

**Казанский (Приволжский) федеральный университет
Биолого-почвенный факультет
Кафедра почвоведения**

**Статистическая обработка результатов полевых
агрохимических исследований
с помощью пакета
STATGRAPHICS PLUS for Windows**

**Учебно-методическое пособие
для студентов биолого-почвенного факультета**

Казань
2012

УДК.631.5

*Печатается по решению Редакционно-издательского совета ФГАОУВПО
«Казанский (Приволжский) федеральный университет»*

методической комиссии биолого-почвенного факультета

Протокол № 1 от 26 января 2012 г.

заседания кафедры почвоведения

Протокол № 11 от 22 марта 2012 г.

Рецензенты:

Д.б.н., профессор Копосов Г.Ф.

Матвеева Н.М., Валеева А.А.

Название: Статистическая обработка результатов полевых агрохимических исследований с помощью пакета STATGRAPHICS PLUS for Windows/
Н.М.Матвеева, А.А. Валеева. – Казань: Казанский университет, 2012. - 63с

Методическое пособие разработано для проведения практических занятий со студентами 4-го курса, обучающихся по специальности «Почвоведение», к дисциплине «Методы полевых агрохимических исследований». В пособие рассматривается решение задач статистической обработки данных полевого опыта, так как математическая интерпретация результатов – неременное условие успешного получения ответов на вопросы, интересующие экспериментатора.

Знание современных методов статистической обработки необходимо не только для количественной характеристики наблюдений и полученных в опыте данных, когда уже нельзя ничего исправить, но и на всех этапах эксперимента – от планирования до интерпретации окончательных результатов.

© Казанский университет, 2012

Оглавление	страница
Введение.....	4
1. Описательная статистика количественной изменчивости.....	5
1.1. Общие сведения о работа с Базовой системой STATGRAPHICS.....	7
1.2. Общие статистики (Summary statistics).....	8
2. Проверка данных на нормальное распределение.....	12
3. Параметрические методы статистического анализа	
3.1 Корреляционный анализ.....	17
3.2. Регрессионный анализ.....	21
3.3. Дисперсионный анализ.....	24
3.3.1. Дисперсионный анализ однофакторного опыта.....	25
3.3.2. Дисперсионный анализ многофакторного опыта.....	31
Приложение.....	37
Сборник задач.....	40
Литература.....	63

Введение

Все статистические методы основаны на теории вероятностей – науке, изучающей общие закономерности в случайных явлениях различной природы, и применяются там, где приходится иметь дело с планированием экспериментов и обследований, с оценкой параметров и проверкой гипотез.

Знание современных методов статистической обработки необходимо не только для количественной характеристики наблюдений и полученных в опыте данных, но и на всех этапах эксперимента – от планирования до интерпретации окончательных результатов.

При планировании и проведение эксперимента существенную роль играют статистические методы, которые дают возможность:

- компактно и информативно описать результаты экспериментов;
- устанавливать степень достоверности сходства и различия исследуемых объектов на основании результатов измерений их показателей;
- анализировать наличие или отсутствие зависимости между различными показателями;
- выявлять информативность показателей.

Свойство условных единиц – растений, урожаев на параллельных делянках полевого опыта – отличаются, друг от друга даже в однородных совокупностях называется изменчивостью, или варьированием. Изменчивость – свойство, присущее всем предметам природы. Так варьирующим признаком у растений является высота, количество и масса зерен в колосе, содержание протеина и т. д. Варьирование возникает вследствие того, что растения одного и того же сорта всегда отличаются своей наследственностью. Кроме того, формирование их часто протекает в относительно различных условиях внешней среды. Следовательно, при любом исследовании данные опыта будут всегда варьировать в тех или иных пределах.

Изменчивость, варьирование признаков создает трудность в тех случаях, когда требуется дать общую характеристику определенной варьирующей группе растений, почв и т.д. по отдельным признакам или сравнить две такие группы и найти различия между ними. Совершенно очевидно, что не всегда возможно исследовать по тому или другому признаку все особенности, всю совокупность. В этих случаях прибегают к изучению части ее, по которой делают общие заключения. Такой метод называют выборочным, и считается основным при статистическом изучении совокупности.

