

**ФГБОУ ВО НОВОСИБИРСКИЙ ГАУ**  
**КАФЕДРА ВЕТЕРИНАРНОЙ ГЕНЕТИКИ И БИОТЕХНОЛОГИИ**

**ПРИКЛАДНАЯ СТАТИСТИКА**

**Учебное пособие**

**Новосибирск 2024**

УДК 519.2

ББК 22.17

Составитель: доктор биол. наук, доцент *М.Л. Кочнева*

Рецензент: канд. с.-х. наук, доцент *И.В. Кондратьева*

Прикладная статистика / М.Л. Кочнева / Учебно-методическое пособие;  
Новосиб. гос. аграр. ун-т., ИЭиПБ. – Новосибирск, 2024.- 48 с.

В пособии представлены задания, вопросы по темам дисциплины «Прикладная статистика» для лабораторных занятий, а также самостоятельного изучения. Даны темы и краткий план контрольной работы, порядок её выполнения и представления. Имеется словарь терминов. Предназначены для обучающихся по направлению подготовки 19.03.01 Биотехнология всех форм обучения.

Практикум утвержден и рекомендован к изданию учебно-методическим советом Института экологической и пищевой биотехнологии (протокол № 2 от 12.02 2024г.).

## Оглавление

Введение .....	4
Условные обозначения .....	<b>Ошибка! Закладка не определена.</b>
1. Первичная обработка экспериментальных данных .....	6
2. Типы распределений объектов и их закономерности .....	10
3. Проверка статистических гипотез .....	16
4. Зависимость случайных величин .....	24
5. Анализ качественной изменчивости .....	30
6. Дисперсионный анализ .....	32
7. Ситуационные задачи .....	35
8. Требования к написанию и защиты контрольной работы .....	36
9. Словарь терминов .....	41
Приложения .....	44

## Введение

Цель предлагаемого пособия состоит в освещении не только основных методов прикладной статистики, но и принципов их применения, так, чтобы каждый опыт был спланирован оптимальным образом и из его результатов максимально полно была бы извлечена важная информация.

В учебном пособии рассмотрены основные понятия математической статистики: генеральная и выборочная совокупности, оценки генеральных и выборочных параметров и их свойства, методы проверки статистических гипотез. Представлены основные методы корреляционного и регрессионного анализов. Изложены основы дисперсионного анализа, позволяющий разложить общую изменчивость признака на отдельные составляющие, обусловленные влиянием анализируемого (-ых) фактора (-ов) или их взаимодействий.

Учебный курс дисциплины «Прикладная статистика» содержит 6 разделов, по каждому из которых осуществляется тестирование. При контроле знаний студентов по любой теме учебной дисциплины предлагается ответить на 5-10 вопросов. Тестовые задания предназначены для оценки успешности овладения обучающимся теоретического и практического материала по изучаемой дисциплине. Вместе с тем тестовые задания также могут быть использованы студентом для самоконтроля, что в определенной степени будет отражать уровень усвоения материалов.

В процессе освоения дисциплины студентам предоставляется возможность провести пробное тестирование, предшествующее основному, с целью устранения ошибок и неточностей в заданиях, а также приблизительной оценки трудности отдельных заданий.

Для оценки знаний используются тестовые задания следующих типов:

а) закрытого (ответы на соответствие, исключение неправильных вариантов, выбор одного или более вариантов правильных ответов);

б) открытого (задание теста, в котором отсутствует варианты ответа и экзаменуемый должен самостоятельно сформулировать ответ).

Контрольная работа является одной из форм оценочных средств текущего контроля освоения обучающимися компетенций. Промежуточный контроль знаний студентов направлен на установление процесса усвоения дисциплины в течение семестра.

## 1. Первичная обработка экспериментальных данных

Дисциплина «Прикладная статистика» направлена на формирование базовых знаний по основам применения статистических методов обработки экспериментальных данных в пищевой промышленности. Основной целью дисциплины является формирование и закрепление системного подхода при получении теоретических и практические знаний в области современных статистических методов обработки данных.

Исходя из цели, в процессе изучения дисциплины решаются следующие задачи:

- изучить математическую основу алгоритмов, используемых при обработке экспериментальных данных в пищевой промышленности;
- научиться составлять репрезентативные выборки, адекватно выбирать методы статистической обработки экспериментальных данных;
- сравнение полученных результатов со стандартами.

**Задание 1.** Для изучения затрат (в рублях) на 100 рублей произведенной продукции обследовано несколько предприятий общественного питания. Результаты исследований представлены ниже:

13,9 15,3 15,0 15,4 14,8 14,9 16,0 14,7 15,8 16,4 15,4 13,6 15,4 13,5 15,2  
15,0 16,6 15,8 13,8 15,1 16,6 16,9 14,7 13,6 16,0 16,0 16,3 14,1 14,8 13,9  
14,7 15,2 15,3 17,1 14,1 14,3 15,6 16,4 16,1 15,9 15,6 15,4 14,9 14,6 15,6  
15,2 13,0 13,7 14,2 15,2 16,6 13,6 14,4 15,7 14,8 15,3 15,1 15,0 15,8 16,8  
13,6 15,0 15,0 12,5 13,9 13,9 16,4 15,4 15,0 15,1 14,8 15,9 14,5 14,4 15,4  
14,8 15,4 15,8 13,4 14,8 15,5 17,3 14,5 13,2 15,4 16,0 14,8 13,8 14,1 15,8  
13,8 14,2 14,6 15,6 14,6 16,7 15,5 14,4 14,7 14,1 16,3 15,5 14,3 13,9 15,7  
14,1 14,8 15,3 17,1 15,8 16,8 14,8 15,4 15,8 15,4 16,6 15,7

1. Создайте базу данных в программе для работы с электронными таблицами. Определите объем выборочной совокупности (n).

2. Рассчитайте среднюю арифметическую. Определите моду, медиану.
3. Укажите границы изменчивости признака.
4. Рассчитайте показатели изменчивости, характеризующие признак.
5. По данным выборки постройте вариационный ряд, распределив значения  $x_i$  по частотам  $n_i$ .
6. Используя вариационный ряд, постройте график распределения объектов.
7. Напишите вывод.

**Задание 2 (для самостоятельной работы).** Проведена оценка жирности молока (%), поставляемого на предприятия общественного питания.

Результаты анализа представлены ниже:

3,86 4,16 3,77 3,97 3,76 3,91 3,96 4,04 3,84 3,94 3,98 3,87 3,87  
 4,07 3,99 3,89 3,86 3,71 3,94 3,82 4,16 3,76 4,04 3,96 4,08 3,88  
 4,01 3,93 3,71 3,81 4,02 4,17 3,72 4,09 3,78 4,02 3,73 3,92 3,85  
 3,92 4,18 4,26 4,03 4,14 3,72 4,33 3,82 4,03 3,82 3,91 3,87 4,23  
 4,12 3,77 3,46 4,43 3,89 3,86 4,01 3,99 3,89 3,97 3,98 4,02 4,01  
 3,93 3,92 3,93 3,90 3,93 4,21 3,93 4,01 3,93 3,71 3,81 4,03 4,14  
 3,72 4,33 3,77 3,97 3,76 3,91

1. Создайте базу данных в программе для работы с электронными таблицами. Определите объем выборочной совокупности ( $n$ ).
2. Рассчитайте среднюю арифметическую. Определите моду, медиану.
3. Укажите границы изменчивости признака.
4. Рассчитайте показатели изменчивости, характеризующие признак.
5. По данным выборки постройте вариационный ряд, распределив значения  $x_i$  по частотам  $n_i$ .
6. Используя вариационный ряд, постройте график распределения объектов.
7. Напишите вывод.

**Задание 3.** Свойства средней арифметической.

1. Каким будет средняя арифметическая, если при исследовании признака в 30 случаях каждая варианта имела значение 8,54?
2. При вычислении средней арифметической восьми вариантов ( $x_1=10$ ,  $x_2=11$ ,  $x_3=14$ ,  $x_4=12$ ,  $x_5=10$ ,  $x_6=15$ ,  $x_7=13$ ,  $x_8=11$ ) получено значение 12. Как изменится величина средней арифметической, если к каждой варианте прибавить 5? А если отнять 4?
3. Как изменится величина средней арифметической ( $\bar{x}=25$ ), если каждую варианту разделить на 2?
4. Чему будет равна сумма отклонений каждой варианты ( $x_1=10$ ,  $x_2=11$ ,  $x_3=14$ ,  $x_4=12$ ,  $x_5=10$ ,  $x_6=15$ ,  $x_7=13$ ,  $x_8=11$ ) от средней арифметической  $\bar{x}=12$ ?

## Текущий контроль

### 1. Малой выборкой принято считать, если её объем составляет:

а)  $< 40$  единиц совокупности; б)  $> 40$  единиц совокупности; в)  $< 30$  единиц совокупности; г)  $> 30$  единиц совокупности.

### 2. Значение отдельно взятого измерения называется \_\_\_\_\_.

### 3. Выборка называется репрезентативной, если она:

а) не дает достаточное представление об особенностях генеральной совокупности; б) дает достаточное представление об особенностях генеральной совокупности; в) не дает достаточное представление об особенностях выборочной совокупности; г) дает достаточное представление об особенностях выборочной совокупности.

