = \quad ,$$

9) по таблице распределения Стьюдента (Приложение 2) для уровня значимости $\alpha = 0,05$ и числа степеней свободы $k = \quad = \quad$ значение критерия составляет $t_{\text{ТАБЛ}} = \quad$. В файле в таблице 4.2 определите табличное значение критерия Стьюдента, используя функцию **СТЮДРАСПОБР**.

Рекомендация:



Табличное значение критерия Стьюдента в MS Excel можно получить с помощью функции **СТЮДРАСПОБР(Вероятность; Степени_свободы)**, в которой в качестве аргумента **Вероятность** указывается принятое при проверке гипотезы значение уровня значимости $\alpha = 0,05$, а аргумента **Степени_свободы** – число степеней свободы k .

10) сравните расчетное и табличное значения критерия Стьюдента, сформулируйте вывод относительно справедливости нулевой гипотезы (по шаблону):

Результаты расчетов показывают, что $t_{\text{РАСЧ}} \quad t_{\text{ТАБЛ}}$, а значит на 5%-ном уровне значимости справедливость нулевой гипотезы H_0 \quad (доказана/не доказана). Таким образом, различия между средними значениями содержания холестерина в крови у здоровых и больных гипотиреозом собак \quad , а значит гипотиреоз \quad повышение уровня холестерина в крови больного животного.

11) представьте результаты исследования уровня содержания холестерина в крови здоровых и больных гипотиреозом собак графически (рис. 4.1), используя в MS Excel тип диаграммы «График», изобразив два ряда данных (Ряд 1 – Норма, Ряд 2 – Гипотиреоз) (пример диаграммы представлен на рисунке 4.1 в учебном пособии [1]).

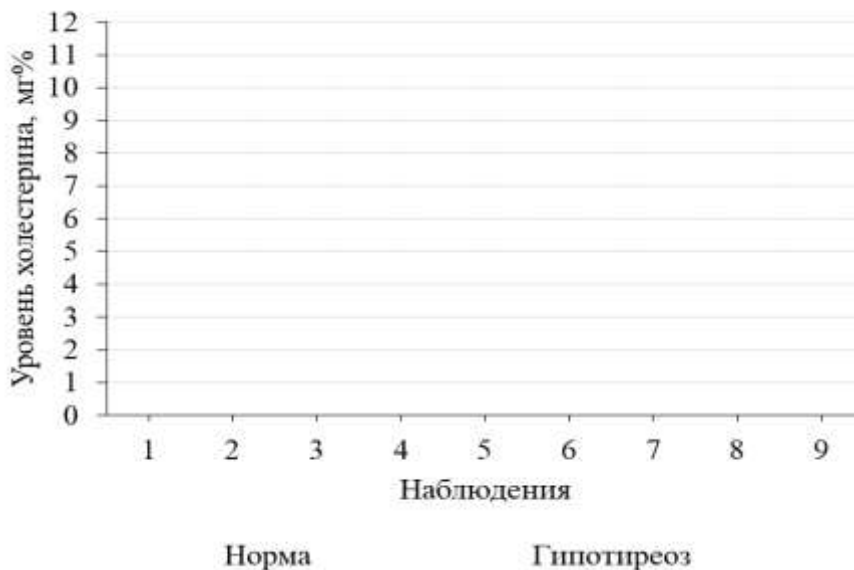


Рисунок 4.1 – Содержание холестерина в крови здоровых и больных гипотиреозом собак

Задание 4.2

Получены следующие результаты измерения влияния микробиологического препарата «Б1-У» («Урга» - полный аналог препарата «Байкал ЭМ1») на продуктивность молочных коров по результатам опытного применения на контрольной группе животных (табл. 4.2). Проверить гипотезу о равенстве двух генеральных средних с использованием критерия Стьюдента. Сформулировать нулевую и альтернативную гипотезы для исследуемого случая опытных данных. Сделать выводы на уровне значимости $\alpha=0,05$.

Таблица 4.3 - Результаты измерения влияния микробиологического препарата «Б1-У»

№ п/п	Инвентарный номер коровы	Средняя удоимость, л/сутки	Контрольный надой после применения «Б1-У», л/сутки
1	447	18,0	19,0
2	509	16,3	18,0
3	546	7,1	7,0
4	418	17,2	19,0
5	511	13,0	17,0
6	585	13,0	13,0
7	542	17,0	17,0
8	541	11,0	12,0
9	049	15,6	18,0
10	457	18,5	14,0

Решение:

1) в файле MS Excel с именем «Группа_ФИО_Тема 4» Лист 2 переименуйте «Задание 4.2» и задайте необходимые параметры форматирования текста в ячейках электронной таблицы;

2) в файле создайте таблицу 4.3 с исходными данными;

3) определите, с каким случаем выборок связан исследуемый эксперимент (зависимые или независимые выборки), сформулируйте нулевую (проверяемую) и альтернативную гипотезы для исследуемого случая опытных данных (заполните пропуски в шаблоне формулировок гипотез):

4)

Эксперимент связан с исследованием уровня удоимости коров до и после применения микробиологического препарата «Б1-У», и представленные в таблице 4.3 два ряда данных представляют собой две _____ выборки.

Проверим с помощью критерия Стьюдента (*t*-критерий) для _____ выборки справедливость нулевой гипотезы H_0 : _____ (различия между средними значениями удоимости коров до и после применения препарата «Б1-У» _____, а значит препарат является _____ (эффективным/не эффективным) в отношении повышения продуктивности коров).

Сформулируем альтернативную гипотезу H_1 : _____ (различия между средними значениями удоимости коров до и после применения препарата «Б1-У» _____, а значит препарат является _____ (эффективным/не эффективным) в отношении повышения продуктивности коров).

Проверим справедливость нулевой гипотезы на уровне значимости $\alpha = 0,05$, используя критерий Стьюдента, расчетное значение которого следует вычислять по формуле (формула 4.6 из учебного пособия [1]):

$$t_{\text{РАСЧ}} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s^2}{n_1} + \frac{s^2}{n_2}}},$$

где _____

5) для выполнения расчетов постройте расчетную таблицу 4.4 (скопируйте таблицу 4.2 и вставьте ее, переименуйте таблицу в «Таблица 4.4 – Расчет средних значений и попарных разностей результатов измерения удойности коров»), добавьте необходимые столбцы и строки (табл. 4.4);

Таблица 4.4 – Расчет средних значений и попарных разностей результатов измерения удойности коров

№ п/п	Инвентарный номер коровы	Средняя удойность, л/сутки	Контрольный надой после применения «Б1-У», л/сутки	Попарные разности (d_i)
1	447	18,0	19,0	
2	509	16,3	18,0	
3	546	7,1	7,0	
4	418	17,2	19,0	
5	511	13,0	17,0	
6	585	13,0	13,0	
7	542	17,0	17,0	
8	541	11,0	12,0	
9	49	15,6	18,0	
10	457	18,5	14,0	
Среднее значение попарных разностей (d_{cp})				
Дисперсия попарных разностей (S^2_d)				
Стандартная ошибка средней попарных разностей (m_d)				
Значение $t_{РАСЧ}$				
Значение $t_{ТАБЛ}$				

