ПРИДНЕСТРОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМ. Т.Г. ШЕВЧЕНКО

Аграрно-технологический факультет

Кафедра садоводства, защиты растений и экологии

ОСНОВЫ БИОЛОГИЧЕСКОЙ СТАТИСТИКИ

Методические указания к практическим занятиям

Тирасполь, 2019

УДК [57:37](072.8)
ББК Е.к95р30
О75

Составитель:
Н.Н. Трескина, к.с.н., доцент

Рецензенты:
Т.Н. Звездина, к.с.н., доцент
Т.В. Пазяева, к.с.н., доцент

Основы биологической статистики: Методические указания к практическим занятиям/ Сост. Н.Н. Трескина. – Тирасполь, 2019. – 72 с.

Методические указания к практическим занятиям по дисциплине «Основы биологической статистики» для обучающихся по направлениям подготовки 4.35.03.04 АГРОНОМИЯ профили «Агробизнес», «Защита растений» и 4.35.03.05 САДОВОДСТВО профили «Плодоовощеводство и виноградарство», «Декоративное садоводство и ландшафтный дизайн» составлены в соответствии с государственными стандартами высшего образования Российской Федерации.

Методические указания включают 13 практических занятий, состоящих из теоретической и расчетной частей, индивидуальных заданий и контрольных вопросов.

УДК[57:37](072.8)
ББК Е.к95р30

Рекомендовано НМС ПГУ им. Т.Г. Шевченко

© ПГУ им. Т.Г. Шевченко
© Трескина Н.Н.,
составление, 2019

ОГЛАВЛЕНИЕ

Практическое занятие № 1. Оценка «сомнительной» варианты 5
Практическое занятие № 2. Способы вычисления средних величин 7
Практическое занятие № 3. Точечная и интервальная оценки распределения 13
Практическое занятие № 4. Построение вариационного ряда 18
Практическое занятие № 5. Статистическая обработка результатов наблюдений при количественной изменчивости 24
Практическое занятие № 6. Статистическая обработка результатов наблюдений при качественной изменчивости 32
Практическое занятие № 7. Оценка существенности разности средних двух выборок при количественной изменчивости 35
Практическое занятие № 8. Оценка существенности разности между выборочными долями при качественной изменчивости 42
Практическое занятие № 9. Оценка соответствия между фактическими и теоретическими частотами по критерию Пирсона 46
Практическое занятие № 10. Оценка существенности различий между двумя независимыми выборками по критерию ван-дер-Вардена 52
Практическое занятие № 11. Коэффициент ранговой корреляции Спирмена 56
Практическое занятие № 12. Критерий знаков 60
Практическое занятие № 13. Пробит-анализ 62
*Приложение 1.*Значения критерия τ на 1 и 5% уровнях значимости 65
Приложение 2. Значение критерия Стьюдента на 1 и 5% уровнях значимости 66

Приложение 3. Значение функции $\varphi\left(\frac{R}{n+1}\right)$ для расчета критерия ван-дер-Вардена 67
Приложение 4. Теоретические значения Х-критерия ван-дер-Вардена на 1 и 5% уровнях значимости 69
Приложение 5. Теоретические значения критерия знаков z на 1 и 5% уровнях значимости 70
Приложение 6. Значения пробитов для экспериментально устанавливаемой гибели вредных организмов от 0 до 99 % 71
Список использованной литературы 72

Практическое занятие №1
ОЦЕНКА «СОМНИТЕЛЬНОЙ» ВАРИАНТЫ

Часто встречаются случаи, когда выборочная совокупность содержит наблюдения (варианты), значения которых сильно отличаются от основной массы наблюдений. Браковка сомнительных вариант на глаз субъективна и совершенно недопустима. Варианта, вызывающая сомнение, может быть забракована только путем статистической проверки.

Проверка гипотезы о принадлежности «сомнительных» вариант, обычно крайних значений x_1 и x_n, к данной совокупности в малых выборках осуществляется по критерию τ (тау).

Чтобы рассчитать значение критерия варианты располагают в порядке возрастания $x_1, x_2, \ldots, x_n$. Сомнительными обычно бывают один или оба крайних члена ряда, т.е. x_1 и x_n.

Фактическое значение критерия τ вычисляют по формулам:

для x_1 $\tau = \dfrac{x_2 - x_1}{x_{n-1} - x_1}$ (1); для x_n $\tau = \dfrac{x_n - x_{n-1}}{x_n - x_2}$ (2)

Разности x_{n-1}-x_1 и x_n-x_2 характеризуют размах варьирования ряда данных без крайних значений, которые сомнительны и, следовательно, нецелесообразно связывать с ними оценку значимости отклонения x_1 с сомнительной x_n, а x_n – с сомнительной вариантой x_1.

Для определения принадлежности к данной выборке варианты, вызывающей сомнение, фактическое значение критерия сравнивают с теоретическим на 1 и 5% уровне значимости. (Уровень значимости – величина ошибки выводов, сделанных на основании расчетов, т.е. при 1% уровне значимости наши выводы ошибочны на 1%, а при 5% соответственно на 5%). Теоретическое значение критерия берут из статистической таблицы в зависимости от количества значений признака в выборке – n (приложение 1).

Если $\tau_{факт} \geq \tau_{теор}$, то варианта отбрасывается, если $\tau_{факт} < \tau_{теор}$, то варианта оставляется.

Пример. На параллельных делянках получили следующие значения урожая, кг*:*

5,2 17,4 18,1 19,8 22,9 26,4

Сомнения вызывают x_1=5,2 и x_6=26,4

Рассчитываем фактическое значение критерия τ:

для x_1 по формуле (1) $\tau = \frac{17,4-5,2}{22,9-5,2} = \frac{12,2}{17,7} = 0,689$

для x_6 по формуле (2) $\tau = \frac{26,4-22,9}{26,4-17,4} = \frac{3,5}{9,0} = 0,389$

Из статистической таблицы берем теоретическое значение критерия τ при количестве значений признака n=6 τ_{01}= 0,805 и τ_{05}=0,689.

Для $\mathbf{x_1}$:

при 1% уровне значимости $\tau_{факт}$ = 0,689 τ_{01} = 0,805, $\tau_{факт} < \tau_{01}$, следовательно варианта x_1 остается в выборке;

при 5% уровне значимости $\tau_{факт}$ = 0,689 τ_{05}=0,689, $\tau_{факт} = \tau_{05}$, следовательно варианта x_1 не принадлежит выборке и отбрасывается.

Таким образом, несмотря на то, что при 5% уровне значимости варианта x_1 отбрасывается, при более строгом подходе – 1% уровень значимости, оснований для отбраковки варианты нет.

Для $\mathbf{x_6}$:

при 1% уровне значимости $\tau_{факт}$ = 0,389 τ_{01} = 0,805, $\tau_{факт} < \tau_{01}$, следовательно варианта x_6 остается в выборке;

при 5% уровне значимости $\tau_{факт}$ = 0,389 τ_{05}=0,689, $\tau_{факт} < \tau_{05}$, следовательно варианта x_6 остается в выборке.

Таким образом, для варианты x_6 оснований для браковки нет ни на 5, ни на 1% уровне значимости.

Задание: Проверить принадлежности крайних вариант к совокупности.

Пример 1. *В образцах почвы было обнаружено следующее содержание гумуса, %:*

1,43; 2,39; 2,51; 2,69; 2,77; 2,82; 3,01; 3,44

Пример 2. *В образцах зерна пшеницы, взятых с параллельных делянок, было обнаружено следующее содержание клейковины, %:*

14,2; 21,4; 22,3; 24,5; 29,6; 32,8; 34,1; 35,7; 36,4; 41,3

Пример 3. *На параллельных делянках получили следующие значения урожая, кг:*

11,6; 22,1; 24,4; 28,2; 31,4; 2,9; 39,2

Пример 4. *В плодах было обнаружено следующее содержание аскорбиновой кислоты, мг/100 г:*

2,17; 3,01; 3,17; 3,22; 3,26; 3,37; 3,48; 3,73

Контрольные вопросы:

1. В каких случаях используется критерий τ?
2. Оценка принадлежности сомнительной варианты к данной совокупности по критерию τ

Практическое занятие № 2
СПОСОБЫ ВЫЧИСЛЕНИЯ СРЕДНИХ ВЕЛИЧИН

Средние величины являются обобщенной, абстрактной характеристикой всей совокупности в целом. Значение средних заключается в их свойстве уравновешивать все индивидуальные отклонения, в результате чего получаются значения, характеризующие своеобразие группового объекта и позволяющие отличать варьирующие объекты друг от друга.

Средние величины могут характеризовать только однородную массу вариант. При наличии разнородных по составу данных их необходимо группировать в отдельные качественно однородные группы и вычислять средние групповые или частные средние.

Существует 2 класса средних величин: ***степенные*** и ***структурные***.

К **структурным** средним относятся *мода* и *медиана.*

Мода (Мо) – значение, которое встречается в данной совокупности с наибольшей частотой.

Медиана (Ме) – средняя, относительно которой ряд значений признака делится на 2 половины: в обе стороны от медианы располагается одинаковое число вариант. В выборках малого размера определить медиану достаточно легко: для этого значения признака ранжируют по возрастанию, и если число членов ряда нечетное, то центральная варианта и будет медианой. При четном числе значений признака медиана определяется по полусумме соседних вариант, расположенных в центре ряда.

Структурные средние применяются достаточно редко.

Степенные средние величины могут быть *простыми* и *взвешенными*.

Простая средняя величина рассчитывается для *не сгруппированных* данных, расположенных в произвольном порядке, по следующей общей формуле:

$$\bar{x} = \sqrt[m]{\frac{\sum X^m}{n}} \qquad (3)$$

Взвешенная средняя величина рассчитывается для *сгруппированных* данных по следующей общей формуле:

$$\bar{x} = \sqrt[m]{\frac{\sum X^m \cdot f}{\sum f}} \qquad (4)$$

где X – значения варьирующего признака или середины группировочных интервалов;

m - показатель степени, от значения которого зависит вид степенных средних величин:

при m = -1 средняя гармоническая;

при m = 0 средняя геометрическая;

при m = 1 средняя арифметическая;

при m = 2 средняя квадратическая;

при m = 3 средняя кубическая.

Из степенных средних наиболее часто используется ***средняя арифметическая***. Для малой выборки и несгруппированных данных определяется ***простая средняя арифметическая*** по формуле:

$$\bar{x} = \frac{\sum X}{n} \qquad (5)$$

где X – значение варьирующего признака,

n – общее число вариант, составляющих выборку.

Для сгруппированных данных, т.е. для вариационного ряда рассчитывается ***взвешенная средняя арифметическая:***

$$\bar{x} = \frac{f_1 \cdot X_1 + f_2 \cdot X_2 + ... + f_n \cdot X_n}{f_1 + f_2 + ... + f_n} = \frac{\sum f \cdot X}{n} \qquad (6)$$

Средняя гармоническая обозначается символом $\mathbf{x_h}$**.** Для расчета средней гармонической используются не абсолютные значения вариант, а их обратные числа. Средняя гармоническая может быть простой и взвешенной.

Простая рассчитывается по формуле: $\bar{x}_h = \dfrac{n}{\sum \dfrac{1}{X}}$ (7)

взвешенная: $\bar{x}_h = \dfrac{n}{\sum f \cdot \dfrac{1}{X}}$ (8)

Пример. Для определения плотности колосьев ржи было отобрано 20 растений, у которых подсчитывалось число зерен в колосе и измерялась длина колоса в см. Плотность определялась как отношение числа зерен к длине колоса. Результаты распределились следующим образом:

Плотность колосьев	4,5	4,2	4,0	3,7	3,5
Частота, f	2	5	10	2	1

Рассчитываем взвешенную среднюю гармоническую по формуле (8):

$$\bar{x}_h = \frac{20}{2 \cdot \frac{1}{4,5} + 5 \cdot \frac{1}{4,2} + 10 \cdot \frac{1}{4,0} + 2 \cdot \frac{1}{3,7} + 1 \cdot \frac{1}{3,5}} =$$

$$= \frac{20}{2 \cdot 0,222 + 5 \cdot 0,238 + 10 \cdot 0,25 + 2 \cdot 0,270 + 1 \cdot 0,286} =$$

$$= \frac{20}{0,444 + 1,19 + 2,5 + 0,54 + 0,286} = \frac{20}{4,96} = 4,03$$

в то время как взвешенная средняя арифметическая рассчитывается по формуле (6):

$$\bar{x} = \frac{2 \cdot 4{,}5 + 5 \cdot 4{,}2 + 10 \cdot 4{,}0 + 2 \cdot 3{,}7 + 1 \cdot 3{,}5}{20} =$$

$$= \frac{9{,}0 + 21{,}0 + 40{,}0 + 7{,}4 + 3{,}5}{20} = \frac{80{,}9}{20} = 4{,}04$$

Задание 1: Рассчитать среднюю арифметическую и среднюю гармоническую.

Пример 1. *У 50 растений льна-долгунца определено процентное соотношение технической длины стебля к общей. Получены следующие результаты:*

Соотношение, %	80	83	85	88	90	92
Частота, f	2	8	13	15	9	3

Пример 2.*У 50 растений винограда определено процентное соотношение вызревшей части однолетнего прироста к общей. Получены следующие результаты:*

Соотношение, %	70	73	75	78	80	82
Частота, f	2	8	13	15	9	3

Если признак выражается мерами площади (например, диаметр корзинок подсолнечника, определяющий его урожайность; величина листовых пластинок, от которой зависит продуктивность фотосинтеза; размеры колоний микроорганизмов) более точной является ***средняя квадратическая***, которая обозначается символом $\boldsymbol{x_q}$**.** Она равна корню квадратному из отношения суммы квадратов вариант к общему числу наблюдений.