### 4. Генеральная совокупность — это:

а) бесконечное множество разнородных объектов; б) конечное множество однородных объектов; в) конечное множество объектов, объединенных по определенному признаку; г) бесконечное множество объектов, объединенных по определенному признаку.

### 5. Ранжированный ряд представляет собой:

а) двойной ряд чисел, состоящий из классов и частот; б) ряд чисел, представляющих собой ранги; в) ряд чисел, расположенных в определенном порядке; г) ряд чисел, представляющих собой отдельные измерения признака.

### 6. Число единиц, входящих в совокупность, называется:

а) предметом совокупности; б) объемом совокупности; в) единицей совокупности; г) группой совокупности.

### 7. К количественным признакам с непрерывной изменчивостью можно отнести:

а) процент бракованных образцов молочной продукции; б) количество роллов в упаковке; в) содержание жира в молоке; г) органолептические показатели молочных продуктов, выраженные в баллах.

### 8. Медианой называют:

а) варианту, расположенную в середине ранжированного или вариационного ряда; б) наиболее часто встречающуюся варианту в ранжированном или вариационном ряду; в) варианту, равную средней арифметической ранжированного или вариационного ряда; г) варианту, расположенную в середине ряда чисел.



**9. Наиболее часто встречающаяся варианта в ранжированном или вариационном ряду – это \_\_\_\_\_.**

**10. К статистическим параметрам изменчивости не относится:**

а) дисперсия; б) медиана; в) стандартное отклонение; г) коэффициент вариации.

**11. Дополните нужное определение:**

отношение стандартного отклонения к среднему арифметическому, выраженное в процентах, это \_\_\_\_\_.

**12. При анализе физико-химических свойств мяса было установлено, что средняя рН образцов составила 4,8, а варианса равна 0,14. Определите коэффициент вариации уровня рН.**

### **Вопросы для самоконтроля**

1. С какой целью отклонение каждой варианты от средней арифметической возводится в квадрат?
2. Дайте классификацию признаков изучаемых объектов.
3. Дайте определение совокупности.
4. В каких пределах лежит вся изменчивость признака относительно средней арифметической?

## 2. Типы распределений объектов и их закономерности

Как известно, признаки делятся на 2 класса (количественные и качественные). Они отличаются друг от друга особенностями варьирования и могут подчиняться тому или иному типу распределения. Каждый тип распределения имеет свои особенности и закономерности и отражает причины варьирования признака и может быть представлен графически в виде кривой.

**Нормальное распределение (Гаусса).** Этот тип распределения является наиболее распространенным. Закон нормального распределения является предельным законом, к которому стремятся другие типы распределения объектов.

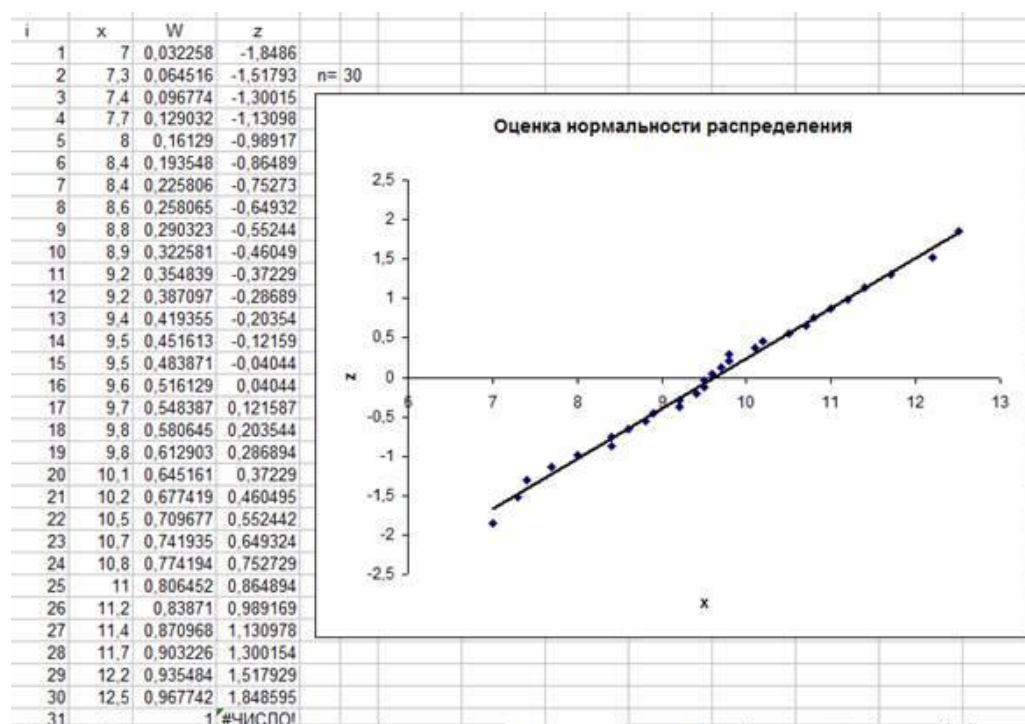
**Асимметричное распределение (распределение Максвелла)** – на графике кривая будет скошена вправо или влево. В тех случаях, когда какие-нибудь причины благоприятствуют более частому появлению значений, которые выше или, наоборот, ниже среднего, образуются асимметричные распределения.

**Экссессивное распределение.** В тех случаях, когда какие-либо причины способствуют преимущественному накоплению частот в классах, расположенных рядом со средним значением признака, образуется распределение с **положительным эксцессом**. Если же в распределении преобладают крайние значения, причем одновременно как более низкие, так и более высокие, то такое распределение характеризуется **отрицательным эксцессом** и в центре распределения может образоваться впадина, превращающая его в двухвершинное.

**Биномиальное распределение.** Это распределение членов совокупности по альтернативным признакам (или дискретное). Т.е. биномиальное распределение характерно для качественных признаков, иногда для количественных признаков с дискретной изменчивостью.

**Распределение Пуассона** – это тип распределения редко происходящих событий. Это распределение дискретно, т.к. частота редких событий очень мала. При графическом изображении распределение Пуассона имеет резко асимметричную форму.

**Задание 1.** Использование критериев согласия и критериев нормальности для оценки вида функции распределения непрерывной случайной величины не всегда удобно, т.к. требует значительных вычислений, а в некоторых случаях большого объёма испытаний. Поэтому вид функции распределения иногда можно оценивать приближёнными методами – графическим или по асимметрии и эксцессу.



В графическом способе результаты испытаний сортируют в ранжированный или вариационный ряд. Для каждого результата  $x_i$  находят накопленную частоту:  $W=i/(n+1)$ , где  $i$  – номер результата в ряду,  $n$  – объём испытаний. Используя данные по затратам на 100 рублей продукции, сортируем данные в порядке от минимального до максимального значения ( $n=105$ ), затем рассчитываем  $W$  по формуле. После этого вычисляем  $z$  (в

Excel статистическая функция НОРМСТОБР, в строку **Вероятность** её диалогового окна вводим ссылку на накопленную частоту **W**).

Затем строим точечную диаграмму по данным столбцов **x** и **z**. Добавляем линейную линию тренда: на какой-либо точке диаграммы открываем контекстное меню и выбираем команду **Добавить линию тренда**.

Если на графике видно, что все точки находятся вблизи прямой, то гипотеза нормальности не отвергается, т.е. фактическое распределение объектов соответствует теоретически ожидаемому (нормальному распределению).

**Задание 2.** Группировка данных выборочной совокупности по количественным признакам.

1. Подготовьте данные для статистического анализа в соответствии с алгоритмом:

- 1) выбор числа классов ( $m$ );
- 2) определение размаха вариационного ряда ( $R = \max - \min$  признака);
- 3) расчет классного промежутка  $k = R/m$
- 4) определение границ классов: начальная граница первого класса – это минимальное значение; начальная граница последующих классов будет отличаться от предыдущей на величину классного промежутка ( $k$ ).

2. Построение гистограммы распределения объектов по уровню изучаемого признака в программе для работы с электронными таблицами.

3. Опишите как распределены объекты в исследованной выборке.

**Задание 3.** Оцените статистические показатели по уровню титруемой кислотности биойогурта ( $^{\circ}T$ ), произведенного на предприятии №1.

Постройте гистограмму распределения объектов по уровню изучаемого признака в программе для работы с электронными таблицами. Напишите вывод.

№ п/п	Кислотность, °Т	№ п/п	Кислотность, °Т
1	76	30	78
2	78	31	80
3	75	32	77
4	78	33	80
5	76	34	78
6	76	35	78
7	76	36	78
8	76	37	78
9	77	38	79
10	78	39	77
11	78	40	79
12	78	41	77
13	78	42	79
14	76	43	78
15	80	44	78
16	79	45	76
17	76	46	78
18	77	47	79
19	78	48	75
20	79	49	80
21	77	50	79
22	78	51	75
23	78	52	80
24	77	53	79
25	77	54	79
26	76	55	78
27	76	56	76
28	76	57	78
29	79	58	78

**Задание 4.** Оценить соответствие фактического распределения образцов молока коров по уровню титруемой кислотности теоретически ожидаемому. Сформулировать нулевую гипотезу и на основании полученных данных сделать вывод.