Таким образом, всю группу объектов, подлежащих изучению, называют совокупностью или генеральной совокупностью, а ту часть объектов, которая попала на проверку, исследование, – выборочной совокупностью или выборкой.

Главная цель выборочного метода – по статистическим показателям малой выборки (средней пробы) возможно точнее охарактеризовать всю совокупность объектов, которая в статистике и называется генеральной совокупностью.

1. Описательная статистика количественной изменчивости

К количественным относят признаки, которые могут быть охарактеризованы количественно – урожай с делянки, число, высота и масса растений, содержание белка и клейковины в зерне и т. д. Различают два вида количественной изменчивости: непрерывную и прерывистую или дискретную. В первом случае значения признака выражены мерами объема, длины, массы и т.д., во втором различия между единицами наблюдения выражаются целыми числами, между которыми нет и не может быть переходов, например число зерен в колосе и т.д.

Основными статистическими характеристиками количественной изменчивости являются средняя арифметическая (\bar{x}), дисперсия (s^2), стандартное отклонение (s), ошибка средней арифметической ($s_{\bar{x}}$), коэффициент вариации (V) и относительная ошибка выборочной средней ($s_{\bar{x}}\%$). Суть их состоит в упорядочивании, уплотнении и графическом представлении результатов наблюдений с тем, чтобы, с одной стороны, сделать их более наглядными, с другой – создать предпосылки для применения к ним методов статистики.

Средняя арифметическая \bar{x} представляет собой обобщенную, абстрактную характеристику всей совокупности в целом.

$$\bar{x} = \sum X / n$$

где X – значение признака; n – общее число измеренных значений

Дисперсия s^2 есть среднее квадратов отклонений от среднего.

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

Чем больше дисперсия или стандартное отклонение, тем более рассеяны около средней индивидуальные значения признака, т. е. больше изменчивость. Дисперсия постоянной величины равна нулю. Недостаток дисперсии заключается в том, что ее размерность соответствует квадратным единицам изменения. Например, если данные измерены в т/га, то дисперсия имеет размерность ($\text{т}^2/\text{га}^2$). Чтобы вернуться к исходной шкале измерений, необходимо произвести операцию, обратную возведению в квадрат, т. е.

извлечь квадратный корень. Результатом является *стандартное отклонение* S .

$$S_x = \sqrt{S_x^2}$$

Коэффициент вариации V - стандартное отклонение, выраженное в процентах к средней арифметической данной совокупности. Коэффициент вариации можно рассчитывать только тогда, когда все измеряемые величины неотрицательны. Коэффициент вариации является относительным показателем изменчивости. Изменчивость принято считать незначительной, если коэффициент вариации не превышает 10%. Если V выше 10%, но менее 20% - средняя изменчивость. Если V более 20% - значительная изменчивость.

$$V = \frac{S}{x} 100$$

Ошибка выборочной средней или ошибка среднего $S_{\bar{x}}$ - является мерой отклонения выборочной средней \bar{x} от средней всей (генеральной) совокупности μ . Ошибки выборки возникают вследствие неполной репрезентативности (представительности) выборочной совокупности и свойственны только выборочному методу исследования. Они связаны с перенесением результатов, полученных при изучении выборки, на всю генеральную совокупность.

$$S_{\bar{x}} = \frac{S}{\sqrt{n}}$$

Относительная ошибка выборочной средней $S_{\bar{x}}\%$ - ошибка выборки, выраженная в процентах от соответствующей средней.

$$S_{\bar{x}}\% = \frac{S_{\bar{x}}}{x} 100$$

Вопросы к занятию:

1. Какие данные являются количественными:

- оценки студентов (1, 2, 3, 4, 5);
- урожай озимой пшеницы (ц/га);
- порода собак;
- поражение пшеницы мучнистой росой (отсутствие, незначительное, среднее сильное);
- количество сорняков на m^2 ;
- молочная продуктивность (литры за лактацию);
- сумма активных температур;
- содержание белка в сухом веществе (%).