Простая средняя квадратическая рассчитывается по формуле: $\bar{x}_q = \sqrt{\frac{\sum x^2}{n}}$ (9); взвешенная - $\bar{x}_q = \sqrt{\frac{\sum fx^2}{n}}$ (10).

Пример. При измерении диаметра у 10 корзинок подсолнечника получены следующие результаты:

Диаметр, см	8	11	13	15	16	17
Частота, f	1	1	2	3	2	1

Рассчитываем взвешенную среднюю квадратическую по формуле (10):

$$\bar{x}_q = \sqrt{\frac{1\cdot 8^2 + 1\cdot 11^2 + 2\cdot 13^2 + 3\cdot 15^2 + 2\cdot 16^2 + 1\cdot 17^2}{10}} =$$

$$= \sqrt{\frac{1\cdot 64 + 1\cdot 121\cdot 2\cdot 169 + 3\cdot 225 + 2\cdot 256 + 1\cdot 289}{10}} =$$

$$= \sqrt{\frac{64 + 121 + 338 + 675 + 512 + 289}{10}} = \sqrt{\frac{1999}{10}} = \sqrt{199{,}9} = 14{,}1$$

в то время как взвешенная средняя арифметическая рассчитывается по формуле (6):

$$\bar{x} = \frac{1\cdot 8 + 1\cdot 11 + 2\cdot 13 + 3\cdot 15 + 2\cdot 16 + 1\cdot 17}{10} =$$

$$= \frac{8 + 11 + 26 + 45 + 32 + 17}{10} = \frac{139}{10} = 13{,}9$$

Задание 2: Рассчитать среднюю квадратическую и среднюю арифметическую.

Пример 3. *При измерении площади листьев яровой пшеницы получены следующие результаты:*

Площадь листа, см^2 14,2; 16,1; 13,3; 17,4; 13,8; 16,5; 15,2; 16,9; 14,7; 5,7.

Пример 4. *При измерении площади листьев яблони получены следующие результаты:*

Площадь листа, см^2 31 28 33 37 32 29 32 36 28 35

Пример 5. *При измерении диаметра луковиц лука севка были получены следующие данные:*

Диаметр луковицы, см	1,6	1,8	2,0	2,2	2,4	2,6	2,8
Частота, f	2	4	5	7	6	3	1

Пример 6. *При измерении диаметра 16 колоний микроорганизмов были получены следующие значения:*

Диаметр колоний, см	3	5	7	9	11
Число колоний	1	3	6	4	2

Для характеристики объемных признаков более точным показателем является ***средняя кубическая***, которая обозначается символом $\mathbf{x_k}$. Простая средняя кубическая рассчитывается по формуле:

$$\bar{x}_k = \sqrt[3]{\frac{\sum X^3}{n}} \quad (11); \quad \text{взвешенная -} \quad \bar{x}_k = \sqrt[3]{\frac{\sum f \cdot X^3}{n}} \quad (12).$$

Средняя геометрическая – более точная характеристика при определении средних прибавок или при увеличении средних размеров, прироста численности популяции за определенный промежуток времени. Обозначается символом $\boldsymbol{x_g}$ и равна корню n-ой степени из произведений членов ряда:

$$\bar{x}_g = \sqrt[n]{X_1 \cdot X_2 \cdot \ldots \cdot X_n} \qquad (13)$$

Обычно средняя геометрическая вычисляется с помощью десятичных логарифмов по формуле:

$$\lg \bar{x}_g = \frac{1}{n-1}\left(\lg X_1 + \lg X_2 + \ldots + \lg X_n\right) = \frac{1}{n-1} \cdot \sum \lg X \qquad (14)$$

Правило мажорантности средних величин

Если рассчитать все виды средних для одних и тех же исходных данных, то значения их окажутся неодинаковыми. Здесь действует правило мажорантности средних: *с увеличением показателя степени m увеличивается и соответствующая средняя величина, т.е.:*

$$\bar{x}_h < \bar{x}_g < \bar{x} < \bar{x}_q < \bar{x}_k$$

Контрольные вопросы:

1. Какие классы и виды средних используются?
2. В каких случаях рассчитывают среднюю квадратическую, среднюю кубическую и среднюю геометрическую?

Практическое занятие № 3
ТОЧЕЧНАЯ И ИНТЕРВАЛЬНАЯ ОЦЕНКИ РАСПРЕДЕЛЕНИЯ

Статистические оценки выборочной совокупности являются приближенными оценками генеральной совокупности. Оценка может быть точечной или интервальной.

Оценка параметра называется точечной, если она выражается одним числом. Точечные оценки являются случайными величинами, причем их значения зависят от объема экспериментальных данных. Точечные оценки могут быть ***состоятельными и несмещенными.***

Состоятельной называется оценка, которая при увеличении объема выборки стремится к истинному значению числовой характеристики.

Несмещенной называется оценка, которая не содержит систематической ошибки в сторону завышения или занижения оцениваемого параметра и математическое ожидание которой совпадает с параметром генеральной совокупности.

Как правило, выборочные характеристики не совпадают с соответствующими параметрами генеральной совокупности. Величина отклонения выборочного показателя от генерального называется *статистической ошибкой* этого показателя, или ошибкой репрезентативности. Статистические ошибки возникают не как результат ошибок при измерении, они возникают в процессе отбора вариант из совокупности. Поэтому точечная оценка генерального параметра записывается с указанием статистической ошибки.

Точечной несмещенной оценкой генеральной средней μ (мю) является средняя $\bar{x}$. Точечная оценка генеральной средней записывается в виде $\bar{x} \pm s_{\bar{x}}$, где $s_{\bar{x}}$ – ошибка выборочной средней. Это означает, что $\bar{x}$ является оценкой генеральной средней μ с ошибкой, равной $s_{\bar{x}}$.

Для практики важно получить не только точечную оценку, но и определить интервал, в пределах находится истинное значение оцениваемого параметра. Для этого используется интервальная оценка. *Интервальная оценка* характеризуется двумя числами – концами интервала, покрывающего оцениваемый параметр. Интервал, который с заданной вероятностью покрывает оцениваемый параметр, называется ***доверительным***. Центром интервала является выборочная оценка точки $\bar{x}$. Границы интервала определяются средней ошибкой оценки и уровнем вероятности. Интервальная оценка является развитием точечной оценки, которая неэффективна при малом объеме выборки.

Доверительный интервал для генеральной средней записывается следующим образом:

$$\bar{x} - t \cdot s_{\bar{x}} \leq \mu \leq \bar{x} + t \cdot s_{\bar{x}}$$

где t – критерий Стьюдента, который показывает отклонение выборочной средней $\bar{x}$ от генеральной средней μ в долях ошибки $s_{\bar{x}}$, принятой за единицу.Значение критерия Стьюдента берется из статистической таблицы (приложение 2) для соответствующего уровня значимости и числа степеней свободы v=n-1

$t \cdot s_{\bar{x}}$ – предельная ошибка выборочной средней при данном числе степеней свободы.

Пример. При определении содержания белка в пшенице были получены следующие значения: $\bar{x}$ = 14,80%; $s_{\bar{x}}$ = 0,20%; n = 4. Необходимо определить 95 и 99% доверительный интервал для генеральной средней μ. Значения t_{01} и t_{05} берем из статистической таблицы для v=4-1=3 степеней свободы (приложение). Для v=3 значения t_{01} =5,84 и t_{05} = 3,18.

Доверительный интервал для 95% уровня вероятности равен:

$\bar{x} \pm t \cdot s_{\bar{x}}$ 14,80±3,18×0,20 14,80±0,64 или **14,16÷15,44**

Доверительный интервал для 99% уровня вероятности равен:

$\bar{x} \pm t \cdot s_{\bar{x}}$ 14,80±5,84×0,20 14,80±1,17 или **13,63÷ 15,97**

Следовательно, генеральная средняя содержания белка в зерне пшеницы с вероятностью 95% находится в интервале от 14,16 до 15,44%; с вероятностью 99% – от 13,63 до 15,97%.

Вероятность того, что значение генеральной средней не попало в границы интервала составляет 5% в первом случае и 1% – во втором.

Интервальная оценка используется для статистической проверки гипотез при сравнении выборочных средних для предварительных данных. Если доверительные интервалы для генеральных перекрывают друг друга, то нулевая гипотеза принимается. Если доверительные интервалы не перекрываются, то нулевая гипотеза отвергается, следовательно, выборки принадлежат различным генеральным совокупностям, то есть между ними есть существенные различия.

Пример. При определении содержания сахара в корнеплодах гибридов сахарной свеклы Гримм и Джакета были получены следующие данные:

Гримм: $\bar{x}_1$= 19,20%; $s_{\bar{x}}$ = 0,21%; n = 12;

Джакета: $\bar{x}_2$= 18,70%; $s_{\bar{x}}$ = 0,19%; n = 14.

Определить доверительные интервалы при 95 и 99% уровне вероятности.

Для v_1 =12-1=11 значения t_{01} = 3,11 и t_{05} = 2,20;

для v_2 =14-1= 13 значения t_{01} = 3,01 и t_{05} = 2,16.

1. Рассчитываем доверительный интервал для 1 выборки на:

99% уровне вероятности:19,20±3,11×0,21 19,20±0,65 8,55÷19,85
95% уровне вероятности: 9,20±2,20×0,21 19,20±0,46 18,74÷19,66
2. Рассчитываем доверительный интервал для 2 выборки на:
99% уровне вероятности:

18,70±3,01×0,19 18,70±0,57 18,13÷19,27

95% уровне вероятности:

18,70±2,16×0,19 18,70±0,41 18,29÷19,11

3. Сравниваем доверительные интервалы на:

99% уровне вероятности: 18,55÷19,85 и 18,13÷19,27 интервалы перекрываются, следовательно между выборками нет существенного различия, т.е. по содержанию сахара в корнеплодах гибриды сахарной свеклы Гримм и Джакета сущуственно не различаются.

95% уровне вероятности: 18,74÷19,66 и 18,29÷19,11 интервалы перекрываются, следовательно между выборками нет существенного различия, т.е. по содержанию сахара в корнеплодах гибриды сахарной свеклы Гримм и Джакета сущуственно не различаются.

***Задание 1*:**

Пример 1. *При определении содержания сухих веществ в яблоках были получены следующие данные:* $\bar{x}$ *=15,34%;* $s_{\bar{x}}$ *= 1,29%; n = 15. Определить доверительные интервалы при 95 и 99% уровне вероятности.*

Пример 2. *При определении содержания крахмала в клубнях сортов картофеля были получены следующие данные:*

Адретта $\bar{x}$ *=17,32%;* $s_{\bar{x}}$ *= 1,71%; n = 12;*

Романо $\bar{x}$ *=13,51%;* $s_{\bar{x}}$ *= 1,43%; n = 18.*

Рассчитать доверительные интервалы при 95 и 99% уровне вероятности, определить существенность различий между средними.

***Задание 2*:**

Пример 1. *При определении содержания крахмала в клубнях картофеля были получены следующие данные:* $\bar{x}$ *=21,19%;* $s_{\bar{x}}$ *= 2,01%; n = 16.Определить доверительные интервалы при 95 и 99% уровне вероятности.*

Пример 2. *При определении содержания сухих веществ в плодах сортов яблонь были получены следующие данные:*

Папировка: $\bar{x}$ *= 14,10%;* $s_{\bar{x}}$ *= 0,13%; n = 15;*

Мелба: $\bar{x}$ *= 13,20%;* $s_{\bar{x}}$ *= 0,12%; n = 13.*

Рассчитать доверительные интервалы при 95 и 99% уровне вероятности, определить существенность различий между средними.

***Задание 3*:**

Пример 1. *При определении содержания белка в зерне сои были получены следующие данные:* $\bar{x}$ *=32,18%;* $s_{\bar{x}}$ *= 3,09%; n = 20. Определить доверительные интервалы при 95 и 99% уровне вероятности.*

Пример 2. *При определении массы зерна в початке гибридов кукурузы были получены следующие данные:*

Альтиус: $\bar{x}$ *= 63,67 г;* $s_{\bar{x}}$ *= 6,45; n = 18;*

Термо: $\bar{x}$ *= 69,74;* $s_{\bar{x}}$ *= 7,01; n = 16.*

Рассчитать доверительные интервалы при 95 и 99% уровне вероятности, определить существенность различий между средними.

***Задание 4*:**

Пример 1. *При определении массы зерна в початке кукурузы были получены следующие данные:* $\bar{x}$ *=67,52 г;* $s_{\bar{x}}$ *= 6,58 г; n = 25.*

Определить доверительные интервалы при 95 и 99% уровне вероятности.

Пример 2. *При определении содержания белка в зерне гибридов сои были получены следующие данные:*

Подолянка: $\bar{x}$ *=34,72%;* $s_{\bar{x}}$ *= 3,28%; n = 22;*

Винничанка: $\bar{x}$ *=37,23%;* $s_{\bar{x}}$ *= 3,31%; n = 24.*

Рассчитать доверительные интервалы при 95 и 99% уровне вероятности, определить существенность различий между средними.