№ п/п	Кислотность молока, °T ( $x_i$ )	Частота (f)
1	14	6
2	15	17
3	16	24
4	17	25
5	18	20
6	19	8

**Задание 6.** Внесите необходимую информацию в таблицу.

Тип распределения	Признак, характерный для данного типа распределения	Статистический показатель, характеризующий данный тип распределения

### Текущий контроль

**1. Какой тип распределения описывает редко происходящие события:**

- а) распределение Стьюдента;
- б) распределение Гаусса;
- в) распределение Пуассона;
- г) распределение  $\chi^2$ .

**2. Какое условие присуще для положительного асимметричного распределения объектов:**

- а)  $\bar{x} = Me$ ;
- б)  $\bar{x} < Mo$ ;
- в)  $\bar{x} > Me$ ;
- г)  $\bar{x} = Mo$ ;
- д)  $\bar{x} > Mo$ .

**3. Какой тип распределения объектов характеризуется значительным накоплением частот в классах, расположенных рядом со средним значением признака**

- а) двухвершинный эксцесс;
- б) положительный эксцесс;
- в) отрицательный эксцесс.

**4. Распределение членов совокупности по признакам с альтернативной или дискретной изменчивостью:**

- а) нормальное распределение;
- б) биномиальное распределение;
- в) асимметричное распределение;
- г) эксцессивное распределение.

**5. При нормальном распределении практические пределы отклонений от среднего значения признака составляют:**

- а)  $\pm 1,3\sigma$ ;
- б)  $\pm 2,3\sigma$ ;
- в)  $\pm 4,3\sigma$ ;
- г)  $\pm 3,3\sigma$ .

**6. При большой величине стандартного отклонения форма нормальной кривой будет:**

- а) пологой;
- б) крутой;
- в) изогнутой;
- г) многовершинной.

**7. Какой процент вариант совокупности будет находиться в пределах  $\bar{X} \pm 3\sigma$ , если варьирование признака подчиняется закону нормального распределения:**

- а) 95,5;
- б) 99,7;
- в) 68,3;
- г) 99,9.

**8. Какими статистическими параметрами определяется положение кривой нормального распределения:**

- а)  $\bar{x}$  и  $M_e$ ;
- б)  $M_e$  и  $M_o$ ;
- в)  $\sigma$  и  $M_o$ ;
- г)  $x$  и  $\sigma$ .

**9. Какими статистическими параметрами определяется положение кривой биномиального распределения:**

- а)  $\bar{M}_e$ ;
- б)  $\bar{M}_o$ ;
- в)  $\bar{x}$ ;
- г)  $\bar{\sigma}$  и  $\bar{M}_e$ ;
- д)  $\bar{x}$  и  $\bar{\sigma}$ .

### **Вопросы для самоконтроля**

1. С помощью какого статистического критерия оценивают соответствие фактического распределения объектов биномиальному?
2. Какие статистические параметры используют для характеристики нормального распределения?
3. Какие типы распределения объектов характерны для количественных признаков, а какие для качественных?

### **3. Проверка статистических гипотез**

Для характеристики выборочных совокупностей вычисляют параметры  $\bar{X}$ ,  $\sigma$ ,  $\sigma^2$ . Генеральная совокупность характеризуется средним значением  $\mu$  (математическое ожидание), показателями разнообразия.

Из генеральной совокупности можно сформировать несколько выборочных совокупностей, которые будут различаться по средним значениям выборочных параметров. Но в том случае, когда несколько выборочных совокупностей имеют достаточно большой объем, то среднее значение выборочных средних должно совпасть со средней генеральной совокупности, то же можно сказать и о показателях изменчивости.

Несовпадение значений параметров выборочных совокупностей и генеральной будет зависеть от изменчивости признака (чем выше  $\sigma$ , тем сложнее оценка). При увеличении объема выборочных совокупностей их средние значения признаков будут приближаться к значению  $\mu$ .

Таким образом, выборочные параметры (статистические параметры выборочных совокупностей) имеют «точечную» оценку, т.е. выражаются единственным числом, а генеральные параметры оцениваются всегда в доверительных границах.

**Проверка статистических гипотез.** Статистическая гипотеза – это предположение о виде распределения или о величинах неизвестных



параметров генеральной совокупности, которая может быть проверена на основании выборочных показателей.

Например: эмпирическое распределение соответствует теоретически ожидаемому; средние двух выборочных совокупностей равны между собой.

Различают следующие типы статистических гипотез: выдвигаемая или нулевая гипотеза ( $H_0$ ); альтернативная или противоречащая нулевой гипотеза ( $H_1$ ).

Решения о принятии или отклонении статистической гипотезы принимаются по статистическим параметрам выборочных совокупностей, что не исключает возможности ошибочного решения.

Различают параметрические и непараметрические статистические критерии. Особенностью параметрических критериев является предположение о том, что распределение признака в генеральной совокупности подчиняется нормальному закону. Тестирование на нормальность фактического распределения объектов в выборке является обязательным до применения любого параметрического критерия.

**Задание 1.** Определить доверительные границы  $\mu$  при  $P=99,7$ , если известно, что среднее значение признака составило 38, а стандартная ошибка среднего — 2,5.

**Задание 2.** Определить доверительные границы  $\sigma$  при  $P=95,5$ , если известно, что стандартное отклонение ( $\sigma$ ) составило 32,2, а объем совокупности был равен 2000.

**Задание 3.** Определить доверительные границы  $C_v$  при  $P=99,9$ , если известно, что стандартное отклонение ( $\sigma$ ) составило 17,8, а среднее значение признака было равно 634, объем совокупности был равен 1000.

**Задание 4.** При определении pH образцов свинины были получены следующие значения: 6,65; 6,60; 6,25; 5,90; 6,85; 5,95; 6,85; 6,35; 5,99; 6,82;

6,47; 6,90. Рассчитать среднее значение  $pH$  и доверительные границы генеральной совокупности ( $\mu$ ) при  $P=99,7$ .

**Задание 5.** По одной и той же рецептуре изготовлено кондитерское изделие на 6-ти предприятиях города. Товар поступил в продажу в один супермаркет. За месяц продаж изделия первого предприятия купило 429 человек, второго предприятия – 468, третьего – 405, четвертого - 501, пятого — 466, шестого – 491.

Оцените, соответствует ли фактическое распределение покупательских предпочтений по качеству выпускаемого изделия теоретически ожидаемому.

Как можно определить ожидаемое число посетителей?

**Задание 6.** Заполните таблицу.

Определение статистического параметра	Обозначение статистического параметра
Объем совокупности	
Стандартная ошибка средней арифметической	
Средняя арифметическая генеральной совокупности	
Средняя арифметическая	
Варианта	
Стандартное отклонение	
Медиана	
Число степеней свободы	
Варианса	
Стандартная ошибка стандартного отклонения	
Дисперсия	
Пределы варьирования признака	
Коэффициент вариации	
Мода	
Уровень вероятности	

**Задание 7.** Оцените достоверность разности между двумя предприятиями по содержанию усвояемых белков в зерне сои. Напишите вывод.

№ п/п	Содержание усвояемых белков в зерне сои предприятия №1 (% на воздушносухое вещество)	Содержание усвояемых белков в зерне сои предприятия №2 (% на воздушносухое вещество)
1	30,7	29,5
2	29,5	29,3
3	27,4	27,2
4	27,6	27,4
5	27,5	27,3
6	27,9	27,7
7	27,6	27,4
8	28,6	28,4
9	28,6	28,4
10	28,1	27,9
11	28,9	28,7
12	29,2	29,1
13	28,6	28,4
14	28,9	28,7
15	29,7	27,3
16	29,1	28,9
17	27,5	27,3
18	28,8	28,6
19	29,9	29,7
20	28,8	28,6
21	29,4	29,2
22	29,8	29,6
23	28,73	28,53
24	30,5	28,3
25	30,8	28,3
26	29,5	28,3
27	28,5	28,3
28	28,1	27,9
29	29,1	28,9

**Задание 8.** Проведена оценка массовой доли сухих веществ (%) и массовой доли влаги (%) образцов киселя с использованием коровьего и миндального молока. Оцените различия между произведенным продуктом с использованием разного вида молока с помощью U-критерия Манна-Уитни (стандартные значения критерия приложение 3). Сделайте вывод.

Массовая доля влаги, %		Массовая доля сухих веществ, %	
кисель на основе коровьего молока	кисель на основе миндального молока	кисель на основе коровьего молока	кисель на основе миндального молока
40	45	60	57
42	46	61	54
40	43	60	55
41	48	63	56
45	49	62	61
43	47	65	55
44	46	64	

**Задание 9.** Проведена дегустационная оценка органолептических качеств (аромат, цвет) киселя с использованием коровьего и миндального молока. Оцените различия между произведенным продуктом с использованием разного вида молока по аромату и цвету на основании баллов дегустаторов с помощью U-критерия Манна-Уитни (стандартные значения критерия приложение 3). Сделайте вывод.