2. Какие данные являются непрерывными, а какие дискретными:

- рост человека;

- масса хлеба;
- число детей в семье;
- урожай пшеницы;
- число абитуриентов поступивших в университет;
- число пораженных зерен в колосе;
- длина стебля кукурузы;
- объем вытяжки;
- число студентов, закрывшие сессию.

3. В вегетационном опыте из трех делянок получили следующие урожай сахарной свеклы (ц/га): $x_1 = 360$, $x_2 = 216$, $x_3 = 369$. Какой средний урожай мы получили?

4. В полевом опыте из 5 делянок рассчитаны коэффициенты вариации урожая озимой пшеницы. В первой делянке коэффициент вариации 21,5%, во второй 12,4%, в третьем – 22,5%, в четвертом – 9,5% и в пятом – 18,1%. В какой делянке наблюдается незначительная, а в какой значительная изменчивость?

5. Рассчитайте дисперсию, если число зерен в двадцати колосках будет 12?

1.1. Общие сведения о работе с Базовой системой StatGraphics

Вид экрана представлен на рис.1.

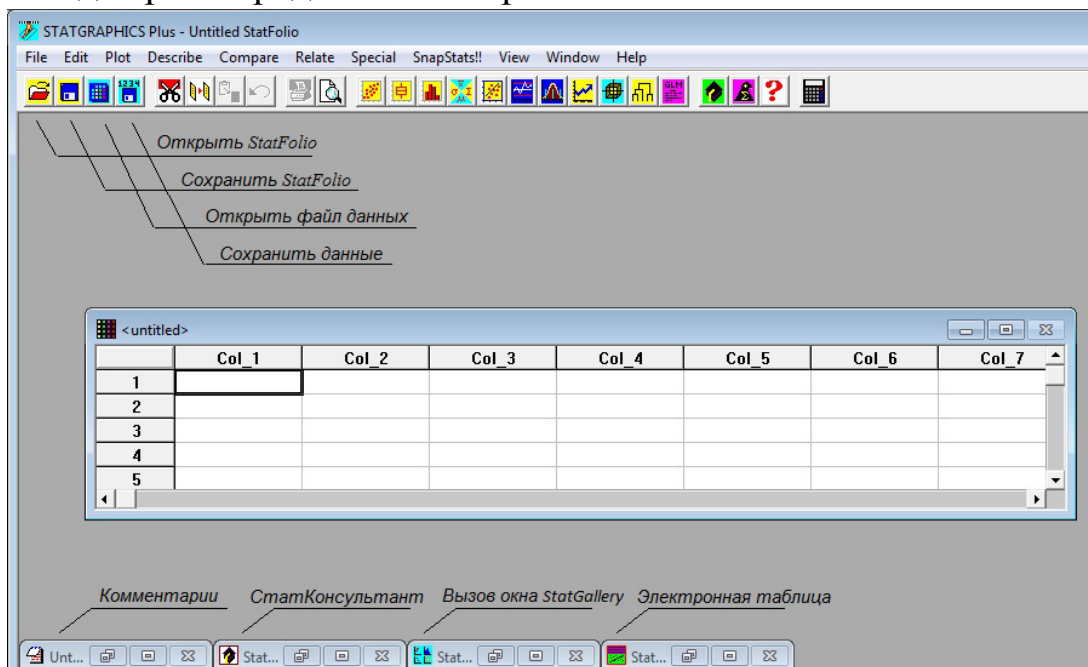


Рис. 1. Вид экрана после инициализации STATGRAPHICS Plus

Набор кнопок в верхней части окна предназначен для открытия готовых статистических проектов StatFolio и записи новых проектов, для открытия файлов данных и их сохранения, для вывода результатов статистического анализа на печать, а также для вызова некоторых статистических и графических процедур. Эти же операции можно осуществить, войдя в меню File, Edit, Plot, Describe, Compare, Relate и Special. Внизу экрана расположен набор пиктограмм, связанных со следующими операциями:

1. Работа с электронной таблицей;
2. Получение консультации у статистической экспертной системы StatAdvisor (СтатКонсультант);
3. Вызов окна StatGallery;
4. Ввод комментариев к проводимому статистическому анализу.