Контрольные вопросы:

1. Что такое точечная оценка? Виды точечной оценки
2. Что такое доверительный интервал?
3. Оценка существенности различий между средними по доверительным интервалам

Практическое занятие № 4
ПОСТРОЕНИЕ ВАРИАЦИОННОГО РЯДА

Использование выборок большого объема позволяет снизить величину статистической ошибки, но получаемый при этом значительный массив данных неудобен для проведения статистической характеристики. Для группировки наблюдений в большой выборке, где число значений варьирующего количественного признака 30 и более, используются вариационные ряды.

Пример. При изучении признака «*Масса плодов перца сорта Золотой Юбилей*» были получены следующие результаты, г:

92	96	86	102	93	77	98	95	83	90
97	82	104	94	99	95	86	105	91	110
93	89	103	75	91	87	100	83	88	94
99	105	93	96	84	98	92	105	107	94

85	87	104	99	91	103	89	93	88	100

Необходимо построить вариационный ряд и гистограмму распределения частот.

При проведении наблюдения был получен достаточно большой объем данных: **n=50**. Построение вариационного ряда начинается с определения количества интервалов, на которые надо разбить выборку. Для определения оптимального количества интервалов ***k*** из общего числа наблюдений **n** извлекается квадратный корень. Если полученное значение не целое, его следует округлить до целого, т.к. количество интервалов не может быть выражено дробным числом. В нашем случае $k = \sqrt{n} = 7$ (15)

Интервалы, кроме первого и последнего, должны быть одинаковых размеров. Величина первого и последнего интервалов может быть немного больше или меньше в связи с тем, что в них попадает меньше значений признака.

Для успешной разбивки выборки на интервалы рассчитывают шаг интервала **i**. Для этого находят самое большое **Xmax** и самое маленькое **Xmin** значение признака, находят между ними разность, которую делят на количество интервалов **k**. Полученное частное округляем до того же уровня точности, которое мы использовали для записи значений признака. Т.е. если значения признака даны в целых числах, округляем до целых; если значения признака даны с точностью до десятых, то округляем до десятых и т.д. В нашем случае

$$i = \frac{X\max - X\min}{k} = \frac{110-75}{7} = 5 \quad (16)$$

Разбиваем весь массив данных на интервалы. Интервал представляет собой ряд данных, для более краткой записи интервала используют только первое и последнее число (нижняя и верхняя границы) между которыми ставят знак интервала $\div$. Нижней границей первого интервала является **Xmin,** для определения верхней границы первого интервала к **Xmin** прибавляют шаг интервала **i**. В нашем случае первый интервал имеет следующий вид **75 ÷ 80.** Второй интервал начинается с числа, идущего после верхней границы интервала, т.е. нижней грани-

цей второго интервала будет число **81**. Для определения верхней границы второго интервала к нижней границе мы прибавляем **«i - единица точности»**. В нашем примере значения признака даны в целых, т.е. единицей точности является 1. Таким образом, верхнюю границу второго интервала мы определяем как ***81+(5-1) = 81+4=85.*** Этот же алгоритм мы используем для определения последующих интервалов.

Мы получили следующие интервалы:

I. **75 ÷ 80**
II. **81 ÷ 85**
II. **86 ÷ 90**
IV. **91 ÷ 95**
V. **96 ÷ 100**
VI. **101 ÷ 105**
VII. **106 ÷ 110**

На следующем этапе необходимо определить сколько значений признака входит в каждый интервал. Для того чтобы не переписывать еще раз значения признака, используют метод штрихов или конвертиков. Более простым является метод штрихов. Сущность этого метода заключается в том, что вместо значения признака напротив соответствующего интервала мы рисуем штрих, затем подсчитываем количество штрихов и определяем, сколько значений признака находятся в полученных интервалах, т.е. определяем частоту встречаемости признака **f**. Сумма частот $\sum f$ должна быть равна общему числу значений признака **n,** т.е. $\sum f = n = 50$

Заполняем таблицу:

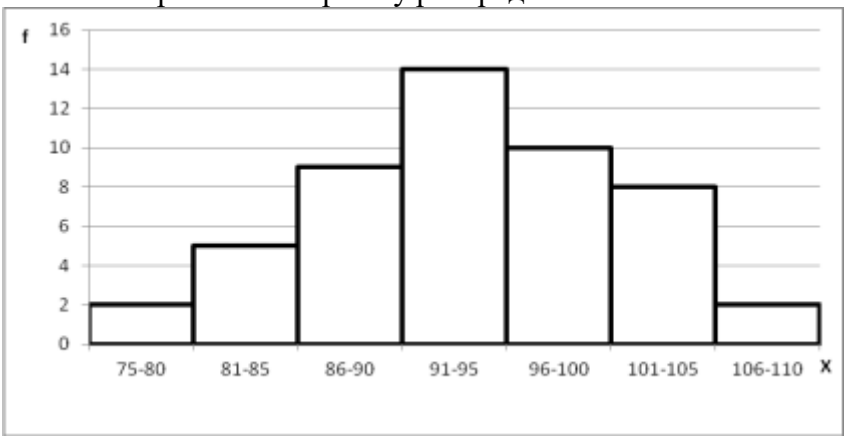
Интервал	Разноска значений признака	Частота, f	X
75 ÷ 80	II	2	78
81 ÷ 85	IIIII	5	83
86 ÷ 90	IIIIIIII	9	88
91 ÷ 95	IIIIIIIIIIII	14	93

96 ÷ 100	IIIIIIIII	10	98
101 ÷ 105	IIIIIII	8	103
106 ÷ 110	II	2	108
Сумма	**-**	**50**	**-**

X – середина интервала, для определения X складываем верхнюю и нижнюю границы интервала и полученное слагаемое делим на 2. Так, для первого интервала X=(75+80):2=155:2=77,5 округляем до целых 78; для второго X=(81+85):2=166:2=83 и т.д.

Таким образом, заполнив столбцы «X» и «f», мы получаем вариационный ряд, где каждому значению признака X соответствует частота f.

Строим гистограмму распределения частот:



Задание. Согласно N примера и варианта построить вариационный ряд и гистограмму распределения частот результатов наблюдения.

Шифр задания к работе

№ варианта	Номера колонок	№ варианта	Номера колонок	№ варианта	Номера колонок
1	1 2 3 4 5	**11**	1 3 4 7 9	**21**	3 7 8 9 10
2	2 3 4 5 6	**12**	4 5 8 9 10	**22**	1 4 5 6 9
3	3 4 5 6 7	**13**	1 2 3 6 9	**23**	1 5 6 7 10
4	4 5 6 7 8	**14**	2 3 4 8 10	**24**	2 4 5 7 8
5	5 6 7 8 9	**15**	3 4 5 9 10	**25**	1 5 7 9 10
6	6 7 8 9 10	**16**	1 3 5 6 8	**26**	2 3 4 8 10
7	1 3 5 7 9	**17**	4 6 7 9 10	**27**	3 5 6 8 9
8	2 4 6 8 10	**18**	2 5 6 8 9	**28**	2 4 6 7 9
9	1 2 5 6 9	**19**	3 5 6 8 10	**29**	3 4 5 8 9
10	2 3 5 7 10	**20**	1 2 4 5 7	**30**	4 6 7 8 9

Пример 1. *Масса плодов сливы сорта Венгерка итальянская, г*

Значение признака по колонкам

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
54	45	44	46	43	38	46	39	49	40
39	37	42	47	35	43	47	36	40	42
31	48	45	57	45	40	40	41	44	60
43	40	58	41	49	35	33	48	52	43
38	35	46	36	52	39	44	45	42	44
45	56	51	45	46	54	42	38	50	35
49	44	41	32	47	41	37	45	43	39
46	53	44	35	39	47	40	48	34	51
34	50	45	42	46	59	53	40	41	47
45	39	36	54	42	46	43	62	46	41

Пример 2. *Высота стеблей пшеницы сорта Крошка, см*

Значение признака по колонкам

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
57	63	74	68	85	73	72	84	66	57
63	58	87	77	72	61	63	78	91	80
84	72	59	71	78	63	71	66	70	69
53	68	68	60	60	72	52	64	58	65
69	79	86	59	61	87	79	83	78	74
75	83	67	78	65	76	60	70	85	67
64	65	61	54	80	82	73	90	62	71
68	69	82	66	73	74	68	69	75	54
80	88	60	79	89	65	81	59	64	79
62	76	70	81	74	80	75	72	77	84

Выполнение работы

1. Записать название примера, номер вариант.
2. Заполнить рабочую таблицу

Рабочая таблица

Номера колонок	Значение признака, X

3. Подсчитать количество значений признака **n =**
4. Определить количество интервалов по формуле (15)

$k = \sqrt{n} =$

5. Рассчитать величину интервала по формуле (16)

$$i = \frac{X\max - X\min}{k} =$$

6. Заполнить таблицу

Интервал	Разноска значений признака	Частота, f	X
Сумма	-		-

7. Построить гистограмму распределения частот:

Контрольные вопросы к работе:

1. Что такое вариационный ряд
2. В каких случаях он рассчитывается
3. Какой показатель является основной характеристикой вариационного ряда
4. Какие показатели являются структурными характеристиками вариационного ряда, их краткая характеристика

Практическое занятие № 5
СТАТИСТИЧЕСКАЯ ОБРАБОТКА РЕЗУЛЬТАТОВ НАБЛЮДЕНИЙ ПРИ КОЛИЧЕСТВЕННОЙ ИЗМЕНЧИВОСТИ

Количественная изменчивость – эта такая изменчивость, при которой значение варьирующего признака имеет числовое выражение. Она бывает прерывистой (дискретной) и непрерывной. При *прерывистой* изменчивости значения признака выражаются только целыми числами; при *непрерывной* количественной изменчивости значение признака могут иметь любую величину в зависимости от точности, принимаемой для характеристики данного признака.

При проведении наблюдения исследователь получает ряд значений изучаемого варьирующего признака X, который подвергается статистической обработке. Основными статистическими характеристиками количественной изменчивости являются: *средняя арифметическая, дисперсия, стандартное отклонение, коэффициент вариации, ошибка выборочной средней и относительная ошибка выборочной средней.*

Средняя арифметическая $\bar{x}$ – это обобщенная, абстрактная характеристика совокупности. Она не содержит полной информации о варьирующих объектах. При одинаковых средних характеризуемые признаки могут отличаться по величине вариации.

Различают простую и взвешенную среднюю арифметическую. *Простая средняя арифметическая* рассчитывается для выборок малого объема по формуле:

$$\bar{x} = \frac{\sum X}{n} \quad (17)$$

где ΣX – сумма всех значений признака,

n – количество значений признака, или объем выборки.

Взвешеная средняя арифметическая рассчитывается для сгруппированных данных по формуле:

$$\bar{x} = \frac{f_1 \cdot x_1 + f_2 \cdot x_2 + \ldots + f_n \cdot x_n}{f_1 + f_2 + \ldots + f_n} = \frac{\sum fX}{n} \quad (18)$$

где n – значение признака,

f – частота встречаемости каждого признака,

n – объем выборки.

Основное свойство средней арифметической заключается в равенстве суммы всех положительных и всех отрицательных отклонений от нее, т.е. сумма всех отклонений вариант равна 0.

Дисперсия s^2, или варианса, или средний квадрат рассчитывается как отношение суммы квадратов отклонений среднего арифметического от каждого значения признака к числу степеней свободы.

Для выборки малого объема она рассчитывается по формуле:

$$s^2 = \frac{\sum (X - x)^2}{n-1} = \qquad (19)$$

где n-1 – число степеней свободы, или количество варьирующих величин.

Для сгруппированных отклонений:

$$s^2 = \frac{\sum f \cdot (X - x)^2}{n-1} = \qquad (20)$$

*Дисперсия показывает квадрат среднего отклонения значений признака **X** от средней арифметической.*

***Стандартное отклонение* s,** или среднее квадратическое отклонение получают путем извлечения квадратного корня из дисперсии, т.е. $s = \sqrt{s^2}$. *Стандартное отклонение показывает величину среднего отклонения значений признака **X** от средней арифметической.*

***Коэффициент вариации* V –** является относительным показателем изменчивости и представляет собой отношение стандартного отклонения к средней арифметической, выраженное в процентах:

$$V = \frac{s}{x} \cdot 100 = \qquad (21)$$

Коэффициент вариации показывает *степень* изменчивости признака: изменчивость считается незначительной, если коэффициент вариации не превышает 10%, средним – если он колеблется от 10 до 20 и значительной – если он более 20%.

Коэффициента вариации используется только в том случае, если признак имеет только ***положительные*** значения.

Ошибка выборочной средней или ошибка выборки s_x является мерой отклонения выборочной средней арифметической от средней генеральной совокупности. Ошибки выборки возникают вследствие неполной репрезентативности (представительности), т.к. не все возможные значения признака, имеющиеся в генеральной совокупности, попадают в выборку. Она свой-

ственна только выборочному методу. Величина ошибки выборочной средней зависит от степени варьирования и от объема выборки.

Ошибка выборочной средней прямо пропорциональна стандартному отклонению и обратно пропорциональна объему выборки, т.е.

$$s_x = \frac{s}{\sqrt{n}} = \quad (22)$$

Ошибка выборочной средней выражается в тех же единицах измерения, что и варьирующий признак и приписывается к соответствующим средним со знаком ±, т.е. $\bar{\chi} \pm \mathbf{s_x}$.