б) используя статистические функции **СРЗНАЧ** и **ДИСПР**, вычислите среднее значение (с точностью до 0,01) и дисперсию (с точностью до 0,001) попарных разностей, перепишите результаты расчетов в рабочую тетрадь в таблицу 4.4;

7) запишите формулу для расчета стандартной ошибки среднего значения попарных разностей (формула (4.7) из учебного пособия [1]), подставив числовые значения показателей в эту формулу, расчеты (с точностью до 0,0001) выполните в файле в соответствующей строке таблицы 4.2, результаты перепишите в рабочую тетрадь:

$$m_d = \quad = \quad = \quad ,$$

8) вычислите (с точностью до 0,01) расчетное значение критерия Стьюдента в расчетном файле и запишите порядок и результат расчета в рабочую тетрадь:

$$t_{РАСЧ} = \frac{\quad}{\quad} = \frac{\quad}{\quad} = \quad ,$$

9) по таблице распределения Стьюдента (Приложение 2) для уровня значимости $\alpha = 0,05$ и числа степеней свободы $k = \quad = \quad$ значение критерия составляет $t_{ТАБЛ} = \quad$. В файле в таблице 4.4 определите табличное значение критерия Стьюдента, используя функцию **СТЮДРАСПОБР**.

10) сравните расчетное и табличное значения критерия Стьюдента, сформулируйте вывод относительно справедливости нулевой гипотезы (по шаблону):

Результаты расчетов показывают, что $t_{\text{РАСЧ}} \dots t_{\text{ТАБЛ}}$, а значит на 5%-ном уровне значимости справедливость нулевой гипотезы H_0 _____ (подтверждается/не подтверждается). Таким образом, различия между средними значениями удойности коров до и после применения препарата «Б1-У» _____, а значит эффективность микробиологического препарата «Б1-У» в отношении повышения продуктивности коров статистически _____ (доказана/не доказана).

11) представьте результаты исследования уровня удойности коров до и после применения препарата «Б1-У» графически (рис. 4.2), используя в MS Excel тип диаграммы «График», изобразив два ряда данных (Ряд 1 – до применения «Б1-У», Ряд 2 – после применения «Б1-У») (пример диаграммы представлен на рисунке 4.2 в учебном пособии [1]).

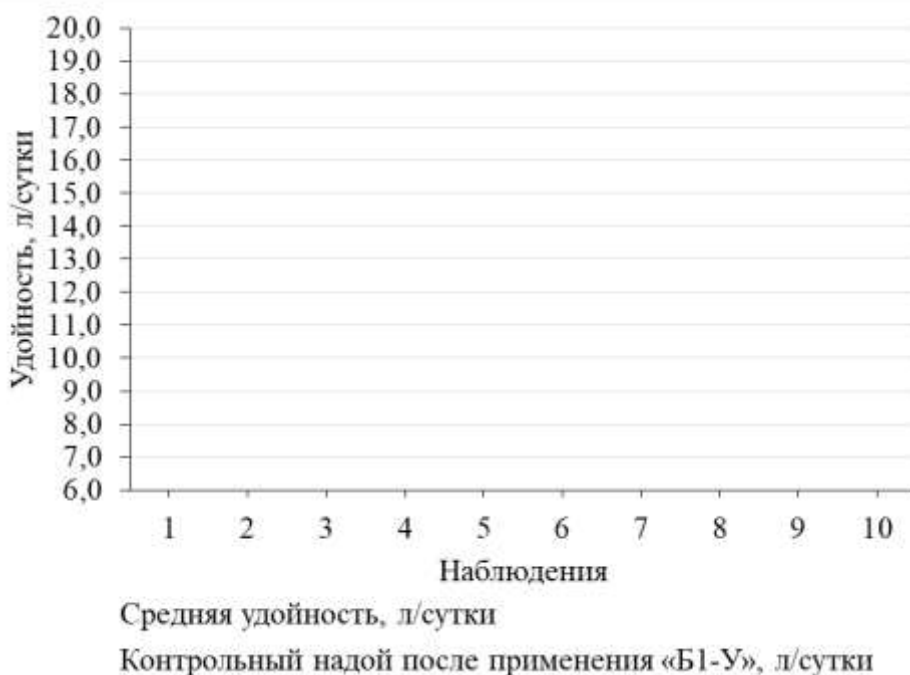


Рисунок 4.2 – Результаты измерения удойности коров (10 голов) до и после приема микробиологического препарата «Б1-У» («Урга» - полный аналог препарата «Байкал ЭМ1»)

Задание 4.3

Проверить гипотезу о равенстве двух генеральных средних с использованием критерия Уилкоксона (Манна-Уитни) по данным задания 4.1 (табл. 4.1) о результатах измерения связанного холестерина крови (мг%) у здоровых собак (норма) и у собак, больных гипотиреозом (заболевание щитовидной железы). Сформулировать нулевую и альтернативную гипотезы. Сделать выводы на уровне значимости $\alpha = 0,05$.

Решение:

1) в файле MS Excel с именем «Группа_ФИО_Тема 4» Лист 3 переименуйте «Задание 4.3» и задайте необходимые параметры форматирования текста в ячейках электронной таблицы;

2) в файле с Листа «Задание 4.1» скопируйте таблицу 4.1 с исходными данными;

3) определите, с каким случаем выборки связан исследуемый эксперимент (зависимые или независимые выборки), сформулируйте нулевую (проверяемую) и альтернативную гипотезы для исследуемого случая опытных данных (заполните пропуски в шаблоне формулировок гипотез):

Эксперимент связан с исследованием содержания холестерина в крови у двух групп подопытных собак, которые представляют собой две _____ выборки.

Проверим с помощью критерия Уилкоксона (Манна-Уитни) (*U*-критерий) справедливость нулевой гипотезы H_0 : _____ (различия между средними значениями содержания холестерина в крови у здоровых и больных гипотиреозом собак _____, а значит гипотиреоз _____ повышение уровня холестерина в крови больного животного).

Сформулируем альтернативную гипотезу H_1 : _____ (различия между средними значениями содержания холестерина в крови у здоровых и больных гипотиреозом собак _____, а значит гипотиреоз _____ повышение уровня холестерина в крови больного животного).

Проверим справедливость нулевой гипотезы на уровне значимости $\alpha = 0,05$, используя *U*-критерий, расчетные значения которого следует вычислять для исследуемых выборок по формуле (формула 4.9 из учебного пособия [1]):

$$U_1 = \quad ,$$

$$U_2 = \quad ,$$

где _____

4) расчеты выполните двумя способами:

1 способ: для выполнения расчетов постройте расчетную таблицу 4.5 («Таблица 4.5 – Расчет суммы рангов»), в соответствии с ее макетом (табл. 4.5);

Таблица 4.5 – Расчет суммы рангов (1 способ)

Номер наблюдения	Норма		Гипотиреоз	
	значение X_1	ранг X_1	значение X_2	ранг X_2
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
Сумма рангов R_i	-		-	
Расчетное значение критерия U_i				

- запишите (можно скопировать и вставить с использованием транспонирования) значения признака в каждой выборке (соответственно в столбцы X_1 и X_2), сортируя их по возрастанию значений, выстраивая при этом как бы в единый ранжированный ряд;