Дегустационный анализ по аромату, баллы		Дегустационный анализ по цвету, баллы	
кисель на основе коровьего молока	кисель на основе миндального молока	кисель на основе коровьего молока	кисель на основе миндального молока
9	7	8	8
9	8	8	9
8	8	9	9
10	8	8	10
9	8	8	9
10	9	9	9
10	10	10	10

## Тестовые задания

### 1. Статистический параметр $\mu$ определяется посредством:

- а) доверительных границ истинного среднего значения признака генеральной совокупности;
- б) точечного значения средней арифметической признака генеральной совокупности;
- в) доверительных границ истинного среднего значения признака выборочной совокупности;
- г) точечного значения средней арифметической признака выборочной совокупности.

### 2. Стандартная ошибка среднего возрастает:

- а) с увеличением объёма выборочной совокупности;
- б) с увеличением изменчивости признака объектов выборочной совокупности;
- в) с уменьшением объёма выборочной совокупности.

### 3. Какая из гипотез не является нулевой ( $H_0$ ):

- а) различие между  $\bar{x}_1$  и  $\bar{x}_2$ , случайные и генеральные параметры  $\bar{x}_1$  и  $\bar{x}_2$ , равны между собой;
- б) фактическое распределение случайной величины соответствует теоретически ожидаемому;
- в) связь между изучаемыми признаками является достоверной;
- г) анализируемый фактор не оказывает влияния на изменчивость признака.

### 4. Уровень значимости $P=0,05$ означает, что:

- а) из 100 повторений опыта в 0,5 случаях будут получены ошибочные результаты;
- б) из 100 повторений опыта в 5 случаях будут получены ошибочные результаты;

в) из 100 повторений опыта в 95 случаях будут получены ожидаемые результаты.

**5. Какой процент выборочных средних находится в пределах  $\pm$  трех стандартных отклонений от средней арифметической?**

**6. Какой из методов не относится к непараметрической статистике:**

- а) критерий Стьюдента;
- б) критерий Манна-Уитни;
- в) критерий Колмогорова-Смирнова;
- г) критерий Вилкоксона.

**7. Разность между средними значениями признака ( $\bar{x}_1$  и  $\bar{x}_2$ ) в двух выборочных совокупностях достоверна и отрицательна. Это означает, что:**

- а)  $\bar{x}_1 > \bar{x}_2$ ;
- б)  $\bar{x}_1 < \bar{x}_2$ ;
- в)  $\bar{x}_1 = \bar{x}_2$ .

**8. Для оценки достоверности разности между средними значениями двух выборочных совокупностей, подчиняющихся нормальному распределению, используют:**

- а) критерий Фишера;
- б) критерий Стьюдента;
- в) дисперсионный анализ;
- г) метод  $\chi^2$ ;
- д) корреляционный анализ.

**9. Фактическое значение критерия  $\chi^2$  составило 12,5 (df=5). Какая гипотеза ( $H_0$  или  $H_1$ ) будет принята?**

**10. Оценивается достоверность разности по содержанию летучих жирных кислот (мг) в свежем мясе двух производителей ( $n_1=105$  и  $n_2=87$ ), использующих разные технологии содержания животных. Фактическое**

значение критерия Стьюдента составило 2,5. Какая гипотеза ( $H_0$  или  $H_1$ ) будет принята?

### **Вопросы для самоконтроля**

1. Перечислите типы статистических гипотез и дайте их характеристику.
2. В чем заключается различие между уровнем значимости и уровнем вероятности?
3. Что такое статистический критерий? Приведите примеры.

## 4. Зависимость случайных величин

Различают следующие типы зависимости: функциональная и корреляционная (статистическая).

**Функциональной** зависимостью называется такая связь, когда при известном значении одной из величин можно точно указать значение другой. Функциональная зависимость может иметь место как между детерминированными (неслучайными) переменными (например, зависимость скорости падения объекта в вакууме от времени), так и между случайными величинами (например, зависимость стоимости реализованных изделий от их количества и т.п.).

**Корреляционная (статистическая) зависимость** — это зависимость между случайными величинами, при которой с изменением одной величины меняется распределение другой (термин “корреляция” впервые ввел Galton, 1888 и происходит он от лат. *correlatio* – соотношение, связь). В отличие от функциональной связи при корреляционной зависимости каждому значению одной переменной соответствует не какое-то определенное, а различное множество возможных значений другой переменной. По мере увеличения тесноты корреляционной зависимости она все ближе будет приближаться к функциональной, а в предельном случае соответствует ей. Крайняя противоположность функциональной связи — полная независимость случайных величин.

Поскольку речь идёт о варьирующих наблюдениях, встаёт вопрос о статистической значимости результата анализа. Для этого необходимо оценить достоверность коэффициента корреляции с помощью t-критерия Стьюдента.

Коэффициент корреляции указывает только на степень связи между признаками. В некоторых случаях необходимо знать характер изменения одного признака в зависимости от изменения другого. Для этих целей используется регрессионный анализ.



**Регрессия** — функция, позволяющая по средней величине одного признака определить среднюю величину другого признака, связанного с первым. Для описания регрессии используют: эмпирические и теоретические ряды, их графики — линии регрессии, а также уравнения регрессии и коэффициент линейной регрессии. Регрессионный анализ даёт возможность графически представить результат — в виде линии (regressio (лат.) — наклонение, склонение), стремящейся максимально точно представить зависимость одной переменной от другой (других).

Различают линейную, нелинейную и множественную регрессию. Линейная регрессия предполагает, что функция зависит от параметров линейно. При этом параметры находят с помощью метода наименьших квадратов. Достоинства: простота алгоритма, быстроедействие. Недостаток: непригодность к решению нелинейных задач.

**Задание 1.** Оценить направление и силу связи между содержанием белка и жира в мясе кур, г в 100 г продукта. Сделайте вывод.

Содержание белка, г в 100 г продукта	Содержание жира, г в 100 г продукта
19	5
18	7
16	9
14	10
19	8
20	5
18	5
19	6
20	4
21	4
19	6
22	4
21	3
20	6
19	5
18	7
25	5

**Задание 2. Оценить связь между сортом мучной продукции и потребительским спросом населения. Сделать вывод.**

Сорт мучной продукции	Потребительский спрос, %
1	57
2	31
1	52
в/с	84
2	32
1	40
в/с	67
2	45
1	70
в/с	81
в/с	82
1	67
2	54
в/с	89
2	43
в/с	91
в/с	89
в/с	91

**Задание 3.** Изучено содержание аминокислот триптофана и оксипролина в образцах говядины. На основании полученных данных вычислите коэффициенты корреляции, детерминации и регрессии. Напишите выводы на основании установленных коэффициентов.

№ п/п	Триптофан, мг%	Оксипролин, мг%
1	421,4	64,5
2	481,6	73,7
3	617,05	94,4
4	301	46,1
5	416,05	48,4
6	511,7	78,3
7	421,4	64,5
8	446,15	53
9	406,35	62,2
10	391,3	59,9
11	466,55	71,4
12	451,5	69,1
13	361,2	55,3
14	436,45	66,8
15	361,2	55,3
16	285,95	43,8
17	346,15	53
18	331,1	50,7
19	376,25	57,6
20	316,05	38,4
21	361,2	55,3
22	421,4	64,5
23	361,2	55,3
24	421,4	64,5
25	331,1	50,7
26	511,7	78,3
27	526,75	75,6
28	436,45	66,8
29	331,1	50,7
30	391,3	59,9
31	316,05	48,4
32	406,35	60,2
33	421,4	64,5
34	346,15	53
35	436,45	66,8
36	361,2	55,3
37	391,3	59,9
38	285,95	43,8
39	316,05	48,4
40	401	57,1

## Тестовые задания

**1. Какова степень связи между признаками при статистически значимой величине коэффициента корреляции  $r = 0,75$ :**

- а) сильная;
- б) слабая;
- в) средняя;
- г) связь отсутствует.

**2. Какой статистический параметр вычисляют по формуле  $r_{x/y} \frac{\sigma_x}{\sigma_y}$  :**

- а) коэффициент внутриклассовой корреляции;
- б) коэффициент регрессии;
- в) коэффициент корреляции Пирсона;
- г) коэффициент корреляции Спирмена;
- д) ранговый коэффициент корреляции.

**3. Величина коэффициента корреляции находится в пределах:**

- а) от  $-10$  до  $+10$ ;
- б) от  $0$  до  $+1$ ;
- в) от  $-1$  до  $0$ ;
- г) от  $-1$  до  $+1$ .

**4. Если коэффициент вариации по одному из признаков будет равен нулю, то коэффициент корреляции между этим признаком и любым другим составит:**

- а)  $r = -1$ ;
- б)  $r = 0$ ;
- в)  $r = +1$ .

**5. Под линейной регрессией понимают:**

- а) зависимость, при которой одинаковые увеличения аргумента  $X$  имеют неодинаковые изменения функции  $Y$ ;

- б) зависимость, при которой неодинаковые приращения аргумента  $X$  имеют неодинаковые изменения функции  $Y$ ;
- в) зависимость, когда при одинаковых приращениях аргумента  $X$  наблюдаются одинаковые изменения функции  $Y$ .