Пример. Масса зерен с одного растения ячменя сорта Донской, г.

Целью данного наблюдения является определение выравненности признака «*масса зёрен с одного растения*» у ячменя сорта Донской. Для достижения этой цели следует провести статистический анализ полученных данных:

10,5	12,4	11,0	11,1	11,6	12,4	11,4	13,2	11,4	10,8
11,9	9,6	11,8	12,2	10,3	11,7	12,0	11,3	12,5	11,0
12,3	12,0	10,4	11,5	10,0	10,8	10,6	11,8	11,6	11,1
11,2	10,9	12,8	11,3	12,6	11,2	11,9	12,7	10,7	11,3

Объем выборки **n=40,** поэтому для проведения статистического анализа следует построить вариационный ряд. Определяем количество интервалов k и шаг интервала i:

$$k = \sqrt{n} = 6{,}3 \approx 6 \ ; \quad i = \frac{X\max - X\min}{k} = \frac{13{,}2 - 9{,}6}{6} = 0{,}6$$

Строим вариационный ряд и заполняем вспомогательную расчетную таблицу:

Интервал	Частота, f	X	$f \cdot X$	X^2	$f \cdot X^2$
9,6 ÷ 10,2	3	9,9	29,7	98,0	294,0
10,3 ÷ 10,8	6	10,6	63,6	112,4	674,4
10,9 ÷ 11,4	12	11,2	134,4	125,4	1504,8
11,5 ÷ 12,0	10	11,8	118,0	139,2	1392,0
12,1 ÷ 12,6	6	12,4	74,4	153,8	922,8
12,7 ÷ 13,2	3	13,0	39,0	169,0	507,0
Суммы	**40**	**-**	**459,1**	**-**	**5295,0**

Проводим статистическую обработку данных наблюдения:

Средняя арифметическая взвешенная

$\bar{x} = \left(\sum f \cdot X\right) : n =$ 459,1:40=**11,5**

Сумма квадратов отклонений $\sum (X - \bar{x})^2 = \sum f \cdot X^2 - (\Sigma\, f \cdot X)^2 : n =$

= 5295,0 - 459,1^2:40 = 5295,0 – 210772,8:40 =5295,0 – 5269,3 = 25,7

Дисперсия $s^2 = \dfrac{\sum (X - \bar{x})^2}{n-1} = \dfrac{25{,}7}{40-1} = \dfrac{25{,}7}{39} = 0{,}66$

Стандартное отклонение $s = \sqrt{s^2} = \sqrt{0{,}66} = 0{,}81$

Коэффициент вариации $V = \dfrac{s}{\bar{x}} \cdot 100 = \dfrac{0{,}81}{11{,}5} \cdot 100 = 7{,}0$

Значение коэффициента вариации меньше 10%, значит изменчивость признака незначительная, следовательно, признак «*масса зерен с одного растения*» у ячменя сорта Донской выровнен.

Ошибка выборочной средней $s_x = \dfrac{s}{\sqrt{n}} = \dfrac{0{,}81}{\sqrt{40}} = \dfrac{0{,}81}{6{,}32} = 0{,}1$

Доверительный интервал для 95% уровня вероятности равен:

$\bar{x} \pm t \cdot s_{\bar{x}}$ 11,5±2,03×0,1 11,5±0,2 или **11,3÷11,7**

Значения t_{05} берем из статистической таблицы (приложение 2) для v=40-1=39 степеней свободы. Для v=39 значение t_{05} = 2,03.

Следовательно, генеральная средняя признака «*масса зерен с 1 растения*» у ячменя сорта Донской с вероятностью 95% находится в интервале от 11,3 до 11,7 г. Вероятность того, что значение генеральной средней не попало в границы интервала, составляет 5%.

Задание. Согласно N варианта провести статистический анализ результатов наблюдения.

Шифр задания к работе

№ варианта	Номера колонок	№ варианта	Номера колонок	№ варианта	Номера колонок
1	1 2 3 4	**6**	1 3 5 7	**11**	3 4 7 8
2	2 3 4 5	**7**	2 4 6 8	**12**	2 3 6 7
3	3 4 5 6	**8**	3 5 7 9	**33**	3 4 7 8
4	4 5 6 7	**9**	4 6 8 10	**14**	5 6 9 10
5	7 8 9 10	**10**	1 2 5 6	**15**	1 4 7 10

Пример 1. *Масса плодов тыквы сорта Чародейка, кг*

Значение признака по колонкам

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
4,0	4,1	5,3	4,5	6,4	5,1	4,2	4,8	4,3	3,6
5,1	5,2	4,5	5,5	5,3	3,7	4,8	4,1	4,7	5,4
5,7	3,3	5,2	6,0	5,6	5,0	3,2	5,2	5,5	4,8
5,4	5,8	4,3	4,9	5,0	4,4	3,9	5,8	6,4	4,5
4,4	5,9	6,3	5,4	4,3	4,0	5,7	4,4	4,6	6,0
4,1	5,3	5,1	4,8	5,5	4,7	4,9	5,0	3,4	5,7
3,7	5,2	5,8	5,4	4,4	5,6	4,3	4,8	5,3	5,3
4,6	4,7	4,3	4,2	3,0	4,9	5,3	4,2	4,5	4,1
6,2	5,4	3,6	5,3	5,1	6,2	4,9	3,8	5,9	6,5
5,5	4,2	5,2	3,7	3,9	5,5	5,1	5,4	3,7	4,6

Пример 2. *Содержание сахара в корнеплодах сахарной свеклы гибрида Ванесса, %*

Значение признака по колонкам

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
13,9	16,6	15,7	14,2	15,5	13,2	16,1	13,1	16,3	18,7
15,1	15,7	13,5	15,3	18,9	15,6	15,4	15,5	17,8	15,1
14,5	12,1	16,7	17,4	19,9	17,5	12,2	14,3	16,0	14,5
16,0	16,4	14,8	16,2	15,0	15,4	15,3	15,7	13,5	17,0
13,1	17,8	14,0	13,7	15,4	18,6	14,9	17,3	14,7	16,4
15,6	15,2	15,8	14,9	16,1	16,5	14,2	15,4	15,9	18,0
18,0	13,3	14,3	12,8	17,2	16,6	13,3	14,8	12,8	16,8
14,6	16,4	15,6	15,5	16,3	14,1	16,5	16,9	15,2	14,6
18,7	14,4	17,6	16,8	13,8	13,7	15,1	18,4	14,4	15,8
16,5	14,7	19,0	16,2	14,1	15,0	17,7	16,2	16,7	13,9

Выполнение работы

1. Записать название примера, номер варианта
2. Заполнить рабочую таблицу

Номера колонок	Значение признака, X

3. Подсчитать количество значений признака **n =**
4. Определить количество интервалов $k = \sqrt{n} =$
5. Рассчитать величину интервала $i = \dfrac{X\max - X\min}{k} =$
6. Заполнить расчетную таблицу

Интервал	Разноска значений признака	Частота, f	X	$f \cdot X$	X^2	$f \cdot X^2$
	Суммы		-		-	

7. Провести статистическую обработку данных

Средняя арифметическая $\bar{x} = \left(\sum f \cdot X\right) : n =$

Сумма квадратов отклонений $\sum (X - \bar{x})^2 = \sum f \cdot X^2 - \left(\sum f \cdot X\right)^2 : n =$

Дисперсия $s^2 = \frac{\sum (X - \bar{x})^2}{n-1} =$

Стандартное отклонение $s = \sqrt{s^2} =$

Коэффициент вариации $V = \frac{s}{\bar{x}} \cdot 100 =$

Ошибка выборочной средней $s_x = \frac{s}{\sqrt{n}} =$

Доверительный интервал $\bar{x} \pm t_{05} \cdot s_x$

при v = n-1= t_{05}=

Контрольные вопросы к работе:

1. Что такое наблюдение и эксперимент?
2. Что называется изменчивостью? Виды изменчивости
3. Какая изменчивость называется количественной? Виды количественной изменчивости

4. Основные статистические характеристики количественной изменчивости

Практическое занятие № 6
СТАТИСТИЧЕСКАЯ ОБРАБОТКА РЕЗУЛЬТАТОВ НАБЛЮДЕНИЙПРИ КАЧЕСТВЕННОЙ ИЗМЕНЧИВОСТИ

Качественная или атрибутивная изменчивость – эта такая изменчивость, при которой значения признака не имеют числового выражения. Основными статистическими характеристиками качественной изменчивости являются: *доля признака, показатель изменчивости, коэффициент вариации и ошибка выборочной доли.*

Доля признака характеризует относительную частоту встречаемости отдельной варианты (значения признака) в данной совокупности и обозначается через p_1, p_2, …, p_n. Выражается доля признака в частях единицы или в процентах. Сумма всех долей признака в пределах совокупности равна 1 или 100%. Доля признака отражает вероятность появления данного признака в изучаемой совокупности и определяется как отношение численности каждого из членов ряда n_1, n_2, …, n_n к объему выборки N:

$$p_1 = \frac{n_1}{N} \quad (23); \quad p_2 = \frac{n_2}{N}; \quad (24)\ldots; \quad p_n = \frac{n_n}{N} \quad (25)$$

При альтернативной изменчивости доля одного признака обозначается через p, другого - через q.

Показатель изменчивости качественного признака **s** характеризует варьирование величин ряда относительно друг друга. При альтернативной изменчивости он определяется по формуле:

$$s = \sqrt{pq} \quad (26)$$

В зависимости от соотношения p и q значение показателя изменчивости изменяется от 0 до 0,5. Максимальная изменчивость качественного признака $\mathbf{s_{max}}$ наблюдается при p=q=0,5.

Если количество градаций признака больше двух, то показатель изменчивости определяется по формуле:

$$s = \sqrt[n]{p_1 \cdot p_2 \cdot \ldots \cdot p_n} \qquad (27)$$

где n – число градаций признака.

Значение максимальной изменчивости уменьшается с увеличением числа градаций признака.

Коэффициент вариации качественного признака V – это отношение показателя изменчивости признака к максимально возможной изменчивости, выраженное в процентах:

$$V = \frac{s}{s_{\max}} \cdot 100 \qquad (28)$$

Коэффициент вариации характеризует относительную степень изменчивости исследуемых признаков и используется для сравнительной оценки выравненности различных совокупностей. Максимальное значение V равное 100% наблюдается при $\mathbf{s=s_{max}}$.

Ошибка выборочной доли $\mathbf{s_p}$ определяет меру отклонения доли признака выборочной совокупности p от доли его во всей генеральной совокупности P вследствие неполной репрезентативности выборки. Вычисляется по формуле:

$$s_p = \frac{s}{\sqrt{n}} = \qquad (29)$$

Пример. В выборке из 200 зерен озимой пшеницы сорта Мироновская 808, взятой для определения стекловидности зерен, 154 зерен было стекловидными, 46 – мучнистыми.

Доля признака. В нашем примере изучается альтернативная изменчивость, когда признак имеет только два взаимо-

исключающих значения. В данном случае «*стекловидные зерна*» и «*мучнистые зерна*». Долю признака «*стекловидные зерна*» обозначаем символом **p**, в долю признака «*мучнистые зерна*» – символом **q**.

Объем выборки, т.е. общее количество наблюдений (общее количество зерен) N = 200.

$$p = \frac{n}{N} = \frac{154}{200} = 0{,}77$$

Т.к. сумма доле признака всегда равна 1 $q = 1 - p = 1\text{-}0{,}77 = 0{,}23$

Показатель изменчивости $s = \sqrt{p \cdot q} = \sqrt{0{,}77 \cdot 0{,}23} = \sqrt{0{,}177} = 0{,}42$

Коэффициент вариации. При альтернативной изменчивости максимальное значение показателя изменчивости s_{max}=0,5

$V = \frac{s}{S_{\max}} \cdot 100 = \frac{0{,}42}{0{,}5} = 0{,}84$ или 84%

Ошибка выборочной доли $S_p = \sqrt{\frac{p \cdot q}{N}} = \sqrt{\frac{0{,}177}{200}} = \sqrt{0{,}00088} = 0{,}03$

Задание. Согласно N варианта провести статистический анализ результатов наблюдения.

Пример 1. *Панцирность семянок подсолнечника*

Вариант	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Объем выборки (N)	250	300	150	230	320	270	240	180	220
Кол-во семянок с панцирным слоем (n)	210	270	125	190	290	250	210	165	195

Пример 2. *Пораженость проростков гороха корневыми гнилями*

Вариант	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Объем выборки (N)	120	150	170	130	240	270	300	320	350
Кол-во проростков, пораженных корневыми гнилями (n)	80	60	100	54	112	136	210	190	220

Выполнение работы

N = , n =

Доля признака $p = \frac{n}{N} =$ $q = 1 - p =$

Показатель изменчивости $s = \sqrt{p \cdot q}$

Коэффициент вариации $V = \frac{s}{s_{max}} \cdot 100 =$

Ошибка выборочной доли $s_p = \sqrt{\frac{p \cdot q}{N}} =$

Контрольные вопросы к работе:

1. Какая изменчивость называется качественной? Что такое альтернативная изменчивость?
2. Основные статистические характеристики качественной изменчивости.