СтатКонсультант представляет собой интерпретацию результатов, определяет значимые эффекты и выявляет возможные изъяны в проведенном анализе.

1.2. Общие статистики (Summary statistics)

Пример 1. Используя описательные статистики проанализируйте количественную изменчивость показателя содержания белка в пшенице. Дана выборка $n=25$ и содержания белка в пшенице, %:

15.5, 16.0, 13.2, 15.2, 17.1, 13.2, 15.4, 16.1, 17.1, 14.2, 16.2, 14.5, 16.2, 12.1, 14.2, 17.1, 13.2, 15.4, 16.1, 17.1, 13.2, 15.4, 16.1, 13.5, 14.2.

- Для проведения статистического анализа создайте файлы данных.

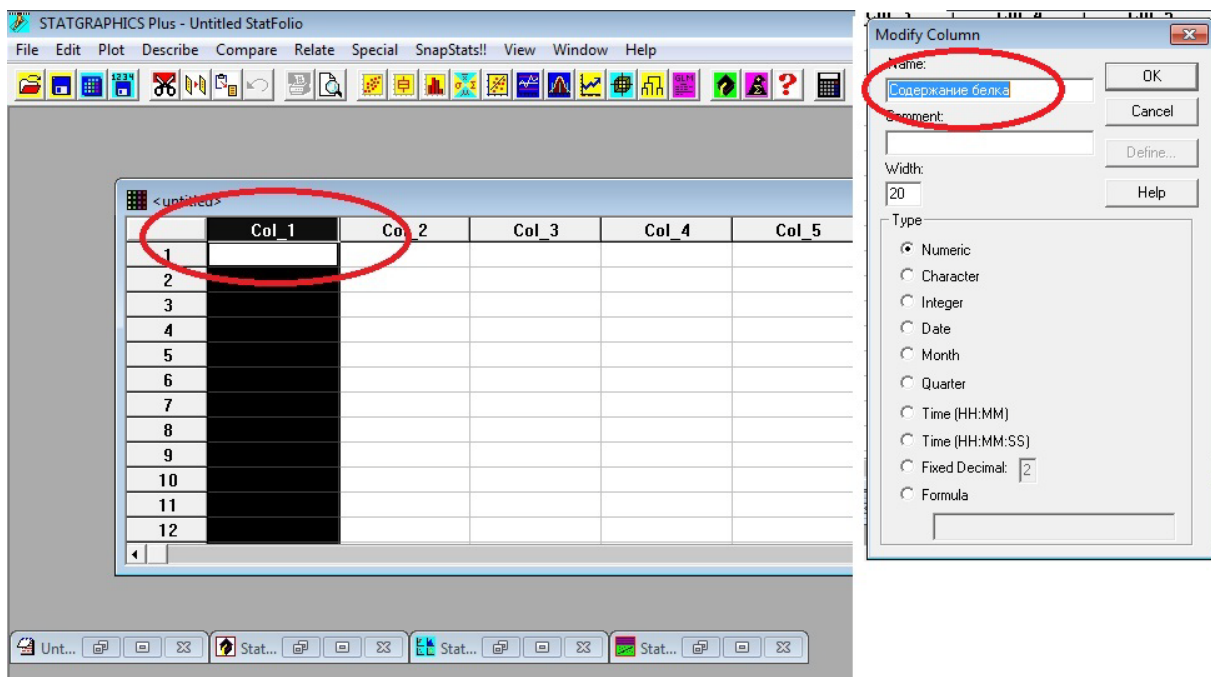


Рис. 2. Электронная таблица

- Создайте новую электронную таблицу.
- Введите имена и типы переменных. Двойным нажатием левой кнопки мыши на значок Col_1 открывается «Окно модификации столбца» (Modify Column), куда вводим название количественной характеристики и нажимаем ОК (рис. 2). Заполняем электронную таблицу.