**6. Как называется статистический параметр  $r^2$ ?**

- а) коэффициент детерминации;
- б) коэффициент корреляции Спирмена;
- в) коэффициент корреляции Пирсона;
- д) коэффициент регрессии.

**7. Что означает в уравнении регрессии одна переменная ( $x$ ) является независимой, а другая ( $y$ ) — зависимой?**

- а) одна переменная определяет другую;
- б) по значению одной переменной предсказывается значение другой;
- в) одна переменная не зависит от другой;
- г) по значению одной переменной нельзя предсказать значение другой.

**8. Линейная зависимость переменной  $y$  от переменной  $x$  определяется формулой:**

- а)  $y=a+bx$ ;
- б)  $y=a+bx^2$ ;
- в)  $y=a+bx+yz$ .

**Вопросы для самоконтроля**

1. В каких пределах варьирует значение коэффициента корреляции?
2. Какой коэффициент корреляции используют при определении силы связи между порядковыми (ранговыми) и количественным признаками?
3. С какой целью используется регрессионный анализ?

## 5. Анализ качественной изменчивости

Многие признаки невозможно измерить числом. Здесь мы имеем дело с качественными признаками. Такие признаки нельзя упорядочить. Описание качественных признаков состоит в том, чтобы подчитать число объектов, имеющих одно и то же значение, или определить долю того или иного значения от общего числа объектов выборки. Исходя из этого, средняя арифметическая для качественных признаков отражает **долю или процент** объектов, имеющих тот или иной признак.

**Задание 1.** Выявление потребительских предпочтений. Из числа опрошенных респондентов ( $n=1600$ ), оценивающих привлекательность нескольких видов вареников, выводимых на продажу в супермаркете №1, 198 человек ответили, что не стали бы покупать вареники с картофелем и грибами, 58 человек – вареники с картофелем, 207 – вареники с творогом, 261 – вареники с вишней. Вычислите процент респондентов, которых не отдадут предпочтение продукции супермаркета №1. Сделайте вывод.

**Задание 2.** На мясоконсервном заводе №1 из 498 проанализированных образцов полукопченой колбасы 15 штук имели повышенное содержание фосфатов. Определите статистические параметры данной выборки.

**Задание 3.** На мясоконсервном заводе №2 из 498 проанализированных образцов полукопченой колбасы 5 штук имели повышенное содержание фосфатов. Определите статистические параметры данной выборки. Оцените достоверность различий между заводами по доле бракованных образцов.

**Задание 4.** Оценка достоверности разности между выборочными доля (процентами). С помощью метода  $\phi$ -преобразования Фишера оцените достоверность разности между мясоконсервными комбинатами (см. задания 2 и 3) по уровню брака. Сделайте вывод.

**Задание 5.** Коэффициент корреляции между альтернативными признаками ( $r_a$ ). Оцените направление и силу связи между временем года и

уровнем брака молочной продукции на перерабатывающем предприятии №16. Сделайте вывод.

Период	Брак	Стандарт	Всего
Лето	52 (a)	278 (b)	
Зима	18 (c)	312 (d)	
Всего			

### Тестовые задания

1. При каких условиях необходимо применять фи-преобразования Фишера:

- а)  $p \leq 0,25$  и  $q \geq 0,75$ ;
- б)  $p < 0,25$  и  $q > 0,75$ ;
- в)  $p \leq 0,50$  и  $q \geq 0,50$ ;
- г)  $p = 0,25$  и  $q = 0,75$ .

2. На мясоконсервном заводе №1 из 602 проанализированных образцов мясных продуктов 25 штук имели содержание кадмия, превышающее ПДК. Определите процент бракованных образцов.

### Вопросы для самоконтроля

- 1. Какой метод используют для оценки достоверности между выборочными частотами?
- 2. Дайте определение альтернативные признаки.
- 3. Что отражает средняя арифметическая для качественных признаков?

## 6. Дисперсионный анализ

Любые объекты в процессе жизнедеятельности подвергаются воздействию различных факторов. С помощью методов статистики возможна оценка влияния факторов (направленных или же случайных) на анализируемые признаки. В 1920 году английский математик Рональд Фишер предложил концепцию дисперсионного анализа (Analysis of variance (ANOVA) – анализ дисперсии). Традиционно в нашей стране используется термин дисперсионный анализ. Дисперсия (от лат. *dispergere* – распылять) – мера изменчивости варьирующей величины, вычисляемая как средний квадрат её отклонений от соответствующего среднего.

Основной целью дисперсионного анализа является оценка вклада фактора (-ов) в изменчивость признака. Это достигается посредством установления статистической значимости различий между выборочными средними арифметическими.

Таким образом, с помощью дисперсионного анализа можно определить достоверность влияния изучаемых факторов, достоверность разности между выборочными средними арифметическими, а также силу влияния фактора (в случайной модели).

**Задание 1.** С помощью дисперсионного анализа оцените влияние страны производителя (A1 - Киргизия, A2 - Таджикистан, 3 -Россия) и типа упаковки (B1 - тетрапак, B2 - стеклянная бутылка) на pH Балтийского кефира. На основании полученных данных сделайте вывод.

Повторности	A1		A2		A3	
	B1	B2	B1	B2	B1	B2
1	4,39	4,19	4,31	4,22	4,49	4,25
2	4,41	4,18	4,35	4,29	4,49	4,21
3	4,42	4,27	4,39	4,27	4,45	4,25
4	4,38	4,15	4,37	4,27	4,41	4,27
5	4,43	4,26	4,33	4,26	4,42	4,21



## Тестовые задания

**1. Какой метод позволяет вычленять отдельные источники изменчивости и измерять их влияние:**

- а) ковариационный анализ;
- б) метод Фишера;
- в) дисперсионный анализ;
- г) метод условных отклонений;
- д) корреляционный анализ.

**2. Подберите к параметрам дисперсионного анализа соответствующие им определения.**

- |             |  |
|-------------|--|
| а) $S_{сл}$ | а) коэффициент внутриклассовой корреляции; |
| б) $F$      | б) критерий достоверности Фишера;          |
| в) $r_w$    | в) случайная дисперсия;                    |
| г) $MS$     | г) средний квадрат.                        |

**3. Кто предложил основы дисперсионного анализа (Analysis of variance - ANOVA)?**

- а) Френсис Гальтон;
- б) Карл Пирсон;
- в) Роберт Фишер;
- г) Даниил Бернулли.

**4. Оценить эффект влияния ряда факторов на изменчивость признака, а также изучить их взаимодействие позволяет:**

- а) многофакторный анализ;
- б) регрессионный анализ;
- в) однофакторный анализ;
- г) корреляционный анализ.

**5. Вычислите  $MS_{сл}$ , если известно, что:**

$$S_{сл} = 82,14$$

$$N = 27$$

$$r = 5.$$

Ответ \_\_\_\_\_.

**6. Коэффициент внутриклассовой корреляции ( $r_w$ ) это:**

- а) показатель, оценивающий доля изменчивости, вызванной действием случайных факторов;
- б) показатель, оценивающий долю влияния фактора в изменчивости анализируемого признака;
- в) показатель, оценивающий разнообразие или изменчивость признаков;
- г) показатель, оценивающий связь между признаками.

**7. Что из ниже перечисленного является целью дисперсионного анализа?**

- а) оценка вклада фактора в изменчивость анализируемого признака;
- б) установление степени и направления связи между признаками;
- в) характеристика вариабельности признака;

г) выявление значимости разности между средними значениями признака двух выборочных совокупностей.

**8. На что указывает большая величина коэффициента внутриклассовой корреляции:**

- а) на значительное варьирование вариант внутри градаций дисперсионного комплекса;
- б) на не значительное варьирование вариант внутри градаций дисперсионного комплекса;
- в) на отсутствие варьирования вариант внутри градаций дисперсионного комплекса.

Вопросы для самоконтроля

1. Что собой представляет дисперсионный комплекс?
2. Опишите различия между случайной и фиксированной моделью дисперсионного комплекса.
3. В каком случае можно вычислять коэффициент внутриклассовой корреляции?

## 7. Ситуационные задачи

**Задание 1.** Проанализируйте варианты ответов 79 дегустаторов (табл. 1) при исследовании вкусовых качеств мясной продукции после добавления вкусовой добавки **X**. Какой статистический метод нужно использовать для проверки, отличается ли это распределение от распределения вариантов ответов, где дегустаторы оценивали продукт без вкусовых добавок? Сделайте вывод.

Таблица 1. Распределение баллов, выставленных дегустаторами

Балл	Частота распределения баллов, выставленных дегустаторами	
	добавка <b>X</b>	без добавки
1	30	9
2	20	15
3	15	40
4	11	10
5	4	5

**Задание 2.** Укажите какой статистический метод (действие) используется для решения поставленной задачи.

Задача	Метод (действие)
Найти достоверность разности между средним содержанием уровня кислотности молока, произведенного на двух разных молзаводах.	
Определить разницу по доле бракованных образцов кондитерского изделия на разных предприятиях.	
Определить величину связи между содержанием триптофана и оксипролина в образцах говядины.	
Сравнить степень изменчивости содержания жира в свинине и белка в говядине.	