Практическое занятие №7

ОЦЕНКА СУЩЕСТВЕННОСТИ РАЗНОСТИ СРЕДНИХ ДВУХ ВЫБОРОК ПРИ КОЛИЧЕСТВЕННОЙ ИЗМЕНЧИВОСТИ

Распределение результатов измерений, полученных при изучении выборки, подчиняется статистическим законам. Для сравнения данных двух или нескольких выборок используют различные статистические методы, которые сводятся к проверке статистических гипотез. ***Статистической гипотезой*** называют научное предположение о тех или иных статистических законах распределения случайных величин, которое может быть проверено на основе выборки. В агрономических исследованиях задача сводится к проверке нулевой гипотезы, которая предполагает, что между выборками нет существенных различий.

Для оценки существенности различий между средними 2 выборок чаще всего используют ***критерий существенности***,

который определяется как отношение разности средних к ее ошибке. Критерий существенности сравнивается с критерием Стьюдента. Значение критерия Стьюдента берется из статистической таблицы с учетом принятого уровня значимости и числа степеней свободы, которое рассчитывается по формуле

$$v = n_1 + n_2 - 2 \quad (30)$$

n_1– число наблюдений в первой выборке, n_2 – число наблюдений во второй выборке.

Знак критерия существенности не несет смысловой нагрузки, поэтому он не учитывается.

Если $t_{факт} \geq t_{теор}$, нулевая гипотеза отвергается, следовательно разница между выборками существенна; если $t_{факт} < t_{теор}$, нулевая гипотеза принимается, следовательно разница несущественна.

Пример. Средняя масса (г) плодов томата сортов Алекс (вариант 1) и Золотая осень (вариант 2).

В выборках по 10 плодов определили массу каждого плода. Был получены следующие данные:

Алекс	62	81	74	59	84	87	93	72	68	79
Золотая осень	48	44	56	52	59	62	60	46	57	64

Следует определить, различаются ли сорта томата по средней массе плодов.

1. *Определяем статистические характеристики данных варианта 1.*

Выборка малого объема, содержит лишь 10 значений (n=10), поэтому рассчитываем ***простую среднюю арифметическую***:

$$\bar{X} = \frac{\sum X}{n} = \frac{62+81+74+59+84+87+93+72+68+79}{10} = \frac{759}{10} = 75{,}9 \approx 76$$

Корректирующий фактор $C = \frac{(\sum X)^2}{n} = \frac{759^2}{10} = \frac{576081}{10} = 57608$

Сумма квадратов отклонений

$$\sum(X-x)^2=\sum X^2-C=62^2+81^2+74^2+59^2+84^2+87^2+93^2+72^2+68^2+$$
$$+79^2-57608=3844+6561+5476+3481+7056+7569+8649+5184+4624+$$
$$+6241-57608=58685-57608=1077$$

Дисперсия $s^2=\frac{\sum(X-x)^2}{n-1}=\frac{1077}{10-1}=\frac{1077}{9}=119{,}7\approx 120$

Стандартное отклонение $s=\sqrt{s^2}=\sqrt{120}=11$

Коэффициент вариации $V=\frac{s}{\bar{x}}\cdot 100=\frac{11}{76}\cdot 100=14$

Ошибка выборочной средней $s_x=\frac{s}{\sqrt{n}}=\frac{11}{\sqrt{10}}=\frac{11}{3{,}2}=3{,}4\approx 3$

2. *Определяем статистические характеристики данных варианта 2.*

Средняя арифметическая

$$\bar{x}=\frac{\sum X}{n}=\frac{48+44+56+52+59+62+60+46+57+64}{10}=\frac{548}{10}=54{,}8\approx 55$$

Корректирующий фактор $C=\frac{(\sum X)^2}{n}=\frac{548^2}{10}=\frac{300304}{10}=30030$

Сумма квадратов отклонений

$$\sum(X-x)^2=\sum X^2-C=48^2+44^2+56^2+52^2+59^2+62^2+60^2+46^2+$$
$$+57^2+64^2-30030=2304+1936+3136+2704+3481+3844+3600+$$
$$+2116+3249+4096-30030=30466-30030=436$$

Дисперсия $s^2=\frac{\sum(X-x)^2}{n-1}=\frac{436}{10-1}=\frac{436}{9}=48{,}4\approx 48$

Стандартное отклонение $s = \sqrt{s^2} = \sqrt{48} = 6{,}9 \approx 7$

Коэффициент вариации $V = \frac{s}{\bar{x}} \cdot 100 = \frac{7}{55} \cdot 100 = 13$

Ошибка выборочной средней $s_x = \frac{s}{\sqrt{n}} = \frac{7}{\sqrt{10}} = \frac{7}{3{,}2} = 2$

3. Проводим оценку существенности разности между средними по ***критерию существенности***

$$t_{факт} = \frac{d}{s_d} = \frac{x_2 - x_1}{\sqrt{s_{x1}^2 + s_{x2}^2}} = \frac{76-55}{\sqrt{3^2+2^2}} = \frac{21}{\sqrt{9+4}} = \frac{21}{\sqrt{13}} = \frac{21}{3{,}6} = 5{,}83$$

при $v=n_1+n_2-2= 10+10-2=18$

$t_{05}=2{,}10$

$t_{факт}$ *больше* t_{05} *, следовательно между вариантами есть существенная разница.*

Вывод: По массе плодов сорт томата Алекс **существенно** превосходит сорт Золотая осень.

Задание. Согласно N задания провести оценку существенности разности между средними 2 выборок.

Шифр задания к работе

№ задания	№№ колонок	№ задания	№№ колонок
1	1 6	**7**	3 6
2	2 7	**8**	4 7
3	3 8	**9**	1 5
4	4 6	**10**	2 5
5	1 7	**11**	3 5
6	1 8	**12**	4 5

Пример 1. *Содержание крахмала (%) в зерне сорго сортов Гранд (вариант 1) и Жемчужина (вариант 2).*

Вариант 1 (Гранд)				Вариант 2 (Жемчужина)			
Номера колонок				Номера колонок			
1	2	3	4	5	6	7	8
65,2	66,3	67,7	68,1	72,3	74,8	77,0	73,8
64,3	68,1	67,8	65,4	74,8	77,1	72,4	72,1
66,8	67,2	66,5	66,2	75,9	76,4	72,9	77,2
65,7	65,9	64,9	68,8	74,3	75,6	73,3	74,5
66,9	67,3	68,0	69,0	72,8	76,2	74,5	75,6
69,3	67,8	67,1	66,4	76,3	72,2	76,3	76,8
68,1	66,7	65,0	67,3	73,8	74,3	77,1	77,3
65,7	65,7	68,8	68,2	77,0	73,0	72,4	74,8
66,4	67,2	66,2	66,4	73,0	75,1	75,8	73,9
67,7	66,9	68,2	65,3	76,5	74,2	76,0	75,0

Пример 2. *Содержание жира (%) в семянках подсолнечника сортов Лакомка (вариант 1) и Пересвет (вариант 2).*

Вариант 1 (Лакомка)				Вариант 2 (Пересвет)			
Номера колонок				Номера колонок			
1	2	3	4	5	6	7	8
42,3	44,1	46,2	45,1	52,8	53,9	55,0	54,2
44,7	43,8	45,4	44,2	51,7	55,0	53,4	53,1
45,8	42,4	44,4	43,0	53,2	54,2	52,1	52,9
45,9	44,9	43,9	47,2	54,9	53,8	54,4	55,5
46,7	45,5	42,8	46,4	55,0	53,0	53,2	52,8
42,8	47,1	46,1	45,5	54,1	52,9	53,6	54,3
44,9	46,8	45,8	42,9	53,9	52,1	64,1	51,9
47,0	45,4	47,2	44,8	52,1	54,8	51,9	54,8
45,6	43,0	44,3	45,4	52,8	55,2	51,7	52,7
46,6	44,8	45,5	46,3	54,3	54,3	52,0	53,4

Пример 3. *Толщина стенки плодов перца сорта Солнышко (вариант 1) и Золотой Юбилей (вариант 2).*

Вариант 1 (Солнышко)				Вариант 2 (Золотой Юбилей)			
Номера колонок				Номера колонок			
1	2	3	4	6	7	8	9
6,5	8,0	8,2	6,4	10,0	8,2	9,7	8,5
7,4	7,2	7,5	8,0	9,1	8,7	8,2	8,0
7,5	6,4	8,3	5,8	8,4	9,2	9,5	8,7
6,3	7,0	7,0	7,3	8,6	8,4	8,1	10,0
6,0	7,8	7,8	7,9	9,7	9,1	8,7	9,4
6,2	6,7	7,2	6,7	8,2	9,8	9,4	9,9
7,9	6,9	8,0	6,4	9,0	8,6	8,6	8,7
6,8	7,6	7,2	6,0	9,4	10,0	10,2	8,2
8,2	6,5	6,9	6,2	8,1	10,1	9,1	9,1
6,9	6,3	6,4	5,9	9,3	8,4	8,4	10,2

Пример 4. *Содержание сахара (%) в ягодах винограда сортов Мускат Оттонель (вариант 1) и Мерло (вариант 2).*

Вариант 1 (Мускат Оттонель)				Вариант 2 (Мерло)			
Номера колонок				Номера колонок			
1	2	3	4	6	7	8	9
17,3	20,9	19,3	18,7	25,4	24,0	22,3	19,8
18,7	18,4	18,4	19,4	22,1	23,4	21,2	21,4
20,4	19,7	19,8	18,0	23,9	22,8	19,0	22,8
18,9	17,8	18,7	20,7	22,0	21,9	21,0	23,3
20,1	19,2	18,0	20,0	22,7	21,4	22,8	20,9
17,6	18,9	20,1	18,1	21,9	2,4	21,4	23,8
19,4	20,7	18,5	19,2	24.5	20,9	23,7	22,1
21,0	19,0	18,2	18,9	25,6	21,4	24,2	23,0
18,0	18,4	18,7	18,4	22,8	22,2	19,1	24,1
19,6	19,5	19,9	15,8	23,2	24,4	18,9	23,9

Выполнение работы

1.Записать название примера, номер варианта.

2.Провести статистический анализ варианта 1.

Средняя арифметическая $\bar{x} = \frac{\sum X}{n} =$

Корректирующий фактор $C = \frac{(\sum X)^2}{n} =$

Сумма квадратов отклонений $\sum (X - x)^2 = \sum X^2 - C =$

Дисперсия $s^2 = \frac{\sum (X - x)^2}{n - 1} =$

Стандартное отклонение $s = \sqrt{s^2} =$

Коэффициент вариации $V = \frac{s}{\bar{x}} \cdot 100 =$

Ошибка выборочной средней $s_x = \frac{s}{\sqrt{n}} =$

3. Провести статистический анализ варианта 2.

Средняя арифметическая $\bar{x} = \frac{\sum X}{n} =$

Корректирующий фактор $C = \frac{(\sum X)^2}{n} =$

Сумма квадратов отклонений $\sum (X - x)^2 = \sum X^2 - C =$

Дисперсия $s^2 = \frac{\sum (X - x)^2}{n - 1} =$

Стандартное отклонение $s = \sqrt{s^2} =$

Коэффициент вариации $V = \frac{s}{\bar{x}} \cdot 100 =$

Ошибка выборочной средней $S_x = \frac{s}{\sqrt{n}} =$

4. Провести оценку существенности разности по критерию существенности

$$t_{факт} = \frac{d}{S_d} = \frac{x_2 - x_1}{\sqrt{S_{x1}^2 + S_{x2}^2}} =$$

при $v=n_1+n_2-2=$ $t_{05}=$

Вывод:

Контрольные вопросы к работе:

1. Что называют статистической гипотезой? Нулевая гипотеза.
2. Оценка существенности разности между средними 2 выборок по критерию существенности.

Практическое занятие №8

ОЦЕНКА СУЩЕСТВЕННОСТИ РАЗНОСТИ МЕЖДУ ВЫБОРОЧНЫМИ ДОЛЯМИ ПРИ КАЧЕСТВЕННОЙ ИЗМЕНЧИВОСТИ

Качественная изменчивость не поддается измерению. Частным случаем качественной изменчивости является *альтернативная изменчивость*, при которой признак может принимать только два взаимоисключающих значения.

Основные статистические показатели качественной изменчивости: *доля признака*, *показатель изменчивости*, *коэффициент вариации, ошибка выборочной доли.*

Пример. В выборках по 150 клубней картофеля количество товарных клубней сорта Романо (вариант 1) было 128, сорта Адретта (вариант 2) – 144.

Следует определить, различаются ли эти сорта картофеля по товарности.

1. Определяем статистические характеристики данных варианта 1.

Доля признака. В нашем примере изучается альтернативная изменчивость, когда признак имеет только два взаимоисключающих значения. В данном случае «*товарные клубни*» и «*нетоварные клубни*». Таким образом, долю признака «*товарные клубни*» обозначаем символом **p**, а долю признака «*нетоварные клубни*» – символом **q**.