- Сохраните файл данных. Выберите команду из меню File/Save Data File As, введите имя файла (1) и нажмите кнопку ОК. В заголовке таблицы появится введенное имя, которое будет указано на пиктограмме файла данных.
- Выберите из меню Describe (описание)/Numeric Data (числовые данные) (рис. 3, шаг 1) /One Variable Analysis (анализ одной переменной) (рис. 3, шаг 2). В появившемся окне задайте переменную – Содержание белка (рис. 3, шаг 3) и нажмите ОК.

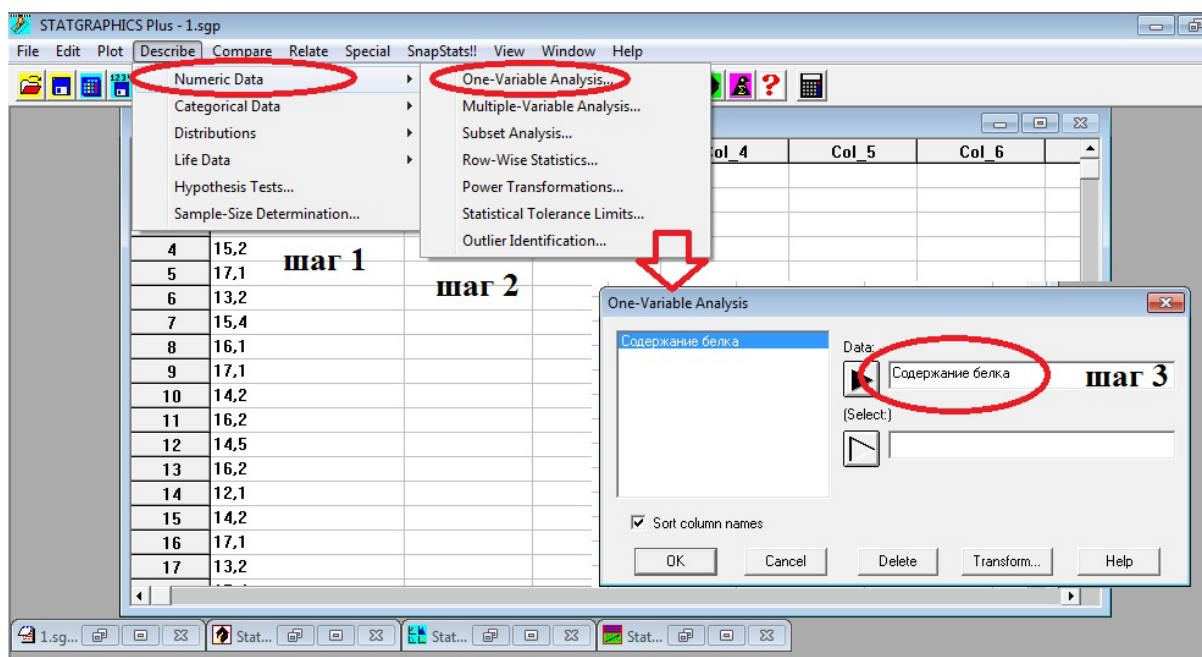


Рис. 3. Выбор данных

- На экране появится рабочее поле анализа одной переменной со сводкой – Analysis Summary . В верхней части рабочего окна расположена панель инструментов, с помощью которой можно изменять входные данные, выбирать табличные и графические опции и сохранять результаты анализа в файле данных (рис. 4).