## 8. Требования к написанию и защиты контрольной работы

Одним из видов самостоятельной работы является выполнение контрольной работы. Вариант контрольной работы определяет преподаватель. Контрольная работа включает три задачи, алгоритм решения которых необходимо отразить в работе. Оформление титульного листа смотри в приложении 4.

### Вариант 1

1. Среднее значение признака составило 12, какими будут доверительные границы математического ожидания:

а) (12; 12,5), б) (11,5; 11,9); в) (11,5; 12,5); д) (11,5; 12). Поясните ответ.

2. Исследовали содержание влаги в образцах вареных колбасных изделий (%) в двух цехах:

1-й цех

64,78; 63,54; 64,75; 63,52; 63,55; 64,78; 63,55; 63,57; 64,65; 63,89; 63,85; 63,24.

2-й цех

65,18; 63,98; 65,15; 64,65; 63,32; 65,38; 64,85; 64,12; 65,67; 65,25; 65,41; 64,87.

Оцените, имеются ли достоверные различия по среднему содержанию влаги в исследованных образцах вареных колбасных изделий между цехами. Сделайте вывод.

3. Определите медиану следующего ряда чисел:

34, 23, 12, 24, 13, 16, 18, 20, 14, 23, 13, 19, 21, 13, 21, 15, 29, 17, 28, 18, 30, 11.

### Вариант 2

1. Если основная гипотеза имеет вид  $H_0 = 20$ , то альтернативной гипотезой может быть:

а)  $H_1 \neq 20$ ; б)  $H_1 \leq 20$ ; в)  $H_1 \geq 20$ ; д)  $H_1 \leq 30$ . Поясните ответ.

2. По данным о содержании летучих жирных кислот (мл 0,2 н. раствора NaOH) в свежем мясе составьте гистограмму с использованием Excel и оцените нормальность распределения.

0,35; 0,50; 0,37; 0,30; 0,32; 0,33; 0,45; 0,44; 0,50; 0,38; 0,44; 0,30; 0,33; 0,44; 0,39; 0,32; 0,39; 0,41; 0,48; 0,41; 0,29; 0,45; 0,30; 0,35; 0,36; 0,35; 0,29; 0,35; 0,34; 0,36; 0,27; 0,28; 0,40; 0,29; 0,35; 0,34; 0,38; 0,39; 0,32; 0,36; 0,47; 0,44; 0,30; 0,29; 0,35.

Сделайте вывод.

3. Определите моду следующего ряда цифр: 3, 5, 2, 4, 1, 8, 4, 3, 2, 4, 7, 3, 4, 5, 6, 3, 2, 4, 6, 7, 8, 3, 8, 9, 4, 6, 4, 9, 2, 3, 4, 5, 6, 3, 4.

### Вариант 3

1. Среднее значение признака составило 14, какими будут доверительные границы математического ожидания:

а) (14; 14,5), б) (13,5; 13,9); с) (13,5; 14,5); d) (13,5; 14). Поясните ответ.

2. Исследовали содержание жира в коровьем молоке (%) в партиях от двух поставщиков:

1 поставщик 3,7 4,1 3,9 3,7 4,3 3,6 3,5 3,6 3,7 3,6 3,5 3,9 3,8 3,6 3,7 3,4 3,5

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

Установите, имеются ли достоверные различия по содержанию жира в

коровьем молоке между исследованными партиями. Сделайте вывод.

3. Вычислите среднюю арифметическую следующего ряда чисел:

5, 4, 7, 3, 5, 8, 3, 5, 8, 1, 5, 4, 5, 6, 10, 14, 10, 9, 13, 12, 18, 4, 2, 7, 3, 9, 16, 12, 18, 14, 15, 16.

#### Вариант 4

1. Если основная гипотеза имеет вид  $H_0 = 25$ , то альтернативной гипотезой может быть:

а)  $H_1 \neq 25$ ; б)  $H_1 \leq 25$ ; в)  $H_1 \geq 25$ ; д)  $H_1 \leq 35$ . Поясните ответ.

2. Определите  $\bar{X} \pm S_{\bar{x}}$ ,  $\sigma^2$ ,  $\sigma$ , а также доверительные границы для генеральной совокупности по данным, полученным при определении кислотного числа жира исследуемых образцов продукта:

0,45; 0,35; 0, 37; 0,36; 0,43; 0,65; 0, 38; 0,52; 0,53; 0,49; 0,39; 0,37; 0, 58; 0, 55; 0,38; 0,35; 0, 42; 0,46; 0,46; 0, 53; 0,49. Сделайте вывод.

3. На мясоконсервном заводе №1 из 652 проанализированных образцов мясных продуктов 25 штук имели содержание кадмия, превышающее ПДК. Определите процент бракованных образцов.

#### Вариант 5

1. Среднее значение признака составило 22, какими будут доверительные границы математического ожидания:

а) (22; 22,5), б) (21,5; 21,9); с) (21,5; 22,5); д) (21,5; 22). Поясните ответ.

2. Исследовали содержание усвояемых белков (% на воздушносухое вещество) в разных партиях гороха:

1-я партия 19,8 19 19,2 19,4 19,3 19,7 19,4 18,1 19,7 19,8 18,2  
21 20,4 20,7 21,5

2-я партия 20,9 19,3 20,6 19,4 20,6 19,2 21,6 19,6 18,9 21,1 20,2  
22,4 21,8 21,5 22,7

Оцените достоверность разности между партиями гороха по среднему содержанию усвояемых белков. Сделайте вывод.

3. Установлен коэффициент корреляции ( $r = 0,28$ ) между содержанием триптофана и оксипролина в образцах говядины ( $n=45$ ). Определите достоверность коэффициента и напишите вывод.

### Вариант 6

1. Если основная гипотеза имеет вид  $H_0 = 12$ , то альтернативной гипотезой может быть:

а)  $H_1 \neq 12$ ; б)  $H_1 \leq 12$ ; в)  $H_1 \geq 12$ ; д)  $H_1 \leq 13$ . Поясните ответ.

2. По данным титруемой кислотности йогурта ( $^{\circ}\text{T}$ ), произведенного на предприятии №3, составьте гистограмму :

77,3 76,4 79,2 78,5 79,4 78,5 78,4 79,6 77,3 76,4 79,2 78,5 79,4  
78,5 78,4 79,6 75,6 80,1 77,9 75,3 78,2 76,9 77,4 79,1  
75,5 77,3 78,0 77,3 78,2 77,3 77,2 78,4 76,6 77,3 79,1 78,3 78,7  
77,8 77,7 78,9 76,6 75,8 78,5 77,8 78,7 77,8 77,7 78,3 79,0 76,8  
74,2 77,1 75,8 76,3 79,4 74,4 76,2 76,9 76,2 77,1 76,2 76,1  
80,1 76,3.

Сделайте вывод.

3. Вычислите среднюю арифметическую следующего ряда чисел:

15, 24, 38, 23, 45, 18, 14, 35, 14, 25, 36, 18, 24, 19, 23, 32, 18, 13, 29, 16, 22, 18, 21.

### Вариант 7

1. Среднее значение признака составило 23, какими будут доверительные границы математического ожидания: а) (23; 23,5), б) (22,5; 22,9); в) (22,5; 23,5); д) (23,5; 24). Поясните ответ.

2. На основании данных о содержании белка в образцах свинины (%) составьте гистограмму с использованием Excel и оцените нормальность распределения.:

15,2; 13,6; 14,5; 13,6; 13,5; 14,2; 13,8; 15,0; 13,8; 15,2; 13,2; 13,7; 14,3; 13,8;  
14,5; 15,1; 13,5; 14,3; 13,2; 14,5; 15,1; 13,7; 14,3; 14,7; 14,2; 15,1; 13,5; 14,4;  
13,5; 13,4; 13,5; 14,7; 14,8; 13,9; 14,6; 13,3; 13,7; 14,8; 14,3; 15,1; 13,1; 13,3;  
13,9; 13,5; 13,6; 14,7; 14,8; 14,0; 13,7; 13,5.

Сделайте вывод.

3. На перерабатывающем заводе №12 из 1276 проанализированных образцов молочных продуктов 13 штук имели нехарактерный вкус и запах. Определите процент бракованных образцов.

### Вариант 8

1. Среднее значение признака составило 25, какими будут доверительные границы математического ожидания:

а) (25; 25,5), б) (24,5; 24,9); в) (24,5; 25,5); д) (24,5; 25). Поясните ответ.

2. По данным о содержании поваренной соли в образцах вареных колбасных изделий (%) составьте гистограмму с использованием Excel и оцените нормальность распределения.:

2,35; 2,65; 3,75; 2,76; 2,40; 2,64; 2,98; 2,87; 3,10; 3,24; 2,89; 3,17; 3,36; 2,95; 3,18; 2,76; 2,46; 2,97; 2,39; 2,41; 2,38; 2,48; 2,44; 2,48; 2,39; 2,35; 2,35; 2,56; 3,54; 3,18; 3,12; 3,17; 3,27; 3,19; 3,53; 3,33; 3,14; 3,28; 2,91; 2,74; 2,85; 2,81; 2,35; 2,95; 3,67; 3,15; 3,28. Сделайте вывод.