- справа соответственно в столбцах «Ранг X_1 » и «Ранг X_2 » напротив каждого значения запишите его ранг (порядковый номер) в получившемся объединенном ряду (если несколько значений признака одинаковые, то ранги у них тоже должны быть одинаковые и равны среднему значению их порядковых номеров в объединенном ряду);

- в ячейках строки «Сумма рангов», используя функцию **СУММ(Массив)**, вычислите суммы рангов в выборках (значения R_1 и R_2);

- запишите формулы и выражения для расчетных значений U-критерия по каждой выборке, выполните расчеты в таблице в файле и перепишите их результаты в таблицу 4.5 рабочей тетради:

$$U_1 = \quad = \quad = \quad ,$$

$$U_2 = \quad = \quad = \quad .$$

5) **2 способ:** в файле MS Excel для определения ранга каждого значения признака можно использовать статистическую функцию **РАНГ.СР(Число; Ссылка; Порядок)**, указывая в аргументах функции в качестве **Числа** ссылку на ячейку со значением признака, в качестве **Ссылки** блок ячеек со значениями обеих выборок, присвоив **Порядок** «1», означающий ранжирование по возрастанию;

- для выполнения расчетов постройте таблицу 4.6 (скопируйте таблицу 4.1, добавив необходимые столбцы и строки);

Таблица 4.6 – Расчет суммы рангов (2 способ)

Показатель	Значения									Сумма рангов R_i	Расчетное значение критерия U_i
	5,9	5,1	6,1	5,3	6	7	5,3	6,1	8,3		
Норма (X_1)	5,9	5,1	6,1	5,3	6	7	5,3	6,1	8,3		
Гипотиреоз (X_2)	10,3	8,7	11,2	10,6	9,1	9,8	9,5	8,1	-		
Ранг X_1											
Ранг X_2									-		

- в строке «Ранг X_1 » для первого значения примените функцию **РАНГ.СР(Число; Ссылка; Порядок)**, указав в качестве аргумента **Число** ссылку на ячейку со значением признака в выборке «Норма (X_1)» в этом столбце, в качестве аргумента **Ссылка** – блок ячеек со значениями признака в обеих выборках (при этом в адресной ссылке на массив ячеек зафиксируйте адреса его начальной и конечной ячеек с помощью функциональной клавиши F4), в строке **Порядок** укажите «1»;

- «протяните» (скопируйте) записанную функцию для остальных ячеек обеих выборок (строки «Ранг X_1 » и «Ранг X_2 »);

- в столбце «Сумма рангов R_i », используя функцию **СУММ(Массив)**, вычислите суммы рангов в выборках (значения R_1 и R_2);

- в столбце «Расчетное значение критерия U_i » вычислите по записанным ранее формулам значения критерия Уилкоксона;

- результаты расчетов перепишите в рабочую тетрадь в таблицу 4.6 и сравните их с расчетами, выполненными 1 способом;

б) сравните полученные значения критерия (U_1 и U_2) и по условиям критерия в качестве тестового расчетного значения выберите наименьшее из них, по таблице значений критерия Манна-Уитни (Приложение 3) для 5%-ного уровня значимости найдите табличное (критическое) значение критерия и, сравнив его с расчетным значением сделайте вывод относительно справедливости проверяемой нулевой гипотезы H_0 :

U_1 ____ U_2 , тогда тестовым является значение $U_{расч} =$ _____.

По таблице Манна-Уитни для объемов выборок $n_1 =$ ____ и $n_2 =$ ____ при 5%-ном уровне значимости $U_{табл} =$ _____.

$U_{расч}$ _____ $U_{табл}$, а значит на 5%-ном уровне значимости можно утверждать, что различия в содержании холестерина в крови у здоровых и больных гипотиреозом собак считаются статистически _____ а значит гипотиреоз _____ повышение уровня холестерина в крови больного животного.

Самостоятельная работа

Задание 4.4

В соответствие с номером варианта, указанным преподавателем, выполните *Индивидуальную работу №2 по теме «Статистические гипотезы и их проверка»*, представленную в методичке [2] на стр. 15. Для выполнения расчетов в MS Excel создайте файл с именем «Группа_ФИО_ИР №2_Вариант». Рукопись работы можно выполнить в тетради или в виде текстового файла MS Word (на титульном листе обязательно укажите ФИО, номер группы и номер варианта работы). Работа может быть сдана в печатном (рукописном) виде или в электронном виде (текстовый файл формата PDF) в комплекте с расчетным файлом MS Excel. Срок выполнения работы две недели после выдачи задания.

Тема 5 Исследование зависимостей на основе корреляционно-регрессионного анализа медико-биологических данных

Наименование работы: применение метода корреляционно-регрессионного анализа к исследованию статистических однофакторных связей на основе опытных данных.

Цель работы: усвоить основные задачи биостатистики при изучении корреляционных зависимостей, научиться оценивать силу корреляционной связи и ее направление, моделировать однофакторную зависимость, оценивать качество и статистическую значимость модели регрессии, вычислять прогнозы.

Приобретаемые навыки и умения: формирование навыка табличной формы выполнения статистических расчетов в рамках решения задач корреляционно-регрессионного анализа, умение применять критерии Стьюдента и Фишера для оценки статистической значимости модели регрессии, формирование и закрепление навыка использования статистических функций СРЗНАЧ, ДИСПР, СТАНДОТКЛОНП, СТЬЮДРАСПОБР, ФРАСПОБР, КОРРЕЛ, закрепление навыка графического представления корреляционной зависимости в виде статистической диаграммы. Овладение навыком выполнения корреляционно-регрессионного анализа с использованием инструмента «Регрессия» в наборе средств анализа данных «Пакет анализа» в MS Excel.

Норма времени: аудиторная работа - 180 мин.; самостоятельная работа – 270 мин.

Оснащение рабочего места: рабочая тетрадь, ПК с табличным процессором MS Excel.

Литература: [1], [2].

Аудиторная работа

Задание 5.1

По данным, представленным в таблице 5.1, выполните корреляционно-регрессионный анализ зависимости между содержанием жира (X) и белка (Y) в молоке коров (в %): определите линейную модель регрессии, дайте интерпретацию ее параметров; рассчитайте показатели тесноты связи между признаками (индекс корреляции, линейные коэффициенты корреляции и детерминации), сделайте выводы по их значениям; проверьте статистическую значимость полученной модели регрессии и коэффициента корреляции на основе применения статистических критериев Стьюдента и Фишера; выполните анализ качества полученной модели и пригодности ее для прогнозирования; выполните прогноз значения результирующей переменной (Y) на основе модели регрессии, приняв ожидаемое значение факторной переменной X , равное среднему значению этого признака по результатам наблюдений; изобразите графически исследуемую зависимость, построив поле корреляции и график линии регрессии, выведите на поле графика уравнение регрессии и коэффициент детерминации (показатель достоверности аппроксимации).