Объем выборки, т.е. общее количество наблюдений (общее количество клубней) N = 1500, число товарных клубней n=128.

$$p = \frac{n}{N} = \frac{128}{150} = 0{,}85$$

Т.к. сумма доле признака всегда равна 1

$$q = 1 - p = 1\text{-}0{,}85 = 0{,}15$$

Показатель изменчивости $s = \sqrt{p \cdot q} = \sqrt{0{,}85 \cdot 0{,}15} = \sqrt{0{,}1275} = 0{,}36$

Коэффициент вариации. При альтернативной изменчивости максимальное значение показателя изменчивости s_{max}=0,5

$$V = \frac{s}{s_{\max}} \cdot 100 = \frac{0{,}36}{0{,}5} \cdot 100 = 72\%$$

Ошибка выборочной доли

$$s_p = \sqrt{\frac{p \cdot q}{N}} = \sqrt{\frac{0{,}1275}{150}} = \sqrt{0{,}00085} = 0{,}03$$

2. Определяем статистические характеристики данных варианта 2.

Доля признака. N = 150, число растений, n=144.

$$p = \frac{n}{N} = \frac{144}{150} = 0{,}96 \quad q = 1 - p = 1\text{-}0{,}96 = 0{,}04$$

Показатель изменчивости

$$s = \sqrt{p \cdot q} = \sqrt{0{,}96 \cdot 0{,}04} = \sqrt{0{,}0384} = 0{,}20$$

Коэффициент вариации $V = \frac{s}{s_{\max}} \cdot 100 = \frac{0{,}2}{0{,}5} \cdot 100 = 40\%$

Ошибка выборочной доли $s_p = \sqrt{\frac{p \cdot q}{N}} = \sqrt{\frac{0{,}0384}{150}} = \sqrt{0{,}000256} = 0{,}02$

3. Проводим оценку существенности разности между выборочными долями по критерию существенности

$$t_{факт} = \frac{d}{s_p} = \frac{p_2 - p_1}{\sqrt{s_{p1}^2 + s_{p2}^2}} = \frac{0{,}96 - 0{,}85}{\sqrt{0{,}03^2 + 0{,}02^2}} = \frac{0{,}11}{\sqrt{0{,}0009 + 0{,}0004}} = \frac{0{,}11}{\sqrt{0{,}0013}} =$$

$$= \frac{0{,}11}{0{,}04} = 2{,}75$$

Знак критерия существенности не несет смысловой нагрузки, т.е. если критерий существенности получился отрицательным, знак «-» отбрасывается.

При $v=N_1+N_2-2= 150+150-2=298$ $t_{05}=1{,}96$

Сравниваем критерий существенности с критерием Стьюдента 2,75 больше 1,96, т.е. $t_{факт} > t_{05}$, следовательно *различие между выборками существенно*

Вывод. Товарность клубней сорта Адретта *существенно* выше, чем сорта Романо.

Задание. Согласно N задания провести оценку существенности разности между выборочными долями.

Пример 1. *В выборках по 200 плодов груши сортов Любимица Клаппа и Золотистая количество товарных плодов было следующим:*

Вариант	Задание									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Любимица Клаппа	181	184	187	183	177	185	174	178	190	171
Золотистая	162	161	174	167	159	187	153	161	187	192

Пример 2. *В выборках по 150 гроздей винограда сортов Восторг и Мускат гамбургский количество гроздей без горошения было следующим:*

Вариант	Задание									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Восторг	140	131	144	146	144	131	138	124	128	130
Мускат гамбургский	72	101	104	83	97	79	87	92	104	65

Пример 3. *При анализе выборок по 200 растений сои в каждой у сортов Альтаир и Марьяна количество растений, пораженных фузариозом, было следующим:*

Вариант	Задание									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Альтаир	12	8	15	9	13	19	11	14	10	7
Марьяна	31	19	22	21	18	27	32	25	17	24

Пример 4. *При анализе выборок по 200 растений картофеля в каждой у сортов Невский и Изора количество растений, пораженных фитофторозом, было следующим:*

Вариант	Задание									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Невский	21	19	13	18	15	19	14	22	24	17
Изора	42	34	31	40	38	44	35	29	37	49

Выполнение работы

1. Записать название примера, номер варианта.
2. Провести статистический анализ каждой выборки.
3. Провести оценку существенности разности между выборочными долями по критерию существенности

$$t_{факт} = \frac{d}{s_p} = \frac{p_2 - p_1}{\sqrt{s_{p1}^2 + s_{p2}^2}} =$$

t_{05}= $v=N_1+N_2-2$=

Вывод:

Контрольные вопросы к работе:

1. Оценка существенности различий между выборочными долями по критерию существенности

Практическое занятие №9
ОЦЕНКА СООТВЕТСТВИЯ МЕЖДУ ФАКТИЧЕСКИМИ И ТЕОРЕТИЧЕСКИМИ ЧАСТОТАМИ ПО КРИТЕРИЮ ПИРСОНА

Закон распределения χ^2, открытый Пирсоном, является специальным распределением, полученным из случаев нормального распределения для определенно поставленной задачи и при ограниченном числе степеней свободы. Критерий χ^2 используется при качественной изменчивости признака для сравнения эмпирических (фактических) и теоретических частот. χ^2 – критерий, или критерий согласия рассчитывают по следующей формуле:

$$\chi^2 = \sum \frac{(f-F)^2}{F} \quad (30)$$

где, f – частота фактическая, F – теоретическая частота.

Фактическое значение критерия сравнивают с теоретическим, которое берут из статистической таблицы. ли χ^2факт$\geq$ χ^2теор, между частотами есть существенные различия; χ^2факт< χ^2теор, между частотами нет существенных различий.

Пример. Обследовали поля озимой пшеницы на поражённость корневыми гнилями при выращивании после черного

пара и яровых зерновых. По степени поражённости поля разделили на три группы: слабая, средняя и сильная поражённость.

Предшественник	Поражённость, %		
	слабая	средняя	сильная
Чистый пар	45	10	5
Яровые зерновые	12	32	15

Следует определить существенно ли различие в поражённости корневыми гнилями озимой пшеницы, выращиваемой по разным предшественникам.

Нулевая гипотеза предполагает, что поражённость озимой пшеницы не зависит от предшествующей культуры. Для проверки нулевой гипотезы рассчитываем критерий χ^2. В результате обследования мы получили ***фактические*** значения част изучаемого признака «*поражённость корневыми гнилями*» в трех градациях. Для расчета критерия χ^2 следует рассчитать ***теоретические*** значения частот изучаемого признака. Для расчета теоретических частот строим вспомогательную таблицу, в которую заносим фактические значения частот (f).

Предшественник	Поражённость, %						Сумма
	слабая		средняя		сильная		
	f	F	f	F	f	F	
Чистый пар	45		10		5		
Яровые зерновые	12		32		15		
Сумма							

Рассчитываем суммы частот по вертикали и по горизонтали:

Предшественник	Поражённость, %						Сумма
	слабая		средняя		сильная		
	f	F	f	F	f	F	
Чистый пар	45		10		5		**60**
Яровые зерновые	12		32		15		**59**
Сумма	**57**		**42**		**20**		**119**

Для расчета теоретической частоты F сумму фактических частот данного столбца умножаем на сумму фактических частот данной строки и полученное произведение делим на общую сумму всех наблюдений:

слабая поражённость по чистому пару - $F1 = \frac{57 \cdot 60}{119} = \frac{3420}{119} = 29$

слабая поражённость по яровому ячменю - $F2 = \frac{57 \cdot 59}{119} = \frac{3363}{119} = 28$

средняя поражённость по чистому пару - $F3 = \frac{42 \cdot 60}{119} = \frac{2520}{119} = 21$

средняя поражённость по яровому ячменю - $F4 = \frac{42 \cdot 59}{119} = \frac{2478}{119} = 21$

сильная поражённость по чистому пару - $F5 = \frac{20 \cdot 60}{119} = \frac{1200}{119} = 10$

сильная поражённость по яровому ячменю - $F6 = \frac{20 \cdot 59}{119} = \frac{1180}{119} = 10$

На основании рассчитанных фактических частот заполняем таблицу:

Предшественник	Поражённость, %						Сумма
	слабая		средняя		сильная		
	f	F	f	F	f	F	
Чистый пар	45	29	10	21	5	10	**60**
Яровые зерновые	12	28	32	21	15	10	**59**
Сумма	**57**		**42**		**20**		**119**

Рассчитываем критерий Пирсона по формуле (30):

$$\chi^2 = \sum \frac{(f-F)^2}{F} = \frac{(45-29)^2}{29} + \frac{(12-28)^2}{28} + \frac{(10-21)^2}{21} + \frac{(32-21)^2}{21} +$$

$$+ \frac{(5-10)^2}{10} + \frac{(15-10)^2}{10} = \frac{16^2}{29} + \frac{-16^2}{28} + \frac{-11^2}{21} + \frac{11^2}{21} + \frac{-5^2}{10} + \frac{5^2}{10} =$$

$$= \frac{256}{29} + \frac{256}{28} + \frac{121}{21} + \frac{121}{21} + \frac{25}{10} + \frac{25}{10} = 8{,}83 + 9{,}14 + 5{,}76 + 5{,}76 + 2{,}5 + 2{,}5 = 34{,}49$$

Фактическое значение критерия Пирсона сравниваем с теоретическим значением, которое берем из статистической таблицы для 5% уровня значимости и числа степеней свободы $v = (c-1)\cdot(k-1)$, где с – число строк, k – число колонок. При $v = (c-1)\cdot(k-1) = (2-1)\cdot(3-1) = 2$ значение критерия χ^2_{05} = 5,99

Фактическое значение критерия больше теоретического, следовательно нулевая гипотеза отвергается, между частотами есть существенные различия.

Вывод. Поражённость пшеницы корневыми гнилями существенно зависит от предшественников. При выращивании озимой пшеницы после яровых зерновых поражённость корневыми гнилями *существенно* выше, чем при выращивании после чистого пара.

Задание. Согласно N примера провести оценку соответствия между эмпирическими и теоретическими частотами.

Пример 1. *Изучали влияние норм расхода волатона на гибель колорадского жука. Из 150 жуков при норме расхода препарата 0,5 л/га погибло 82, а при норме 0,75 л/га – 108. Определить существенно ли различие в гибели колорадского жука от нормы расхода волатона.*

Норма расхода волатона, л/га	Количество жуков				Сумма
	живые		погибшие		
	f	F	f	F	
0,5					
0,75					
Сумма					

χ^2_{05} = 3,84

Пример 2. *Проведено обследование полей картофеля на поражённость фитофторозом. Посадка была произведена перебранными и не перебранными клубнями. Определить, существует ли зависимость между пораженностью полей фитофторозом и подготовкой клубней к посадке.*

Варианты подготовки клубней к посадке	Число полей, поражённых фитофторозом						Сумма
	слабо		средне		сильно		
	f	F	f	F	f	F	
Перебранные	30		15		5		
Не перебранные	23		12		15		
Сумма							

$\chi^2_{05} = 5{,}99$

Пример 3. *У яровой пшеницы проверяли эффективность отборов по продуктивности растений. Проверить нулевую гипотезу о независимости продуктивности родительских форм и урожая с делянки у потомства.*

Продуктивность родительских форм	Число делянок с урожаем у потомства F1						Сумма
	высокий		средний		низкий		
	f	F	f	F	f	F	
Высокая	10		17		7		
Низкая	7		13		9		
Сумма							

$\chi^2_{05} = 5{,}99$

Пример 4. *На 200 полях учтена поражённость озимой пшеницы корневыми гнилями и урожайность. Поля распределены по урожайности на 3 группы и по поражённости на 4.*

Урожай по группам	Поражённость растений по группам								Сумма
	до 10%		11-20%		21-50%		более 30%		
	f	F	f	F	f	F	f	F	
35	25		24		15		16		
45	32		25		10		7		
56	25		14		4		2		
Сумма									

$\chi^2_{05} = 12{,}59$

Пример 5. *В хранилище изучали влияние газовой среды №1 (12% кислорода и 9% углекислого газа) и газовой среды №2 (3% кислорода и 5% углекислого газа) на сохранность плодов груши. Определить, является ли сохранность плодов следствием действия среды или случайных факторов.*

Тип газовой среды	Количество плодов				Сумма
	товарные		нетоварные		
	f	F	f	F	
Газовая среда № 1	142				230
Газовая среда №2	204				260
Сумма					

$\chi^2_{05} = 3{,}84$

Выполнение работы

1. Рассчитать теоретические частоты F и записать в таблицу. Для расчета теоретической частоты F сумму фактических частот данной строки умножаем на сумму фактических частот данного столбца и полученное произведение делим на общую сумму всех наблюдений.

2. Рассчитать значение критерия χ^2: $\chi^2 = \sum \frac{(f-F)^2}{F} =$

3. Провести оценку существенности

χ^2теор = при $v = (c-1)\cdot(k-1)$

Вывод:

Контрольные вопросы к работе:

1. В каких случаях используется критерий χ^2 Пирсона?
2. Оценка существенности при критерию χ^2.