3. Определите медиану следующего ряда цифр:

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

### Вариант 9

1. Если основная гипотеза имеет вид  $H_0 = 120$ , то альтернативной гипотезой может быть:

а)  $H_1 \neq 120$ ; б)  $H_1 \leq 120$ ; в)  $H_1 \geq 120$ ; г)  $H_1 \leq 130$ . Поясните ответ.

2. По данным о содержании сухих веществ в образцах сладкой консервной продукции составьте гистограмму с использованием Excel и оцените нормальность распределения.

55,25; 57,55; 64,35; 55,35; 55,25; 56,25; 57,35; 56,35; 57,30; 57,75; 58,25; 59,35; 59,25; 60,55; 60,15; 60,35; 61,25; 60,65; 61,85; 62,45; 60,15; 63,65; 64,05; 60,85; 56,65; 56,35; 59,25; 58,35; 61,45; 59,75; 57,95; 60,15; 62,25; 60,65; 59,95; 58,55; 56,75; 58,35; 59,15; 58,35; 60,35; 58,15; 55,55; 61,75; 56,45; 61,55; 58,75; 64,05; 63,35; 61,75; 63,25; 63,75; 63,85; 63,65.

Сделайте вывод.

3. При определении pH образцов свинины были получены следующие значения: 6,65; 6,60; 6,25; 5,90; 6,85; 5,95; 6,85; 6,35; 6,90. Рассчитать среднее значение pH и его доверительные границы при вероятности 99,7.

### Вариант 10

1. Среднее значение признака составило 17, какими будут доверительные границы математического ожидания:

а) (17; 17,5), б) (17,5; 17,9); в) (16,5; 17,5); г) (17,5; 18). Поясните ответ.

2. Исследовали содержание аминокислотного азота в образцах свежего мяса в разных партиях продукции:

1-я партия

77,2; 76,3; 78,0; 77,8; 79,0; 78,1; 79,4; 78,6; 79,3; 78,1; 79,7; 78,2; 79,1; 79,5; 79,3; 77,5; 79,2; 78,6; 76,9.

2-я партия

79,2; 78,1; 76,3; 79,2; 78,7; 77,9; 79,1; 79,2; 78,9; 79,5; 79,6; 78,9; 79,6; 78,6; 79,5; 79,1; 79,3; 79,2; 79,9.

Оцените достоверность разности между партиями по среднему содержанию аминокислотного азота в образцах свежего мяса. Сделайте вывод.

3. Установлен коэффициент корреляции ( $r = -0,75$ ) между массовой долей сахарозы и сухих веществ в образцах молока ( $n=89$ ). Определите достоверность коэффициента и напишите вывод.

#### Вариант 11

1. Если основная гипотеза имеет вид  $H_0=18$ , то альтернативной гипотезой может быть: а)  $H_1 \neq 18$ ; б)  $H_1 \leq 18$ ; в)  $H_1 \geq 18$ ; д)  $H_1 \leq 19$ . Поясните ответ.

2. Исследовали pH образцов свинины в 2-х партиях, проанализированных в разные смены:

1-я смена

6,55; 6,70; 6,10; 5,95; 6,80; 6,25; 6,55; 6,80; 6,00; 5,95; 5,90; 6,10; 6,35; 6,25; 6,75; 6,30; 6,45; 6,30; 6,15; 6,20.

2-я смена

6,60; 6,70; 6,15; 6,15; 6,85; 6,15; 6,25; 6,20; 6,25; 6,15; 6,20; 6,30; 6,15; 6,10; 6,20; 6,25; 6,30; 6,35; 6,45; 6,15.

Оцените достоверность разности между партиями, исследованными в разные смены, по данным pH образцов свинины. Сделайте вывод.

3. Каково будет максимальное содержание углеводов в шоколаде, если среднее значение углеводов составило 36 г, а стандартное отклонение 0,9 г.

#### Вариант 12

1. Среднее значение признака составило 114, какими будут доверительные границы математического ожидания:

а) (114; 114,5), б) (113,5; 113,9); в) (113,5; 114,5); д) (113,5; 114). Поясните ответ.

2. Исследовали содержание жира в коровьем молоке (%) двух поставщиков:

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

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

Установите, имеются ли достоверные различия по содержанию жира в молоке у коров из 2-х ферм.

3. Определите моду следующего ряда чисел:  
15, 21, 14, 15, 18, 14, 31, 22, 14, 17, 31, 24, 15, 26, 13, 23, 14, 26, 17, 28, 13, 18, 17.



## 9. Словарь терминов

Асимметрия (коэффициент асимметрии) – мера несимметричности распределения.

Бимодальное распределение – распределение объектов совокупности, имеющее две моды (т.е. два «пика»).

Биномиальное распределение – распределение объектов совокупности по признакам с качественной или дискретной количественной изменчивостью.

Варианта ( $X_i$ ) – отдельно взятое измерение изучаемого признака.

Вариационный ряд - это ряд чисел, состоящий из классов и чисел.

Верхняя квартиль (квантиль) (75-я процентиль переменной) - это такое значение, ниже которого попадают 75% значений переменной.

Внутриклассовый коэффициент корреляции – коэффициент, отражающий долю или силу влияния фактора на изменчивость признака.

Выборка – набор чисел, случайно отобранных из независимых и одинаково распределенных случайных величин генеральной совокупности.

Генеральная совокупность – бесконечное множество объектов, объединенных по определенному признаку.

Дисперсионный анализ (ANOVA) – метод проверки значимости различий между средними арифметическими в разных группах с помощью сравнения дисперсий этих групп. Деление общей дисперсии на несколько источников, позволяет сравнить дисперсию, вызванную различием между группами, с дисперсией, вызванной внутригрупповой изменчивостью.

Доверительные границы – пределы изменчивости признака, в которых находится истинное значение средней арифметической.

Доверительные интервалы для среднего задают область вокруг среднего, в которой с заданным уровнем доверия содержится «истинное» среднее популяции.

Квартильный размах переменных - разность значений 75-й процентиля и 25-й процентиля, т.е. это тот диапазон вокруг медианы, который содержит 50% наблюдений исследуемой выборки.

Корреляция - связь между двумя переменными (признаками).

Коэффициент корреляции Пирсона – мера оценки степени линейных связей между переменными (признаками).

Коэффициент детерминации – это квадрат коэффициента корреляции Пирсона между двумя переменными.

Медиана (median) – варианта, занимающая в вариационном ряду срединное положение и разбивает выборку на две равные части.

Метод наименьших квадратов. Общий смысл оценивания по методу наименьших квадратов заключается в минимизации суммы квадратов отклонений наблюдаемых значений зависимой переменной от значений, предсказанных моделью.

Множественная регрессия - анализ связи между несколькими независимыми переменными (называемыми также регрессорами или предикторами) и зависимой переменной (откликом).

Мода (mode)- варианта, наиболее часто встречающаяся в вариационном ряду.

Нижняя квартиль (квантиль) (25-я процентиль переменной) - это такое значение, ниже которого попадают 25% значений переменной.

Нормальное распределение (Гауссовское) – непрерывное распределение вероятностей с пиком в центре и симметричными боковыми сторонами, которое в одномерном случае задаётся функцией плотности вероятности, совпадающей с функцией Гаусса.

Объем генеральной совокупности  $N$  и объем выборки  $n$  – число объектов в рассматриваемой совокупности.

Полимодальное (мультимодальное) распределение - распределение объектов совокупности, имеющее несколько мод («пиков»).

Признак – это свойство, которое характеризует определенные черты и особенности, присущие единицам изучаемой совокупности.

Распределение Пуассона - тип распределения редко происходящих событий.

Свободные эффекты в дисперсионном анализе — факторы с неизвестными заранее уровнями.

Совокупность генеральная – бесконечное множество объектов, объединенных по какому-либо признаку.

Совокупность выборочная – часть генеральной совокупности, члены которой были отобраны случайным образом.

Средняя арифметическая – усредненное значение, характеризующее какую-либо группу наблюдений.

Среднее квадратичное отклонение или стандартное отклонение – показатель указывает, насколько в среднем каждая варианта отклоняется от средней арифметической.

Стандартная ошибка – статистический параметр, отражающий репрезентативность выборок по отношению к генеральным совокупностям.

Статистический уровень значимости (р-уровень) - статистическая значимость результата представляет собой меру уверенности в его «истинности» (в смысле «репрезентативности выборки»).

Унимодальное распределение - это распределение, имеющее только одну моду.

Фиксированные эффекты в дисперсионном анализе — факторы, уровни которых заранее определяются исследователем, а не случайно выбираются в ходе эксперимента.

Частота (р) – статистический параметр, указывающий сколько раз встречается данная варианта в вариационном ряду,

Эксцесс (коэффициент эксцесса) – величина, измеряющая «пикообразность» распределения объектов в совокупности.