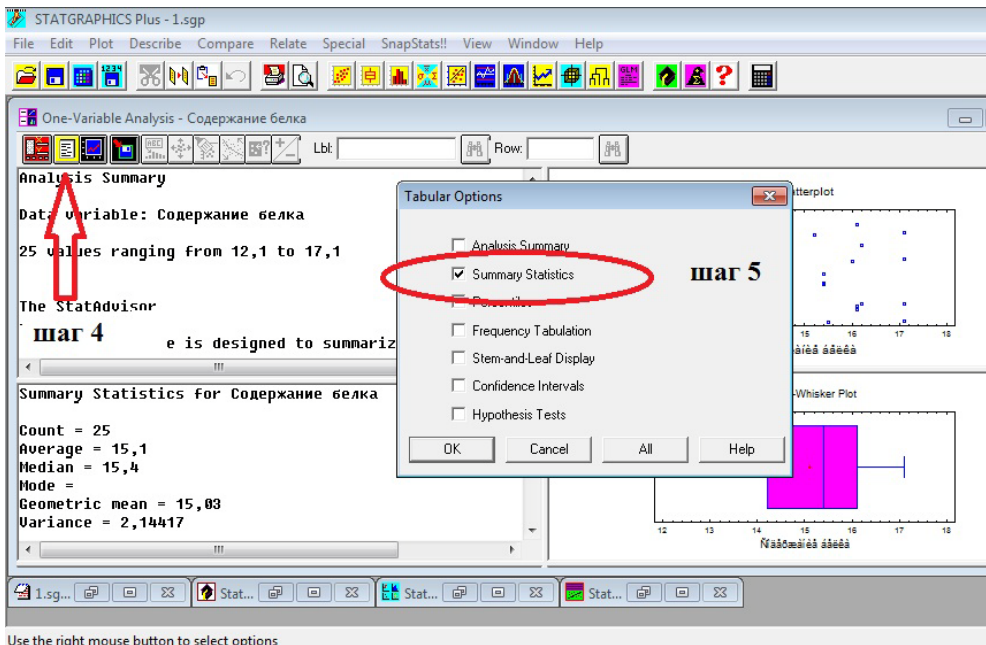


Рис. 4. Рабочее поле анализа

- На панели инструментов нажмите кнопку Tabular options (Табличные опции) (рис. 4, шаг 4). Установите флажок на Summary Statistics (рис. 4, шаг 5) и нажмите ОК.
- Если раскрыть окно общих статистик и нажать правую клавишу мыши, то на экране возникает окно диалога, в котором можно задать необходимые изменения в наборе выдаваемых статистик (рис. 5, шаг 6)

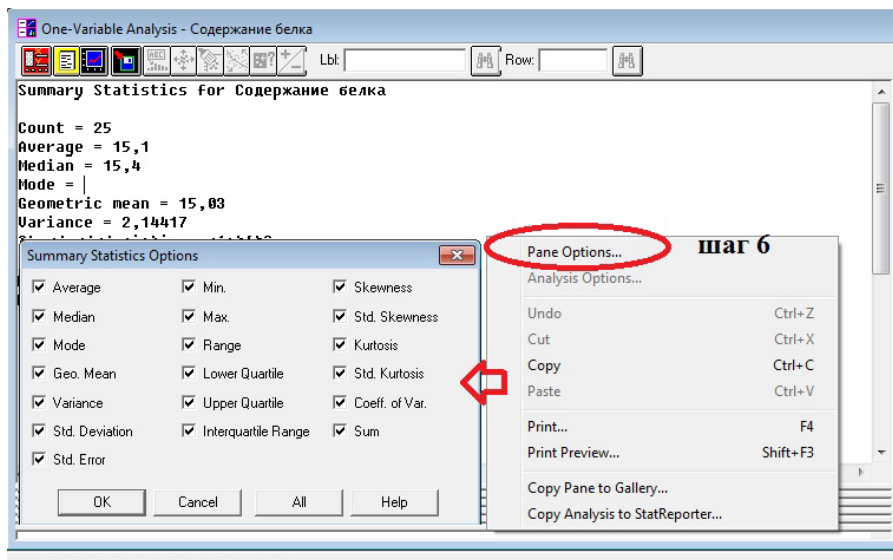


Рис. 5. Изменение выдаваемых статистик

Показатели общей статистики:
Average (среднее арифметическое);