Практическое занятие №10
ОЦЕНКА СУЩЕСТВЕННОСТИ РАЗЛИЧИЙ МЕЖДУ ДВУМЯ НЕЗАВИСИМЫМИ ВЫБОРКАМИ ПО КРИТЕРИЮ ВАН-ДЕР-ВАРДЕНА

Х-критерий ван-дер-Вардена применяется для проверки нулевой гипотезы при сравнении выборок с попарно несвязанными вариантами. При расчете этого критерия обе сравниваемые выборки ранжируют в один общий ряд по возрастающим значениям признака. Затем каждой варианте присваивается порядковый номер (R) того места, которое она занимает в этом ряду. Далее по порядковым номерам ***одной из выборок*** (обычно, но необязательно, меньшей по объему) находят отношения R:(n+1), где R – порядковый номер вариант, их ранг, n – количество всех членов сравниваемых групп, увеличенное на единицу. Определив отношения R:(n+1), находят значение функции этой величины φ{R:(n+1)} в статистической таблице. Суммируя результаты, получают величину критерия Х:

$$X = \sum \varphi\left(\frac{R}{n+1}\right) \quad (31)$$

Фактическое значение критерия Х сравнивается с критическим (или теоретическим), которое берется из статистической таблицы с учетом принятого уровня значимости и $n = n_1 + n_2$ с учетом разницы $n_1 - n_2$. Нулевая гипотеза или предположение о том, что сравниваемые выборки извлечены из генеральных совокупностей с одинаковыми функциями распределения опровергаются, если Хфакт.≥Хтеор.

Пример. Длина тела личинок жуков-щелкунов (мм) в посевах озимой ржи и проса варьировала следующим образом:

в посевах ржи:	7, 10, 14, 15, 12, 16, 12
в посевах проса:	11, 12, 16, 13, 18, 15

Следует определить, существенна ли разность по длине тела личинок жуков щелкунов. Нулевая гипотеза предполагает, что между выборками нет существенного различия, т.е. длина личинки жука-щелкуна не зависит от места обитания.

1. *Определяем объем и среднюю арифметическую 1 выборки:* $n_1=7; x_1=(7+10+14+15+12+16+12):7=86:7=12,3$

2. *Определяем объем и среднюю арифметическую 2 выборки:* $n_2=6$; $x_2=(11+12+16+13+18+15):6=85:6=14,2$

3. Определяем разность между средними: $n_2-n_1=14,2-12,3=1,9$

4. Составляем расчетную таблицу:

Длина личинок щелкуна (мм), обитающих в посевах		Ранг R	$\frac{R}{n+1}$	$\varphi\left(\frac{R}{n+1}\right)$
ржи	проса			
7		1		
10		2		
	11	3	3:14=0,214	-0,79
12		4		
	12	5	5:14=0,357	-0,37
12		6		
	13	7	7:14=0,500	0,00
14		8		
	15	9	9:14=0,643	+0,37
15		10		
	16	11	11:14=0,786	+0,79
16		12		
	18	13	13:14=0,929	+1,47
$n_1=7$	$n_2=6$	-	-	**1,47**

Отношение $\frac{R}{n+1}$ находим по порядковым номерам выборки меньшего объема, где $n= n_1+ n_2=7+6=13$

Значение функции $\varphi\left(\frac{R}{n+1}\right)$ находим в приложении 3, а величину критерия $X=\sum\varphi\left(\frac{R}{n+1}\right)$ определяем, суммируя все значения $\varphi\left(\frac{R}{n+1}\right)$ с учетом знаков.

Теоретическое значение критерия **X** на 5% уровне значимости берем из статистической таблицы (приложение 4). При n=13 с учетом разности n1- n2=1,9 **X=3,71**.

Т.к. фактическое значение критерия меньше теоретического, нулевая гипотеза принимается, следовательно, разница между средними двух выборок статистически недостоверна, или несущественна,

Вывод: ***длина личинок жука-щелкуна не зависит от местообитания***.

Задание. Согласно N примера провести оценку существенности различий между 2 выборками.

Пример 1. *При подсчете числа бутонов сортов яблонь Монтуанское и Голден Делишес, поврежденных яблонным цветоедом, были получены следующие данные:*

Сорт	Количество поврежденных бутонов									
Монтуанское	10	13	18	12	17	7	15	11	14	16
Голден Делишес	5	15	14	11	8	12	13	10	12	-

Пример 2. *При определении количества личинок колорадского жука на растениях картофеля Светлячок и Спринтер были получены следующие данные:*

Сорт	Количество личинок колорадского жука							
Светлячок	10	11	13	16	12	9	21	19
Спринтер	8	14	18	19	17	20	-	-

Выполнение работы

1. *Определяем объем и среднюю арифметическую 1 выборки:*

   n_1=

   x_1=

2. *Определяем объем и среднюю арифметическую 2 выборки:*

   n_2=

   x_2=

3. Определяем разность между средними: n_2-n_1=
4. Составляем расчетную таблицу:

		Ранг R	$\frac{R}{n+1}$	$\varphi\left(\frac{R}{n+1}\right)$
n_1=	n_2=	-	-	**X=**

5. Сравниваем фактическое значение критерия X с теоретическим:

Вывод:

Контрольные вопросы к работе:

1. Какие критерии называются непараметрическими?
2. В каких случаях используется критерий ван-дер-Вардена?
3. Оценка существенности различий 2 выборок по критерию ван-дер-Вардена.

Практическое занятие №11
КОЭФФИЦИЕНТ РАНГОВОЙ КОРРЕЛЯЦИИ СПИРМЕНА

Коэффициент ранговой корреляции Спирмена используется для попарно связанных значений признаков. При ранжировании попарно связанных значений признаков необходимо обращать внимание на то, как они распределяются относительно друг друга. Если возрастающим значениям одного признака соответствуют возрастающие значения другого признака, то между ними существует положительная связь, если же при возрастании значений одного признака значения другого последовательно уменьшаются, между ними существует отрицательная связь. Ранги значений признака определяют, обозначив ранжированные значения признаков порядковыми номерами. Коэффициент ранговой корреляции определяют по формуле

$$r_s = 1 - \frac{6 \cdot \sum d^2}{n \cdot (n^2 - 1)} \quad (32)$$

где d- разность между рангами сопряженных значений признаков X и Y,

n – число парных наблюдений.

При полной связи ранги коррелируемых признаков совпадут и разность между ними будет равна нулю. В таких случаях коэффициент корреляции окажется равным единице. Если же признаки варьируют независимо друг от друга, коэффициент корреляции будет равен нулю. Коэффициент корреляции выражается в долях единицы и может изменяться от –1 до +1.

Достоверность коэффициента ранговой корреляции устанавливается на основании нулевой гипотезы, которая сводится к предположению, что в генеральной совокупности он равен нулю.

Проверяется нулевая гипотеза с помощью t-критерия Стьюдента, с которым сравнивается критерий значимости коэффициента ранговой корреляции. Теоретическое значение критерия Стьюдента берется из статистических таблиц для числа степеней свободы v = n-2 с учетом принятого уровня значимости.

Если $t \geq t_{теор}$, то корреляционная связь значима, если $t < t_{теор}$, то связь незначима.

Пример. Определить зависимость урожайности лука от засоренности полей

Кол-во сорных растений на 1м^2 (X)	15	18	30	39	45	51	35	70	42	53
Урожайность т/га (Y)	22	20	18	17	13	8	18	6	12	8

1. Ранжируем значения признаков по ряду X. Одинаковым значениям признака присваиваем один и тот же ранг.

X, балл	Y, ц/га	Ранги рядов		$d=R_x - R_y$	d^2
		R_x	R_y		
15	22	1	8	7	49
18	20	2	7	5	25
30	18	3	6	3	9
35	18	4	6	2	4
39	17	5	5	0	0
42	12	6	4	2	4
45	13	7	3	4	16
51	8	8	2	6	36
53	8	9	2	7	49
70	6	10	1	9	81
-	-	-	-	-	$\sum d^2 = 273$

2. Определяем коэффициент ранговой корреляции Спирмена:

$$r_s = 1 - \frac{6 \cdot \sum d^2}{n \cdot (n^2 - 1)} = 1 - \frac{6 \cdot 273}{10(100-1)} = 1 - \frac{1638}{990} = 1 - 1{,}65 = -0{,}65$$

Коэффициент ранговой корреляции имеет знак «-», следовательно связь между признаками отрицательная, или обратная.

3. Рассчитываем ошибку коэффициента корреляции:

$$s_{r_s} = \sqrt{\frac{1-r^2}{n-2}} = \sqrt{\frac{1-(-0{,}65)^2}{10-2}} = \sqrt{\frac{1-0{,}42}{8}} = \sqrt{\frac{0{,}58}{8}} = \sqrt{0{,}07} = 0{,}26$$

4. Определяем критерий значимости коэффициента корреляции и сравниваем его с критерием Стьюдента:

$$t_{r_s} = \frac{r_s}{s_{r_s}} = \frac{-0{,}65}{0{,}26} = -2{,}50$$ при v = n-2=10-2=8 t_{05} = 2,31

Критерий значимости по модулю больше критерия Стьюдента, следовательно связь между признаками значимая.

Вывод. ***Установлена обратная значимая зависимость между X и Y, с увеличением количества сорных растений урожайность лука значительно снижается.***

Задание. Согласно N примера провести оценку значимости корреляционной связи между признаками.

Пример 1. *Определить корреляционную зависимость между интенсивностью поражения листовой ржавчиной и урожайностью пшеницы.*

Балл (X)	1	3	2	5	4	1	3	0	4	2
Урожайность ц/га (Y)	44	37	38	32	33	43	35	45	36	40

Пример 2. *Определить влияние семенной инфекции пшеницы на массу зерен.*

% пораженных растений (X)	23	6	60	18	42	16	35	6	18	32
Масса 1000 зерен, г (Y)	34	40	13	35	28	39	24	43	33	29

Пример 3. *Определить корреляционную зависимость между урожайностью огурца и выпуском галлицы против тли*

Норма выпуска галлицы на 1м^2, экз (X)	50	70	80	60	100	90	70	60	100	80
Урожайность кг/м^2 (Y)	17	24	26	19	27	26	23	20	26	25

Выполнение работы

1. Ранжируем значения признаков по ряду X. Одинаковым значениям признака присваиваем один и тот же ранг.

X, балл	Y, ц/га	Ранги рядов		$d=R_x - R_y$	d^2
		R_x	R_y		
-	-	-	-	-	$\sum d^2 =$

2. Определяем отклонения между рангами сопряженных значений признаков X и Y и возводим их в квадрат.
3. Вычисляем коэффициент ранговой корреляции:

$$r_s = 1 - \frac{6 \cdot \sum d^2}{n \cdot (n^2 - 1)}$$

4. Рассчитываем ошибку коэффициента корреляции: $s_{r_s} = \sqrt{\frac{1 - r^2}{n - 2}}$
5. Вычисляем критерий значимости коэффициента корреляции:

$$t_{r_s} = \frac{r_s}{s_{r_s}}$$

t_{05} = при v = n-2

Вывод:

Контрольные вопросы к работе:

1. В каких случаях используется коэффициент ранговой корреляции Спирмена?
2. Оценка значимости корреляционной зависимости между признаками по коэффициенту ранговой корреляции Спирмена.

Практическое занятие №12
КРИТЕРИЙ ЗНАКОВ

Критерий знаков – z, используется для выборок с попарно связанными вариантами. При расчете критерия знаков различия, наблюдаемые при сравнении парных значений, обозначают знаками + (положительный эффект) или – (отрицательный эффект). В основе расчета критерия лежит предположение, что если сравниваемые выборки существенно не различаются, то число плюсовых и минусовых разностей будет одинаковым. Если наблюдается преобладание плюсов или минусов, то это может быть следствием воздействия изучаемого фактора.

Величина критерия знаков определяется большим числом однозначных разностей. При этом нулевые разности, т.е. разности не давшие ни положительного, ни отрицательного результата, в расчет не принимаются и число парных наблюдений уменьшается. Фактическое значение критерия знаков сравнивают с теоретическим, которое берется из статистической таблицы (приложение 5) для принятого уровня значимости и количества парных наблюдений без нулевых значений. Если фактическое значение критерия больше или равно теоретическому, то нулевая гипотеза отвергается.

Пример. На 12 яблонях выборочно подсчитывалось число бутонов, поврежденных яблоневым цветоедом. Учет повреждений проводился одновременно на южной и западной стороне каждой яблони. Результаты представлены в таблице:

Сторона дерева		Разница	Сторона дерева		Разница
южная	западная		южная	западная	
10	5	+	15	13	+
13	15	-	14	12	+
28	22	+	16	7	+
12	11	+	9	16	-
17	8	+	8	11	-
7	11	-	15	10	+

В 8 случаях из 12 больше бутонов было повреждено цветоедом на южной стороне и в 4 случаях – на западной. Следовательно, zфакт = 12 – 4 = 8. При n =12 на 5% уровне значимости zst = 10, на 1% - zst = 11. zфакт<zst, следовательно нулевая гипотеза принимается.

Вывод: ***яблоневый цветоед одинаково повреждает бутоны на западной и южной сторонах яблони***.

Задание. Провести оценку существенности различий между 2 выборками.

Пример. *В опыте по изучению влияния облучения гамма-лучами семян гороха на устойчивость растений к корневым гнилям поражение проростков гороха возбудителями корневых гнилей было следующим (%):*

семена необлученные 24; 28; 20; 28; 24; 29; 21; 25; 31; 27; 30; 27; 26; 32

семена облученные 22; 23; 20; 23; 18; 20; 21; 19; 24; 19; 17; 24; 21; 25

Определить влияет ли облучение семян гороха гамма-лучами на поражение проростков корневыми гнилями.

Контрольные вопросы к работе:

1. В каких случаях используется критерий знаков?
2. Оценка существенности различий 2 выборок по критерию знаков.

Практическое занятие № 13
ПРОБИТ-АНАЛИЗ

Пробит-анализ используется при изучении силы действия повреждающих факторов, чаще всего химических средств защиты растений, на биологические объекты: сорные растения, вредители, патогенные микроорганизмы.