Таблица 5.1 – Результаты измерения содержания жира (X) и белка (Y) в молоке коров (в %)

Номер пробы	1	2	3	4	5	6	7	8	9	10
Содержание жира, %	3,5	2,8	3,6	4,0	3,2	3,1	3,5	3,7	3,9	4,0
Содержание белка, %	3,4	2,8	3,4	3,8	3,1	3,0	3,3	3,5	3,5	3,9

Решение:

- 1) создайте файл MS Excel с именем «Группа_ФИО_Тема 5», в котором Лист 1 переименуйте «Задание 5.1» и задайте необходимые параметры форматирования текста в ячейках электронной таблицы;
- 2) в файле создайте таблицу 5.1 с исходными данными;
- 3) для выполнения расчетов постройте вспомогательную таблицу 5.2;

4) в файле для расчета параметров модели регрессии постройте таблицу 5.3, в тетради запишите формулы и вычислите параметры уравнения регрессии (используя формулы (5.3) - (5.21) из учебного пособия [1]), все необходимые промежуточные расчеты выполните в файле в таблице 5.2 (в столбцах 2-7) и перепишите результаты в рабочую тетрадь (для расчетов в строке «Итого» используйте функцию СУММ(Массив), в строке «Среднее» - функцию СРЗНАЧ(Массив)):

$$\bar{x} = \frac{\sum x_i}{n} = \frac{\dots}{\dots} = \dots ;$$

$$\bar{y} = \frac{\sum y_i}{n} = \frac{\dots}{\dots} = \dots ;$$

$$b = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} = \frac{\dots}{\dots} = \dots ;$$

$$a = \bar{y} - b \cdot \bar{x} = \dots - \dots \cdot \dots = \dots .$$

Таблица 5.3 – Параметры модели регрессии

Показатель	Значение
Константа, a	
Коэффициент регрессии, b	

5) запишите получившееся уравнение регрессии, подставив найденные значения параметров a и b в формулу: $\tilde{y}_x = a + b \cdot x$, сформулируйте вывод по полученной модели (раскройте смысл параметра b , стоящего в модели при факторной переменной X):

Модель регрессии, описывающая зависимость уровня белка в молоке от процента содержания жира имеет вид:

$$\tilde{y}_x = \dots + \dots \cdot x.$$

В соответствии с данной моделью, при увеличении жирности коровьего молока (X) на \dots можно ожидать \dots (увеличение/снижение) процента содержания белка в молоке (Y) в среднем на \dots .

б) выполните контроль правильности расчетов параметров уравнения регрессии:

- вычислите теоретические значения зависимой переменной (\tilde{y}_x), подставив в найденное уравнение регрессии исходные значения факторной переменной (X) (столбец 8 в табл. 5.2);

- вычислите сумму теоретических значений ($\sum \tilde{y}_x$) и сравните ее с суммой фактических исходных значений ($\sum y$), запишите вывод:

Поскольку $\sum \tilde{y}_x$ \dots $\sum y$, то параметры уравнения регрессии вычислены \dots .

7) для расчета показателей силы корреляционной связи между исследуемыми признаками постройте таблицу 5.4:

Рекомендация:

В *MS Excel* для расчета значения парного линейного коэффициента корреляции применяется статистическая функция **КОРРЕЛ(Массив1; Массив2)**, где в качестве аргумента **Массив1** указывают ссылку на блок ячеек со значениями признака X, а в качестве аргумента **Массив2** – ссылку на блок ячеек со значениями признака Y.

Полученное значение парного коэффициента корреляции, равное $r = \underline{\hspace{2cm}}$, указывает на наличие по направлению $\underline{\hspace{2cm}}$, по силе $\underline{\hspace{2cm}}$ корреляционной зависимости между процентом содержания жира и белка в коровьем молоке.

Значение коэффициента корреляции показывает, что $\underline{\hspace{2cm}}\%$ вариации значений уровня содержания белка в коровьем молоке объясняется влиянием процента жирности молока, и $\underline{\hspace{2cm}}\%$ вариации объясняется совокупным влиянием всех прочих факторов.

11) выполните проверку статистической значимости (достоверности) полученной модели регрессии с помощью критерия Стьюдента (t-критерий), вычислив:

- стандартные ошибки коэффициентов регрессии и корреляции (формулы (5.14) и (5.15) в учебном пособии [1]) с точностью до 0,0001:

$$m_b = \underline{\hspace{2cm}} = \underline{\hspace{2cm}} = \underline{\hspace{2cm}} ;$$

$$m_r = \underline{\hspace{2cm}} = \underline{\hspace{2cm}} = \underline{\hspace{2cm}} ;$$

- расчетные значения критерия для этих коэффициентов (формулы (5.12) и (5.13) в учебном пособии [1]) с точностью до 0,001:

$$t_b = \underline{\hspace{2cm}} = \underline{\hspace{2cm}} = \underline{\hspace{2cm}} ;$$

$$t_r = \underline{\hspace{2cm}} = \underline{\hspace{2cm}} = \underline{\hspace{2cm}} ;$$

- для выполнения расчетов в файле постройте таблицу 5.5, в которой при вычислении значений показателей используйте адресные ссылки на соответствующие в таблице 5.2 суммы в итоговой строке, результаты вычислений перепишите в рабочую тетрадь;

Таблица 5.5 – Расчет значений критериев Стьюдента и Фишера

Показатель	Значение
Стандартная ошибка коэффициента регрессии (m_b)	
Стандартная ошибка коэффициента корреляции (m_r)	
t_b	
t_r	
$t_{табл}$	
$F_{расч}$	
$F_{табл}$	
Средняя ошибка аппроксимации (\bar{A}), %	

- определите по таблице значений критерия Стьюдента (Приложение 2) для уровня значимости $\alpha=0,05$ и числа степеней свободы $k = \underline{\hspace{2cm}} = \underline{\hspace{2cm}}$ критическое значение критерия $t_{табл} = \underline{\hspace{2cm}}$ и запишите его в таблицу 5.5;

- в файле для определения табличного значения критерия Стьюдента воспользуйтесь статистической функцией **СТЮДРАСПОБР** (см. рекомендации на стр. 41);
- сравните расчетные и табличное значения критерия и сделайте вывод о статистической значимости коэффициентов регрессии и корреляции:

$|t_b| = |t_r| = \underline{\hspace{2cm}}$, что $\underline{\hspace{2cm}}$ (больше/меньше) $t_{\text{табл}}$, а значит на 5%-ном уровне значимости параметры b и r являются статистически $\underline{\hspace{2cm}}$ (значимыми/не значимыми).

12) выполните проверку статистической значимости (достоверности) полученной модели регрессии с помощью критерия Фишера (F-критерий), вычислив:


- расчетное значение критерия (формула (5.16) в учебном пособии [1]) с точностью до 0,001:

$$F_{\text{расч}} = \underline{\hspace{2cm}} = \underline{\hspace{2cm}} = \underline{\hspace{2cm}} ;$$

- определите по *таблице значений критерия Фишера* (Приложение 4) для уровня значимости $\alpha=0,05$ и числа степеней свободы $k_1 = \underline{\hspace{1cm}}$ и $k_2 = \underline{\hspace{1cm}} = \underline{\hspace{1cm}} = \underline{\hspace{1cm}}$ табличное значение критерия $F_{\text{табл}} = \underline{\hspace{2cm}}$ и запишите его в таблицу 5.5;

- в файле для определения табличного значения критерия Фишера воспользуйтесь статистической функцией **FRASPOBR** (см. рекомендации);

Рекомендация:

 Табличное значение критерия Фишера в *MS Excel* можно получить с помощью функции **FRASPOBR(Вероятность; Степени_свободы1; Степени_свободы2)**, в которой в качестве аргумента **Вероятность** указывается принятое при проверке гипотезы значение уровня значимости $\alpha = 0,05$, а аргументы **Степени_свободы1** – число степеней свободы k_1 , **Степени_свободы2** – число степеней свободы k_2 .