Median (медиана);
 Mode (мода);
 Geo. Mean (геометрическое среднее);
 Variance (дисперсия);
 Std. Deviation (стандартное отклонение);
 Std. Error (ошибка среднего);
 Min. (минимум);
 Max. (максимум);
 Range (размах);
 Lower Quartile (нижний квартиль);
 Upper quartile (верхний квартиль);
 Interquartile Range (межквартильный размах);
 Skewness (коэффициент асимметрии);
 Std. Skewness (нормированный коэффициент асимметрии);
 Kurtosis (коэффициент эксцесса);
 Std. Kurtosis (нормированный коэффициент эксцесса).

Не приводится коэффициент вариации, но его можно рассчитать по формуле:

$$V = \frac{S}{\bar{x}} 100$$

где s - стандартное отклонение; \bar{x} - среднее арифметическое.

Относительную ошибку выборочной средней $S_x\%$ рассчитывают по формуле:

$$S_x\% = \frac{S_x}{\bar{x}} 100$$

где S_x - ошибка среднего, \bar{x} - среднее арифметическое.

Вопросы к занятию:

1. Какие признаки называются количественными?
2. Чем отличаются прерывистые от дискретных количественных изменений.
3. Какие выборки называются малым, а какие большими?
4. В чем отличия генеральной совокупности от выборки?
5. Чем отличается дисперсия от стандартного отклонения?
6. Назовите основные статистические характеристики количественной изменчивости.
7. Какие недостатки имеет дисперсия?

8. В каких случаях нельзя рассчитать коэффициент вариации?
9. В каких случаях возникает ошибка среднего?
10. При каком значении коэффициента вариации наблюдается незначительная изменчивость признака?

2. Проверка данных на нормальное распределение.

Различают эмпирические и теоретические распределения частот совокупности результатов наблюдений.

Эмпирическое распределение – распределение результатов измерений, полученных при изучении выборки, например распределение растений по высоте и массе, распределение делянок дробного учета по урожаю и т.д. В основе его лежат определенные математические закономерности, которые в генеральной совокупности, т. е. при очень большом числе наблюдений, характеризуются некоторыми теоретическими распределениями.

Необходимо помнить, что для конкретных статистических критериев имеются ограничения. Так, например, нельзя использовать коэффициент корреляции Пирсона для признаков, не имеющих нормального распределения.

Поскольку понятие нормальности играет важную роль для многих статистических процедур, важно иметь возможность проверить, что выборка значимо не отличается от нормальной, а используемые статистики имеют выборочные нормальные распределения.

Значение α -ошибки, называемой уровнем значимости (или р-значение). Оно зависит от желаемой доверительной вероятности. Для доверительной вероятности 95%, например $\alpha = 1 - 0,95 = 0,05$. В практике агрономических исследований считается возможным пользоваться вероятностями 0,95-95% и 0,99-99%, к которым соответствуют 0,05-5%-ный и 0,01-1%-ный уровни значимости. Эти вероятности получили название доверительных вероятностей, т.е. таких значений, которым можно доверять и уверенно пользоваться ими.

Пример 2. Проверить на нормальность данные содержания белка в пшенице.

- Откройте файл данных «1» (сохраненный на предыдущем занятии).
- Для первичной проверки постройте график выборочной функции для переменной «Содержание белка». Выберите меню Describe (рис. 6, шаг 1)/ Distribution Fitting (подбор распределения) (рис. 6, шаг 2) для вызова окна

выбора данных. Введите показатель в окно Date (рис. 6, шаг 3) и нажмите ОК. На экране появляется окно анализа Distribution Fitting с Analysis Summary (описательная функция с общим анализом).

- Нажмите кнопку <Графические опции> на панели инструментов. Установите флажок на Symmetry Plot и Frequency Histogram и нажмите ОК (рис. 7, шаг 4). Если точки графика выборочной функции распределения лежат в разумной близости от прямой линии, то это показывает приблизительную нормальность распределения рассматриваемой совокупности.