## **Приложения**

Стандартные значения критерия  $\chi^2$ 

Число степеней свободы (df)	Вероятность			Число степеней свободы (df)	Вероятность		
	0,95	0,99	0,999		0,95	0,99	0,999
1	3,8	6,6	10,8	11	19,7	24,7	31,3
2	6,0	9,2	13,8	12	21,0	26,2	32,9
3	7,8	11,3	16,3	13	22,4	27,7	34,5
4	9,5	13,8	18,5	14	23,7	29,1	36,1
5	11,1	15,1	20,5	15	25,0	30,6	37,7
6	12,6	16,8	22,5	16	26,3	32,0	39,3
7	14,1	18,5	24,3	17	27,6	33,4	40,8
8	15,5	20,1	26,1	18	28,9	34,8	42,3
9	16,9	21,7	27,9	19	30,1	36,2	43,8
10	18,3	23,2	29,6	20	31,4	37,6	45,3
11	19,7	24,7	31,3	21	32,7	39,0	46,9
12	21,0	26,2	32,9	22	34,0	40,3	48,4
13	22,4	27,7	34,5	23	35,3	41,7	50,0
14	23,7	29,1	36,1	24	36,5	43,0	51,6
15	25,0	30,6	37,7	25	37,7	44,3	53,2
16	26,3	32,0	39,3	26	39,0	45,7	54,9
17	27,6	33,4	40,8	27	40,2	47,0	56,6
18	28,9	34,8	42,3	28	41,4	48,3	58,3
19	30,1	36,2	43,8	29	42,6	49,6	60,1
20	31,4	37,6	45,3	30	43,8	50,9	61,9
21	32,7	39,0	46,9	31	45,0	52,2	63,7
22	34,0	40,3	48,4	32	46,2	53,5	65,6
23	35,3	41,7	50,0	33	47,4	54,8	67,5
24	36,5	43,0	51,6	34	48,6	56,1	69,4
25	37,7	44,3	53,2	35	49,8	57,4	71,4
26	39,0	45,7	54,9	36	51,0	58,7	73,4
27	40,2	47,0	56,6	37	52,2	60,0	75,5
28	41,4	48,3	58,3	38	53,4	61,3	77,6
29	42,6	49,6	60,1	39	54,6	62,6	79,7
30	43,8	50,9	61,9	40	55,8	63,9	81,9
31	45,0	52,2	63,7	41	57,0	65,2	84,1
32	46,2	53,5	65,6	42	58,2	66,5	86,3
33	47,4	54,8	67,5	43	59,4	67,8	88,6
34	48,6	56,1	69,4	44	60,6	69,1	90,9
35	49,8	57,4	71,4	45	61,8	70,4	93,2
36	51,0	58,7	73,4	46	63,0	71,7	95,6
37	52,2	60,0	75,5	47	64,2	73,0	98,0
38	53,4	61,3	77,6	48	65,4	74,3	100,5
39	54,6	62,6	79,7	49	66,6	75,6	103,0
40	55,8	63,9	81,9	50	67,8	76,9	105,6
41	57,0	65,2	84,1	51	69,0	78,2	108,2
42	58,2	66,5	86,3	52	70,2	79,5	110,9
43	59,4	67,8	88,6	53	71,4	80,8	113,7
44	60,6	69,1	90,9	54	72,6	82,1	116,5
45	61,8	70,4	93,2	55	73,8	83,4	119,4
46	63,0	71,7	95,6	56	75,0	84,7	122,3
47	64,2	73,0	98,0	57	76,2	86,0	125,3
48	65,4	74,3	100,5	58	77,4	87,3	128,3
49	66,6	75,6	103,0	59	78,6	88,6	131,4
50	67,8	76,9	105,6	60	79,8	89,9	134,6
51	69,0	78,2	108,2	61	81,0	91,2	137,8
52	70,2	79,5	110,9	62	82,2	92,5	141,1
53	71,4	80,8	113,7	63	83,4	93,8	144,5
54	72,6	82,1	116,5	64	84,6	95,1	148,0
55	73,8	83,4	119,4	65	85,8	96,4	151,6
56	75,0	84,7	122,3	66	87,0	97,7	155,3
57	76,2	86,0	125,3	67	88,2	99,0	159,1
58	77,4	87,3	128,3	68	89,4	100,3	163,0
59	78,6	88,6	131,4	69	90,6	101,6	167,0
60	79,8	89,9	134,6	70	91,8	102,9	171,1
61	81,0	91,2	137,8	71	93,0	104,2	175,3
62	82,2	92,5	141,1	72	94,2	105,5	179,6
63	83,4	93,8	144,5	73	95,4	106,8	184,0
64	84,6	95,1	148,0	74	96,6	108,1	188,5
65	85,8	96,4	151,6	75	97,8	109,4	193,1
66	87,0	97,7	155,3	76	99,0	110,7	197,8
67	88,2	99,0	159,1	77	100,2	112,0	202,6
68	89,4	100,3	163,0	78	101,4	113,3	207,5
69	90,6	101,6	167,0	79	102,6	114,6	212,5
70	91,8	102,9	171,1	80	103,8	115,9	217,6
71	93,0	104,2	175,3	81	105,0	117,2	222,8
72	94,2	105,5	179,6	82	106,2	118,5	228,1
73	95,4	106,8	184,0	83	107,4	119,8	233,5
74	96,6	108,1	188,5	84	108,6	121,1	239,0
75	97,8	109,4	193,1	85	109,8	122,4	244,6
76	99,0	110,7	197,8	86	111,0	123,7	250,3
77	100,2	112,0	202,6	87	112,2	125,0	256,1
78	101,4	113,3	207,5	88	113,4	126,3	262,0
79	102,6	114,6	212,5	89	114,6	127,6	268,0
80	103,8	115,9	217,6	90	115,8	128,9	274,2
81	105,0	117,2	222,8	91	117,0	130,2	280,5
82	106,2	118,5	228,1	92	118,2	131,5	287,0
83	107,4	119,8	233,5	93	119,4	132,8	293,6
84	108,6	121,1	239,0	94	120,6	134,1	300,4
85	109,8	122,4	244,6	95	121,8	135,4	307,3
86	111,0	123,7	250,3	96	123,0	136,7	314,4
87	112,2	125,0	256,1	97	124,2	138,0	321,6
88	113,4	126,3	262,0	98	125,4	139,3	329,0
89	114,6	127,6	268,0	99	126,6	140,6	336,5
90	115,8	128,9	274,2	100	127,8	141,9	344,2

## Приложение 3

Таблица. Критические значения U-критерия Манна-Уитни

$n_1 \backslash n_2$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1																				
2								0	0	0	0	1	1	1	1	1	2	2	2	2
3					0	1	1	2	2	3	3	4	4	5	5	6	6	7	7	8
4				0	1	2	3	4	4	5	6	7	8	9	10	11	11	12	13	13
5			0	1	2	3	5	6	7	8	9	11	12	13	14	15	17	18	19	20
6			1	2	3	5	6	8	10	11	13	14	16	17	19	21	22	24	25	27
7			1	3	5	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34
8		0	2	4	6	8	10	13	15	17	19	22	24	26	29	31	34	36	38	41
9		0	2	4	7	10	12	15	17	20	23	26	28	31	34	37	39	42	45	48
10		0	3	5	8	11	14	17	20	23	26	29	33	36	39	42	45	48	52	55
11		0	3	6	9	13	16	19	23	26	30	33	37	40	44	47	51	55	58	62
12		1	4	7	11	14	18	22	26	29	33	37	41	45	49	53	57	61	65	69
13		1	4	8	12	16	20	24	28	33	37	41	45	50	54	59	63	67	72	76
14		1	5	9	13	17	22	26	31	36	40	45	50	55	59	64	67	74	78	83
15		1	5	10	14	19	24	29	34	39	44	49	54	59	64	70	75	80	85	90
16		1	6	11	15	21	26	31	37	42	47	53	59	64	70	75	81	86	92	98
17		2	6	11	17	22	28	34	39	45	51	57	63	67	75	81	87	93	99	105
18		2	7	12	18	24	30	36	42	48	55	61	67	74	80	86	93	99	106	112
19		2	7	13	19	25	32	38	45	52	58	65	72	78	85	92	99	106	113	119
20		2	8	13	20	27	34	41	48	55	62	69	76	83	90	98	105	112	119	127

**ФГБОУ ВО НОВОСИБИРСКИЙ ГАУ**  
**КАФЕДРА ВЕТЕРИНАРНОЙ ГЕНЕТИКИ И БИОТЕХНОЛОГИИ**

Контрольная работа

по дисциплине «Прикладная статистика»

Выполнил: студент 2411 группы...

ФИО.. ..

Проверил: .....

док. биол. наук, проф. Кочнева М.Л.

Новосибирск 20\_\_

Составитель:  
Кочнева Марина Львовна

Прикладная статистика

Учебно-методическое пособие

Компьютерная верстка: М.Л. Кочнева

Объем 3,0 п. л.

ФГБОУ ВО Новосибирский ГАУ  
630039, Новосибирск, ул. Добролюбова, 160.