- сравните расчетные и табличное значения критерия и сделайте вывод о статистической значимости модели регрессии:

$F_{\text{расч}} = \underline{\hspace{2cm}}$, что $\underline{\hspace{2cm}}$ (больше/меньше) $F_{\text{табл}}$, а значит на 5%-ном уровне значимости модель регрессии признается статистически $\underline{\hspace{2cm}}$ (значимой/не значимой).

13) проверьте качество полученной модели и пригодность ее для прогнозирования, вычислив в таблице 5.5 среднюю ошибку аппроксимации (формула (5.17) в учебном пособии [1]): запишите формулу, выполните промежуточные расчеты в последнем столбце таблицы 5.2, сделайте вывод о качестве модели:

$$\bar{A} = \underline{\hspace{2cm}} = \underline{\hspace{2cm}} = \underline{\hspace{2cm}} \text{ или } \underline{\hspace{2cm}} \% ;$$

Полученное значение средней ошибки аппроксимации, равное $\underline{\hspace{1cm}}\%$, $\underline{\hspace{1cm}}$ 5-7%, что свидетельствует о $\underline{\hspace{2cm}}$ качестве подбора уравнения регрессии и $\underline{\hspace{2cm}}$ его для прогнозирования.

14) с целью прогнозирования по модели в файле постройте таблицу 5.6 и вычислите точечный прогноз (с точностью до 0,01) предельного уровня содержания белка в пробе молока при нормативном значении жирности сырого коровьего молока, равной 2,8% (в соответствии с ГОСТ 31449-2013 Молоко коровье сырое. Технические условия массовая доля

жира сырого коровьего молока должна быть не менее 2,8%), используя формулу (5.18) из учебного пособия [1] и найденное уравнение регрессии:

Таблица 5.6 – Прогноз уровня содержания белка в коровьем молоке по модели регрессии

Показатель	Значение
Прогнозный уровень жирности молока (x_p), %	
Точечный прогноз уровня белка в молоке по модели регрессии (y_p), %	
Стандартная ошибка уравнения регрессии ($S_{\tilde{y}}$)	
Средняя ошибка прогноза (m_p)	
Коэффициент доверия (t) для надежности $P=0,95$ (по таблице Стьюдента)	
Предельная ошибка прогноза (Δ_p)	
Нижняя граница интервала прогноза	
Верхняя граница интервала прогноза	

$$\tilde{y}_p = \text{_____} + \text{_____} \cdot \text{_____} = \text{_____} + \text{_____} \cdot \text{_____} = \text{_____} \%$$

В соответствии с моделью при нормативной жирности сырого коровьего молока не менее 2,8% уровень содержания белка в нем составит в среднем не менее _____%.

15) с надежностью $P=0,95$ рассчитайте границы доверительного интервала прогноза (формулы и порядок расчета запишите в тетрадь, расчеты выполните в файле в таблице 5.6, переписав результаты в тетрадь):

- вычислите стандартную ошибку модели регрессии (формула (5.20) в учебном пособии [1]) с точностью до 0,0001:

$$S_{\tilde{y}_x} = \text{_____} = \text{_____} = \text{_____} ;$$

- рассчитайте среднюю ошибку прогноза (формула (5.19) в учебном пособии [1]) с точностью до 0,001:

$$m_p = \text{_____} = \text{_____} = \text{_____} ;$$

- определите по таблице значений критерия Стьюдента (Приложение 2) для уровня значимости $\alpha=0,05$ и числа степеней свободы $k = \text{_____} = \text{_____}$ значение коэффициента доверия $t = \text{_____}$ и запишите его в таблицу 5.5 тетради;

- в файле для определения табличного значения критерия Стьюдента воспользуйтесь статистической функцией **СТЮДРАСПОБР** (см. рекомендации на стр. 41);

- вычислите с точностью до 0,001 предельную ошибку прогноза по формуле:

$$\Delta_p = t \cdot m_p = \text{_____} = \text{_____} ;$$

- вычислите по формулам границы доверительного интервала прогноза (с точностью до 0,01):

$$\text{нижняя граница интервала: } \tilde{y}_p - \Delta_p = \text{_____} = \text{_____} ;$$

$$\text{верхняя граница интервала: } \tilde{y}_p + \Delta_p = \text{_____} = \text{_____} .$$

Таким образом, с надежностью 0,95 можно утверждать, что в соответствии с найденной моделью регрессии и нормативным значением жирности сырого коровьего молока, равным 2,8%, среднее значение уровня содержания белка в молоке составит от _____% до _____%.

1б) представьте графически исследуемую зависимость между признаками, построив в файле **поле корреляции** с помощью **Мастера диаграмм** в *MS Excel* (тип диаграммы «Точечная»):

- выделите блок ячеек со значениями признаков – сначала значения X (содержание жира), потом Y (содержание белка), в **Меню** выберите команду **Вставка→Диаграммы→Точечная**;

- оформите рисунок в соответствии с требованиями (формат шрифта, подписи и параметры осей, название рисунка в соответствии с рисунком 5.1);

- покажите линию регрессии и ее уравнение, а также величину коэффициента детерминации (показатель достоверности аппроксимации) на поле графика: щелчком *правой кнопки мыши* по любой из точек вызовите диалоговое окно, в котором выберите команду «Добавить линию тренда...», в окне «**Параметры линии тренда**» выберите тип линии «**Линейная**», установив флажки в строках «**показывать уравнение на диаграмме**» и «**поместить на диаграмму величину достоверности аппроксимации**», нажмите **ОК**;

- перенесите изображение из файла в рабочую тетрадь (рис. 5.1).