На оси абсцисс откладывают логарифмы доз изучаемого препарата, по оси ординат – процентное выражение гибели подопытных объектов, преобразованное в условные вероятностные величины – ***пробиты***. Значение пробит находят в статистической таблице (приложение 6).

Пример. Изучали эффективность инсектицида в различных дозах на мухах. В 45 чашек Петри поместили по 20-25 мух, затем по 9 чашек обработали различными дозами инсектицида. По данным гибели в 9 чашках вычисляли среднюю гибель мух для каждой дозы. Были получены следующие данные:

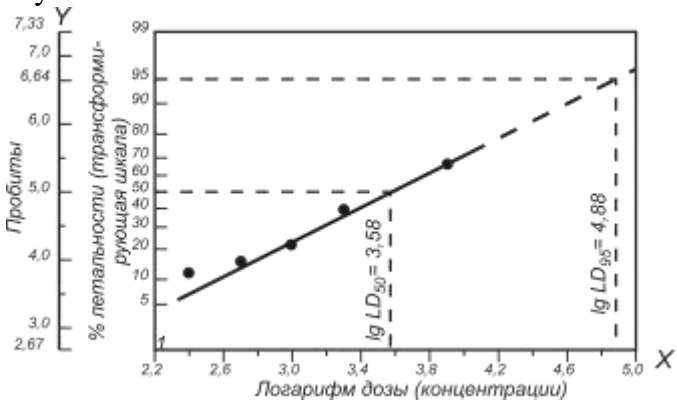
Доза (концентрация, мг/л)	250	500	1000	2000	8000
Средняя гибель, %	13	15	22	39	66

Следует определить дозу, соответствующую 50% гибели мух.

По таблице логарифмов преобразуем значения доз, а по приложению 6 находим значение пробит.

Доза (концентрация, мг/л)	Средняя гибель, %	Преобразованные значения	
		lg дозы (осьX)	Значение пробит (ось Y)
250	13	2,398	3,87
500	15	2,699	3,96
1000	22	3,000	4,23
2000	39	3,301	4,72
8000	66	3,903	5,41

Откладываем преобразованные значения доз инсектицида и гибели мух



Прямую линию проводим таким способом, чтобы она располагалась как можно ближе ко всем точкам.

Определяем дозу, соответствующую 50% летальности. Для 50% гибели пробит равен 5, ему соответствует lg = 3,58.

По таблице антилогарифмов находим дозу LD_{50}, которая равна 3800 мг/л.

Задание. Согласно N примера рассчитать $LD_{50.}$

Пример 1. *При обработке гусениц капустной белянки волатоном были получены следующие результаты*

Доза волатона, мг/10л	50	75	100	125	150
Гибель, %	27,1	42,4	67,5	80,4	92,1

Пример 2. *При обработке личинок колорадского жука банколом были получены следующие результаты:*

Доза банкола, мг/10 л	100	200	400	600	800
Гибель, %	12,3	34,8	45,2	79,4	90,7

Выполнение работы

1. По таблице логарифмов (приложение 5) преобразуем значения доз, а по приложению 6 находим значение пробит.

Доза препарата	Гибель, %	Преобразованные значения	
		lg дозы	Значение пробит

2. Откладываем на графике преобразованные значения доз препарата и гибели вредителей.
3. Определяем lg дозы, соответствующей пробиту равному 5.
4. По таблице антилогарифмов (приложение 3) находим дозу LD_{50}.

Вывод:

Контрольные вопросы к работе:

1. В каких случаях используется пробит-анализ?
2. Сущность пробит-анализа.

Приложение 1

Значения критерия τ на 1 и 5% уровнях значимости

n	τ	
	01	05
4	0,991	0,955
5	0,916	0,807
6	0,805	0,689
7	0,740	0,610
8	0,683	0,554
9	0,635	0,512
10	0,597	0,477
11	0,566	0,450
12	0,541	0,428
14	0,502	0,395
16	0,472	0,364
18	0,449	0,349
20	0,430	0,334
22	0,414	0,320

Приложение 2

Значение критерия t на 5 и 1% уровне значимости

Число степеней свободы	Уровень значимости	
	05	01
3	3.18	5.84
4	2.78	4.60
5	2.57	4.03
6	2.45	3.71
7	2.37	3.50
8	2.31	3.36
9	2.26	3.25
10	2.23	3.17
11	2.20	3.11
12	2.18	3.06
13	2.16	3.01
14	2.15	2.98
15	2.13	2.95
16	2.12	2.92
17	2.11	2.90
18	2.10	2.88
19	2.09	2.86
21	2.08	2.83
22	2.07	2.82
23	2.07	2.81
24	2.06	2.80
25	2.06	2.79
26	2.06	2.78
27	2.05	2.77
28	2.05	2.76
29	2.05	2.76
30	2.04	2.75
50	2.01	2.68
100	1.98	2.63
более 100	1.96	2.58

Приложение 3

Значение функции $\varphi\left(\frac{R}{n+1}\right)$ для расчета критерия ван-дер-Вардена (по И.П. Ашмарину и А.А. Воробьеву)

R \ n	8	9	10	11	12	13	14	15	16
1	-1,22	-1,28	-1,33	-1,39	-1,43	-1,47	-1,50	-1,54	-1,56
2	-0,77	-0,84	-0,91	-0,97	-1,02	-1,07	-1,11	-1,15	-1,19
3	-0,43	-0,52	-0,60	-0,67	-0,74	-0,79	-0,84	-0,89	-0,93
4	-0,14	-0,25	-0,35	-0,43	-0,50	-0,57	-0,62	-0,67	-0,72
5	0,14	0	-0,12	-0,21	-0,29	-0,37	-0,43	-0,49	-0,54
6	0,43	0,25	0,12	0	-0,10	-0,18	-0,25	-0,32	-0,38
7	0,77	0,52	0,35	0,21	0,10	0	-0,09	-0,16	-0,22
8	1,22	0,84	0,60	0,43	0,29	0,18	0,09	0	-0,07
9		1,28	0,91	0,67	0,50	0,37	0,25	0,16	0,07
10			1,33	0,97	0,74	0,57	0,43	0,32	0,22
11				1,39	1,02	0,79	0,62	0,49	0,38
12					1,43	1,07	0,84	0,67	0,54
13						1,47	1,11	0,89	0,72
14							1,50	1,15	0,93
15								1,54	1,19
16									1,56

R \ n	17	18	19	20	21	22	23	24	25
1	-1,59	-1,62	-1,64	-1,66	-1,70	-1,72	-1,73	-1,75	-1,77
2	-1,22	-1,25	-1,28	-1,31	-1,33	-1,36	-1,39	-1,41	-1,43
3	-0,97	-1,00	-1,04	-1,07	-1,10	-1,13	-1,15	-1,18	-1,20
4	-0,77	-0,81	-0,84	-0,88	-0,91	-0,94	-0,97	-0,99	-1,02
5	-0,59	-0,63	-0,67	-0,71	-0,75	-0,78	-0,81	-0,84	-0,87
6	-0,43	-0,48	-0,52	-0,57	-0,60	-0,64	-0,67	-0,71	-0,74
7	-0,28	-0,33	-0,39	-0,43	-0,47	-0,51	-0,55	-0,58	-0,62
8	-0,14	-0,20	-0,25	-0,30	-0,35	-0,39	-0,48	-0,47	-0,50
9	0	-0,07	-0,13	-0,18	-0,28	-0,28	-0,32	-0,36	-0,40
10	0,14	0,07	0	-0,06	-0,11	-0,16	-0,21	-0,25	-0,29
11	0,28	0,20	0,13	0,06	0	-0,06	-0,10	-0,15	-0,19
12	0,43	0,33	0,25	0,18	0,11	0,06	0	-0,05	-0,10
13	0,59	0,48	0,39	0,30	0,28	0,16	0,10	0,05	0
14	0,77	0,63	0,52	0,43	0,35	0,28	0,21	0,15	0,10
15	0,97	0,81	0,67	0,57	0,47	0,39	0,32	0,25	0,19
16	1,22	1,00	0,84	0,71	0,60	0,51	0,48	0,36	0,29
17	1,59	1,25	1,04	0,88	0,75	0,64	0,55	0,47	0,40
18		1,62	1,28	1,07	0,91	0,78	0,67	0,58	0,50
19			1,64	1,31	1,10	0,94	0,81	0,71	0,62
20				1,66	1,33	1,13	0,97	0,84	0,74
22					1,70	1,36	1,15	0,99	0,87
23						1,72	1,39	1,18	1,02
24							1,73	1,41	1,20
25								1,75	1,43
26									1,77

Приложение 4

Теоретические значения Х-критерия ван-дер-Вардена
на 1 и 5% уровнях значимости

N	n_x-n_y=0 или 1		n_x-n_y=2 или 3	
	05	01	05	01
8	2,40		2,30	
9	2,38		2,20	
10	2,60	3,20	2,49	3,10
11	2,72	3,40	2,58	3,40
12	2,86	3,60	2,79	3,58
13	2,96	3,71	2,91	3,64
14	3,11	3,94	3,06	3,98
15	3,24	4,07	3,19	4,05
16	3,39	4,26	3,36	4,25
17	3,49	4,44	3,44	4,37
18	3,63	4,60	3,60	4,58
19	3,73	4,77	3.69	4,71
20	3,86	4,94	3,84	4,92
21	3,96	5,10	3,92	5,05
22	4,08	5,26	4,06	5,24
23	4,18	5,40	4,15	5,36
24	4,28	5,55	4,27	5,53
25	4,39	5,68	4,36	5,65
26	4,50	5,83	4,48	5,81

Приложение 5

Теоретические значения критерия знаков z
на 1 и 5% уровнях значимости

n	05	01	**n**	05	01	**n**	05	01
5	0		**26**	8	6	**47**	17	15
6	0	-	**27**	8	7	**48**	17	15
7	0	0	**28**	8	7	**49**	18	15
8	1	0	**29**	9	7	**50**	18	16
9	1	0	**30**	10	8	**52**	19	17
10	1	0	**31**	10	8	**54**	20	18
11	2	1	**32**	10	8	**56**	21	18
12	2	1	**33**	11	9	**58**	22	19
13	3	1	**34**	11	9	**60**	23	20
14	3	2	**35**	12	10	**62**	24	21
15	3	2	**36**	12	10	**64**	24	22
16	4	2	**37**	13	10	**66**	25	23
17	4	3	**38**	13	11	**68**	26	23
18	5	3	**39**	13	11	**70**	27	24
19	5	4	**40**	14	12	**72**	28	25
20	5	4	**41**	14	12	**74**	29	26
21	6	4	**42**	15	13	**76**	30	27
22	6	5	**43**	15	13	**78**	31	28
23	7	5	**44**	16	13	**80**	32	29
24	7	5	**45**	16	14	**82**	33	30
25	7	6	**46**	16	14	**84**	33	30

Приложение 6

Значения пробитов для экспериментально устанавливаемой гибели вредных организмов от 0 до 99 %

Гибель, %	0	1	2	3	4	5	6	7	8	9
0	-	2,67	2,95	3,12	3,25	3,35	3,45	3,52	3,59	3,66
10	3,72	3,77	3,82	3,83	3,92	3,96	4,01	4,05	4,08	4,12
20	4,16	4,19	4,23	4,26	4,29	4,33	4,36	4,39	4,42	4,45
30	4,48	4,50	4,53	4,56	4,59	4,61	4,64	4,67	4,69	4,72
40	4,75	4,77	4,80	4,82	4,85	4,87	4,90	4,92	4,95	4,97
50	5,00	5,03	5,05	5,08	5,10	5,13	5,15	5,18	5,20	5,23
60	5,25	5,28	5,31	5,33	5,36	5,39	5,40	5,44	5,47	5,50
70	5,52	5,55	5,58	5,61	5,64	5,67	5,71	5,74	5,77	5,81
80	5,84	5,88	5,92	5,95	5,99	6,04	6,08	6,13	6,18	6,23
90	6,28	6,34	6,41	6,48	6,55	6,64	6,75	6,88	7,05	7,33

Список использованной литературы

1. Доспехов Б.А. Методика полевого опыта (с основами статистической обработки результатов исследования) - М.: Агропромиздат, 1985. – 351 с.
2. Ивантер Э.В., Коросов А.В. Элементарная биометрия. Учебное пособие - Петрозаводск: Изд-во ПГУ, 2005. – 104 с.
3. Каюмов А.Р., Маркелов О.А. Математические методы в биологии и генетике. Часть 1. Основы теории вероятностей. Учебно-методическое пособие – Казань: Казань, КФУ, 2016. - 46 с.
4. Лакин Г.Ф. Биометрия - М.: Высшая школа, 1990. - 352с.
5. Пересыпкин В.Ф., Коваленко С.И. и др. Практикум по методике полевого опытного дела в защите растений – М.: Агропромиздат, 1989. – 174 с.
6. Рокицкий П.Ф. Биологическая статистика – Минск: Высшая школа, 1973. - 327с.

Учебное издание

ОСНОВЫ БИОЛОГИЧЕСКОЙ СТАТИСТИКИ

Методические указания к практическим занятиям

Трескина Наталья Новомировна

Издается в авторской редакции

Подписано в печать ___2019 г. Формат 60 х 90/16. Усл. печ. л. 4,0. Тираж 20 экз.