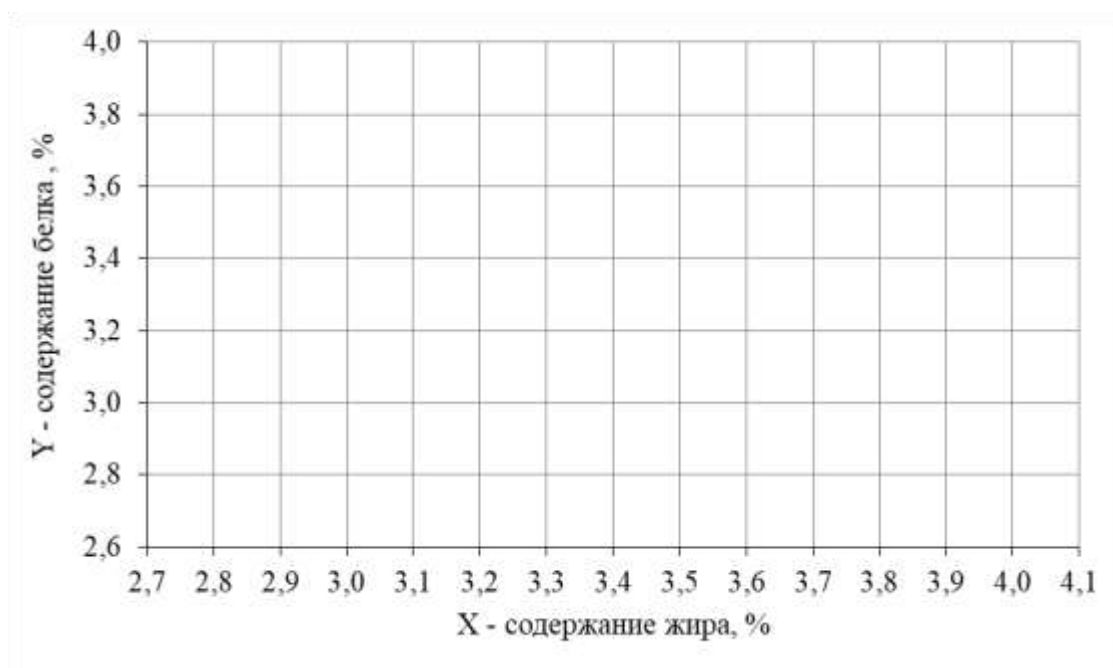


Рисунок 5.1 – Поле корреляции между жирностью (X, %) и уровнем содержания белка (Y, %) в коровьем молоке

Задание 5.2

По данным, представленным в таблице 5.1, выполните корреляционно-регрессионный анализ зависимости между содержанием жира (X) и белка (Y) в молоке коров (в %), используя **Пакет анализа** в *MS Excel* (см. рекомендации, а также стр. 66-70 учебного пособия [1]). Сравните результаты вывода итогов статистической обработки с результатами, полученными в задании 5.1, выделите цветом ячейки со значениями показателей, которые в рамках корреляционно-регрессионного анализа вычисляет инструмент «**Регрессия**».

Рекомендация:



В MS Excel для осуществления однофакторного корреляционно-регрессионного анализа применяется в **Пакете анализа** инструмент **Регрессия**, в окне которого в качестве **Входной интервал Y** указывают блок ячеек со значениями зависимой переменной *Y*, а в качестве **Входной интервал X** – блок ячеек с соответствующими значениями факторной переменной *X*. При включении в указываемые блоки ячеек названий переменных в строке **Метки** необходимо установить флажок . В строке **Уровень надежности** также следует установить флажок , указав значение 95% (в этом случае уровень значимости при применении критериев Стьюдента и Фишера составит 0,05 или 5%). В разделе **Параметры вывода** для вывода отчета о статистической обработке можно выбрать **Выходной интервал**, выделив при этом пустой блок ячеек на текущем рабочем листе, либо выбрать **Новый рабочий лист**, тогда результаты анализа будут представлены на другом рабочем листе, автоматически созданным программой.

Решение:

1) в файле MS Excel с именем «Группа_ФИО_Тема 5» Лист 2 переименуйте «**Задание 5.2**» и задайте необходимые параметры форматирования текста в ячейках электронной таблицы;

2) с Листа «**Задание 5.1**» скопируйте и вставьте таблицу 5.1 с исходными данными, используя вариант вставки с транспонированием (поскольку инструмент **Регрессия** работает корректно только для рядов данных представленных в столбцах);

3) в **Меню** перейдите в раздел **Данные** (или **Сервис**), выберите команду **Анализ данных**, в диалоговом окне **Пакета анализа** выберите инструмент **Регрессия**;

4) в окне **Регрессия** задайте (рис. 5.2):

- **Входной интервал Y** (выделите в таблице 5.1 столбец со значениями содержания белка, включая ячейку с названием столбца) и **Входной интервал X** (выделите в таблице 5.1 столбец со значениями содержания жира, включая ячейку с названием столбца);

- в строке **Метки** установите флажок , в строке **Уровень надежности** установите флажок , указав значение 95% ;

- в разделе **Параметры вывода** выберите **Выходной интервал**, выделив при этом пустой блок ячеек на текущем рабочем листе ниже таблицы 5.1;

Номер пробы	Содержание жира (X), %	Содержание белка (Y), %
1	3,5	3,4
2	2,8	2,8
3	3,6	3,4
4	4,0	3,8
5	3,2	3,1
6	3,1	3,0
7	3,5	3,3
8	3,7	3,5
9	3,9	3,5
10	4,0	3,9

Регрессия

Входные данные:
Входной интервал Y: \$C\$2:\$C\$12
Входной интервал X: \$B\$2:\$B\$12
 Метки
 Уровень надежности: 95 %
 Константа - ноль

Параметры вывода:
 Выходной интервал: \$A\$14:\$D\$14
 Новый рабочий лист:
 Новая рабочая книга

Остатки:
 Остатки
 Стандартизованные остатки
 График остатков
 График подбора

Нормальная вероятность:
 График нормальной вероятности

Рисунок 5.2 – Пример установки параметров в окне «Регрессия» Пакета анализа в MS Excel

5) сравните результаты вывода итогов статистической обработки с результатами, полученными в задании 5.1, выделите цветом ячейки со значениями показателей, которые в рамках корреляционно-регрессионного анализа вычисляет инструмент «Регрессия»;

б) по показателям, выделенным цветом, сформулируйте общий вывод о результатах построения модели регрессии, описывающей зависимость между признаками:

Результаты анализа показывают, что статистическая зависимость между исследуемыми признаками X (содержание жира, %) и Y (содержание белка, %) может быть описана с помощью регрессионной модели:

$$\tilde{y}_x = \underline{\hspace{2cm}} + \underline{\hspace{2cm}} \cdot x,$$

которая указывает на то, что при увеличении значения жирности молока на _____ можно ожидать _____ значения уровня содержания белка в среднем на _____.

Статистическая значимость и надежность данной модели подтверждается как критерием Фишера: $F = \underline{\hspace{2cm}}$, что будет _____ (больше/меньше) табличного значения, так как значимость значения F равна _____, что _____ (не превышает/превышает) выбранный уровень значимости $\alpha = 0,05$; так и критерием Стьюдента: значения t -статистики для параметров регрессии соответственно равны $t_a = \underline{\hspace{2cm}}$ и $t_b = \underline{\hspace{2cm}}$, _____ (превышают/не превышают) табличное значение критерия, так как значимость значений t (P -значение) соответственно равна $P(t_a) = \underline{\hspace{2cm}}$ и $P(t_b) = \underline{\hspace{2cm}}$, что _____ (меньше/больше) выбранного уровня значимости $\alpha = 0,05$.

Описываемая корреляционная зависимость между признаками является по направлению _____ (т.к. $b \underline{\hspace{1cm}} 0$) и по силе _____ ($R = \underline{\hspace{2cm}}$), при этом изменчивость значений уровня содержания белка в молоке на _____% объясняется влиянием фактора - уровня жирности молока, включенного в данную модель ($R^2 = \underline{\hspace{2cm}}$ или _____%). Результаты дисперсионного анализа _____ (подтверждают/не подтверждают) существенность влияния этого фактора на уровень содержания белка в сыром коровьем молоке по сравнению с влиянием всех прочих факторов, не учтенных в модели.

Самостоятельная работа

Задание 5.3

В соответствии с номером варианта, указанным преподавателем, выполните *Индивидуальную работу №3 по теме «Исследование зависимостей на основе корреляционно-регрессионного анализа медико-биологических данных»*, представленную в методичке [2] на стр. 39. Для выполнения расчетов в MS Excel создайте файл с именем «Группа_ФИО_ИР №3_Вариант». Рукопись работы можно выполнить в тетради или в виде текстового файла MS Word (на титульном листе обязательно укажите ФИО, номер группы и номер варианта работы). Работа может быть сдана в печатном (рукописном) виде или в электронном виде (текстовый файл формата PDF) в комплекте с расчетным файлом MS Excel. Срок выполнения работы две недели после выдачи задания.

Список рекомендуемой литературы

1. Шихова, О.А. Математическая биостатистика: Учебное пособие / О.А. Шихова. – Вологда – Молочное: ИЦ Вологодской ГМХА, 2016. – 90 с.
2. Математическая биостатистика: Методические указания и индивидуальные задания / Сост. О.А. Шихова. – Вологда – Молочное: ФГБОУ ВО Вологодская ГМХА, 2017. – 45 с.
3. Козлов, А.Ю. Статистический анализ данных в MS Excel: Учебное пособие / А.Ю. Козлов, В.С. Мхитарян, В.Ф. Шишов - М.: НИЦ ИНФРА-М, 2019. - 320 с. – Режим доступа: <https://znanium.com/read?id=329892>.
4. Хуснутдинов, Р.Ш. Математическая статистика: учеб. Пособие / Р.Ш. Хуснутдинов. – М.: ИНФРА-М, 2019. – 205 с. – Режим доступа: <https://znanium.com/read?id=354383>.

Приложение 1
Значения интегральной функции стандартного нормального распределения

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$$

x	Φ(x)	x	Φ(x)	x	Φ(x)	x	Φ(x)
0,00	0,0000	0,41	0,1591	0,82	0,2939	1,23	0,3907
0,01	0,0040	0,42	0,1628	0,83	0,2967	1,24	0,3925
0,02	0,0080	0,43	0,1664	0,84	0,2995	1,25	0,3944
0,03	0,0120	0,44	0,1700	0,85	0,3023	1,26	0,3962
0,04	0,0160	0,45	0,1736	0,86	0,3051	1,27	0,3980
0,05	0,0199	0,46	0,1772	0,87	0,3078	1,28	0,3997
0,06	0,0239	0,47	0,1808	0,88	0,3106	1,29	0,4015
0,07	0,0279	0,48	0,1844	0,89	0,3133	1,30	0,4032
0,08	0,0319	0,49	0,1879	0,90	0,3159	1,31	0,4049
0,09	0,0359	0,50	0,1915	0,91	0,3186	1,32	0,4066
0,10	0,0398	0,51	0,1950	0,92	0,3212	1,33	0,4082
0,11	0,0438	0,52	0,1985	0,93	0,3238	1,34	0,4099
0,12	0,0478	0,53	0,2019	0,94	0,3264	1,35	0,4115
0,13	0,0517	0,54	0,2054	0,95	0,3289	1,36	0,4131
0,14	0,0557	0,55	0,2088	0,96	0,3315	1,37	0,4147
0,15	0,0596	0,56	0,2123	0,97	0,3340	1,38	0,4162
0,16	0,0636	0,57	0,2157	0,98	0,3365	1,39	0,4177
0,17	0,0675	0,58	0,2190	0,99	0,3389	1,40	0,4192
0,18	0,0714	0,59	0,2224	1,00	0,3413	1,41	0,4207
0,19	0,0753	0,60	0,2257	1,01	0,3438	1,42	0,4222
0,20	0,0793	0,61	0,2291	1,02	0,3461	1,43	0,4236
0,21	0,0832	0,62	0,2324	1,03	0,3485	1,44	0,4251
0,22	0,0871	0,63	0,2357	1,04	0,3508	1,45	0,4265
0,23	0,0910	0,64	0,2389	1,05	0,3531	1,46	0,4279
0,24	0,0948	0,65	0,2422	1,06	0,3554	1,47	0,4292
0,25	0,0987	0,66	0,2454	1,07	0,3577	1,48	0,4306
0,26	0,1026	0,67	0,2486	1,08	0,3599	1,49	0,4319
0,27	0,1064	0,68	0,2517	1,09	0,3621	1,50	0,4332
0,28	0,1103	0,69	0,2549	1,10	0,3643	1,51	0,4345
0,29	0,1141	0,70	0,2580	1,11	0,3665	1,52	0,4357
0,30	0,1179	0,71	0,2611	1,12	0,3686	1,53	0,4370
0,31	0,1217	0,72	0,2642	1,13	0,3708	1,54	0,4382
0,32	0,1255	0,73	0,2673	1,14	0,3729	1,55	0,4394
0,33	0,1293	0,74	0,2704	1,15	0,3749	1,56	0,4406
0,34	0,1331	0,75	0,2734	1,16	0,3770	1,57	0,4418
0,35	0,1368	0,76	0,2764	1,17	0,3790	1,58	0,4429
0,36	0,1406	0,77	0,2794	1,18	0,3810	1,59	0,4441
0,37	0,1443	0,78	0,2823	1,19	0,3830	1,60	0,4452
0,38	0,1480	0,79	0,2852	1,20	0,3849	1,61	0,4463
0,39	0,1517	0,80	0,2881	1,21	0,3869	1,62	0,4474
0,40	0,1554	0,81	0,2910	1,22	0,3888	1,63	0,4484

Окончание приложения 1

Значения интегральной функции стандартного нормального распределения

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$$

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
1,64	0,4495	2,05	0,4798	2,46	0,4931	2,87	0,4979
1,65	0,4505	2,06	0,4803	2,47	0,4932	2,88	0,4980
1,66	0,4515	2,07	0,4808	2,48	0,4934	2,89	0,4981
1,67	0,4525	2,08	0,4812	2,49	0,4936	2,90	0,4981
1,68	0,4535	2,09	0,4817	2,50	0,4938	2,91	0,4982
1,69	0,4545	2,10	0,4821	2,51	0,4940	2,92	0,4982
1,70	0,4554	2,11	0,4826	2,52	0,4941	2,93	0,4983
1,71	0,4564	2,12	0,4830	2,53	0,4943	2,94	0,4984
1,72	0,4573	2,13	0,4834	2,54	0,4945	2,95	0,4984
1,73	0,4582	2,14	0,4838	2,55	0,4946	2,96	0,4985
1,74	0,4591	2,15	0,4842	2,56	0,4948	2,97	0,4985
1,75	0,4599	2,16	0,4846	2,57	0,4949	2,98	0,4986
1,76	0,4608	2,17	0,4850	2,58	0,4951	2,99	0,4986
1,77	0,4616	2,18	0,4854	2,59	0,4952	3,00	0,4987
1,78	0,4625	2,19	0,4857	2,60	0,4953	3,01	0,4987
1,79	0,4633	2,20	0,4861	2,61	0,4955	3,02	0,4987
1,80	0,4641	2,21	0,4864	2,62	0,4956	3,03	0,4988
1,81	0,4649	2,22	0,4868	2,63	0,4957	3,04	0,4988
1,82	0,4656	2,23	0,4871	2,64	0,4959	3,05	0,4989
1,83	0,4664	2,24	0,4875	2,65	0,4960	3,06	0,4989
1,84	0,4671	2,25	0,4878	2,66	0,4961	3,07	0,4989
1,85	0,4678	2,26	0,4881	2,67	0,4962	3,08	0,4990
1,86	0,4686	2,27	0,4884	2,68	0,4963	3,09	0,4990
1,87	0,4693	2,28	0,4887	2,69	0,4964	3,10	0,4990
1,88	0,4699	2,29	0,4890	2,70	0,4965	3,11	0,4991
1,89	0,4706	2,30	0,4893	2,71	0,4966	3,12	0,4991
1,90	0,4713	2,31	0,4896	2,72	0,4967	3,13	0,4991
1,91	0,4719	2,32	0,4898	2,73	0,4968	3,14	0,4992
1,92	0,4726	2,33	0,4901	2,74	0,4969	3,15	0,4992
1,93	0,4732	2,34	0,4904	2,75	0,4970	3,16	0,4992
1,94	0,4738	2,35	0,4906	2,76	0,4971	3,17	0,4992
1,95	0,4744	2,36	0,4909	2,77	0,4972	3,18	0,4993
1,96	0,4750	2,37	0,4911	2,78	0,4973	3,19	0,4993
1,97	0,4756	2,38	0,4913	2,79	0,4974	3,20	0,4993
1,98	0,4761	2,39	0,4916	2,80	0,4974	3,21	0,4993
1,99	0,4767	2,40	0,4918	2,81	0,4975	3,22	0,4994
2,00	0,4772	2,41	0,4920	2,82	0,4976	3,23	0,4994
2,01	0,4778	2,42	0,4922	2,83	0,4977	3,24	0,4994
2,02	0,4783	2,43	0,4925	2,84	0,4977	3,25	0,4994
2,03	0,4788	2,44	0,4927	2,85	0,4978	3,26	0,4994
2,04	0,4793	2,45	0,4929	2,86	0,4979	3,27	0,4995

Приложение 2
Распределение Стьюдента (t – распределение) для уровня значимости α и
числа степеней свободы k

k	<i>Вероятность α</i>								
	<i>0,5</i>	<i>0,4</i>	<i>0,3</i>	<i>0,2</i>	<i>0,1</i>	<i>0,05</i>	<i>0,02</i>	<i>0,01</i>	<i>0,001</i>
1	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,657	636,619
2	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925	31,598
3	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,841	12,941
4	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604	8,610
5	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,043	6,859
6	0,718	0,906	1,134	1,440	1,953	2,447	3,143	3,707	5,959
7	0,711	0,896	1,119	1,415	1,895	2,365	2,998	3,499	5,405
8	0,706	0,889	1,108	1,397	1,860	2,306	2,896	3,355	5,041
9	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250	4,781
10	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169	4,583
11	0,697	0,876	1,088	1,363	1,796	2,201	2,718	3,106	4,437
12	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055	4,318
13	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012	4,221
14	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977	4,140
15	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947	4,073
16	0,690	0,865	1,071	1,337	1,746	2,120	2,583	2,921	4,015
17	0,689	0,863	1,069	1,333	1,740	2,110	2,567	2,898	3,965
18	0,688	0,862	1,067	1,330	1,734	1,101	2,552	2,878	3,922
19	0,688	0,861	1,066	1,328	1,729	2,093	2,539	2,861	3,833
20	0,687	0,860	1,064	1,325	1,725	2,086	2,528	2,845	3,850
21	0,686	0,859	1,063	1,323	1,721	2,080	2,518	2,831	3,819
22	0,686	0,858	1,061	1,321	1,717	2,074	2,508	2,819	3,792
23	0,685	0,868	1,060	1,319	1,714	2,069	2,500	2,807	3,767
24	0,685	0,857	1,059	1,318	1,711	2,064	2,402	2,797	3,745
25	0,684	0,856	1,058	1,316	1,708	2,060	2,485	2,787	3,725
26	0,684	0,856	1,058	1,315	1,706	2,056	2,479	2,778	3,707

Приложение 3
Значения U -критерия Уилкоксона (Манна-Уитни) для уровня
значимости $\alpha=0,05$ и объема выборок n_1 и n_2

n_1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
n_2																			
3	...	0																	
4	...	0	1																
5	0	1	2	4															
6	0	2	3	5	7														
7	0	2	4	6	8	11													
8	1	3	5	8	10	13	15												
9	1	4	6	9	12	15	18	21											
10	1	4	7	11	14	17	20	24	27										
11	1	5	8	12	16	19	23	27	31	34									
12	2	5	9	13	17	21	26	30	34	38	42								
13	2	6	10	15	19	24	28	33	37	42	47	51							
14	3	7	11	16	21	26	31	36	41	46	51	56	61						
15	3	7	12	18	23	28	33	39	44	50	55	61	66	72					
16	3	8	14	19	25	30	36	42	48	54	60	65	71	77	83				
17	3	9	15	20	26	33	39	45	51	57	64	70	77	83	89	96			
18	4	9	16	22	28	35	41	48	55	61	68	75	82	88	95	102	109		
19	4	10	17	23	30	37	44	51	58	65	72	80	87	94	101	109	116	123	
20	4	11	18	25	32	39	47	54	62	69	77	84	92	100	107	115	123	130	138

Приложение 4
Значения F- критерия Фишера
 для $\alpha=0,05$ и чисел степеней свободы k_1 и k_2

k_2	k_1									
	1	2	3	4	5	6	7	8	9	10
1	161	200	216	225	230	234	237	239	240	242
2	18,5	19,0	19,2	19,2	19,3	19,3	19,3	19,4	19,4	19,4
3	10,10	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,47
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99
120	3,92	3,07	2,68	2,45	2,29	2,17	2,09	2,02	1,96	1,91
∞	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,83	1,83

СОДЕРЖАНИЕ

Цели и задачи дисциплины, ее место в учебном процессе	3
Вводный инструктаж.....	4
Тема 1 Анализ медико-биологических данных на основе их группировки и графического представления	5
Аудиторная работа	5
Самостоятельная работа	11
Тема 2 Оценка характеристик генеральной совокупности по выборке	16
Аудиторная работа	16
Самостоятельная работа	26
Тема 3 Нормальный закон распределения.....	30
Аудиторная работа	30
Самостоятельная работа	34
Тема 4 Статистические гипотезы и их проверка	38
Аудиторная работа	38
Самостоятельная работа	46
Тема 5 Исследование зависимостей на основе корреляционно-регрессионного анализа медико-биологических данных	47
Аудиторная работа	47
Самостоятельная работа	56
Список рекомендуемой литературы.....	57
Приложение 1 Значения интегральной функции стандартного нормального распределения	58
Приложение 2 Распределение Стьюдента (t – распределение) для уровня значимости α и числа степеней свободы k	60
Приложение 3 Значения U -критерия Уилкоксона (Манна-Уитни) для уровня значимости $\alpha=0,05$ и объема выборок n_1 и n_2	61
Приложение 4 Значения F - критерия Фишера.....	62

Ответственный за выпуск О.А. Шихова

Заказ № -Р. Тираж 100 экз. Подписано в печать 2021 г.
ФГБОУ ВО Вологодская ГМХА 160555, г. Вологда, с. Молочное, ул. Емельянова, 1