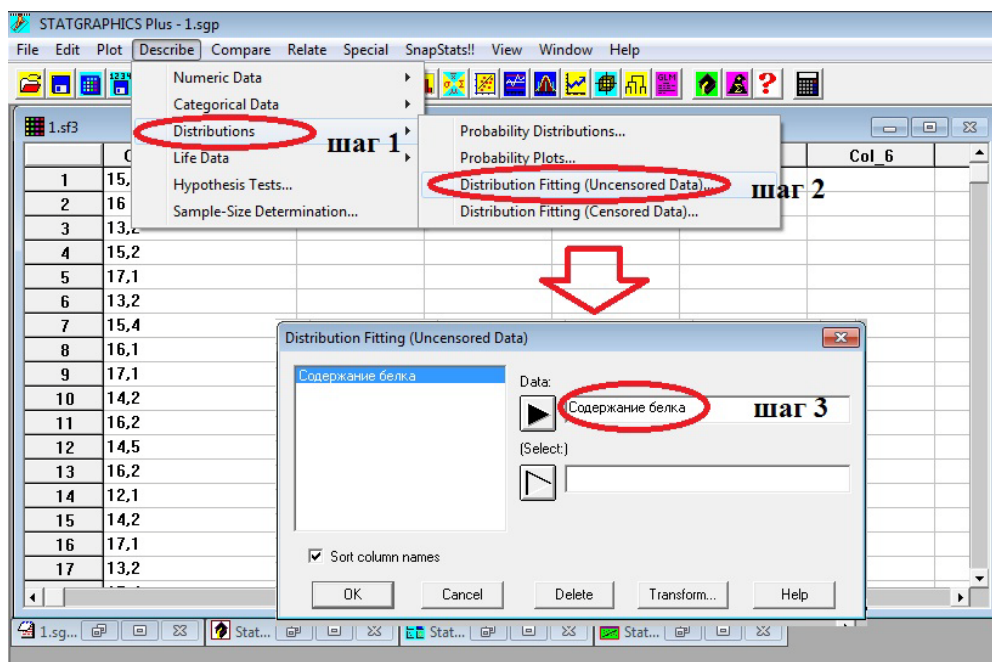


Рис. 6. Проверка данных на нормальность распределения

- Для объективной проверки допущения нормальности могут применяться более тонкие методы: с использованием третьих и четвертых моментов, критерия χ^2 или критерия равенства дисперсии. В STATGRAPHICS реализованы следующие процедуры проверки нормальности распределения (Tests for Normality):
 - скорректированный χ^2 ;
 - тест Шапиро-Уилка;
 - тесты для малых выборок (с помощью асимметрии и эксцесса);
 - тесты согласия (Goodness-of-Fit Tests).

Большинство критериев согласия, применяемых на практике, основаны на одной общей идее, принадлежащей К. Пирсону. Критерий согласия

позволяет выяснить, насколько согласуются между собой наблюдаемые частоты и ожидаемые, иными словами существенны или нет различия между ними. Наблюдаемые частоты – частоты, полученные по выборке. Ожидаемые частоты – частоты, полученные путем вычисления на основе теоретических представлений о нормальном распределении.

Критерий Шапиро-Уилка (W) применим при $8 \leq n \leq 50$, где n – объем выборки. Малые выборки – выборки менее 30 единиц. Большие выборки – выборки более 100 единиц. Когда $n < 8$ при обнаружении отклонений от нормального распределения не дают достоверных результатов. Если число наблюдений (n) менее 50 принадлежность распределения к нормальному проводится с помощью критерия согласия W. При большем экспериментальном материале проверку выполняют с помощью критерия хи-квадрат (χ^2). Критерий χ^2 предпочтителен, когда используются большие объемы выборок (более 100 единиц). При малых объемах выборок этот критерий практически не пригоден.

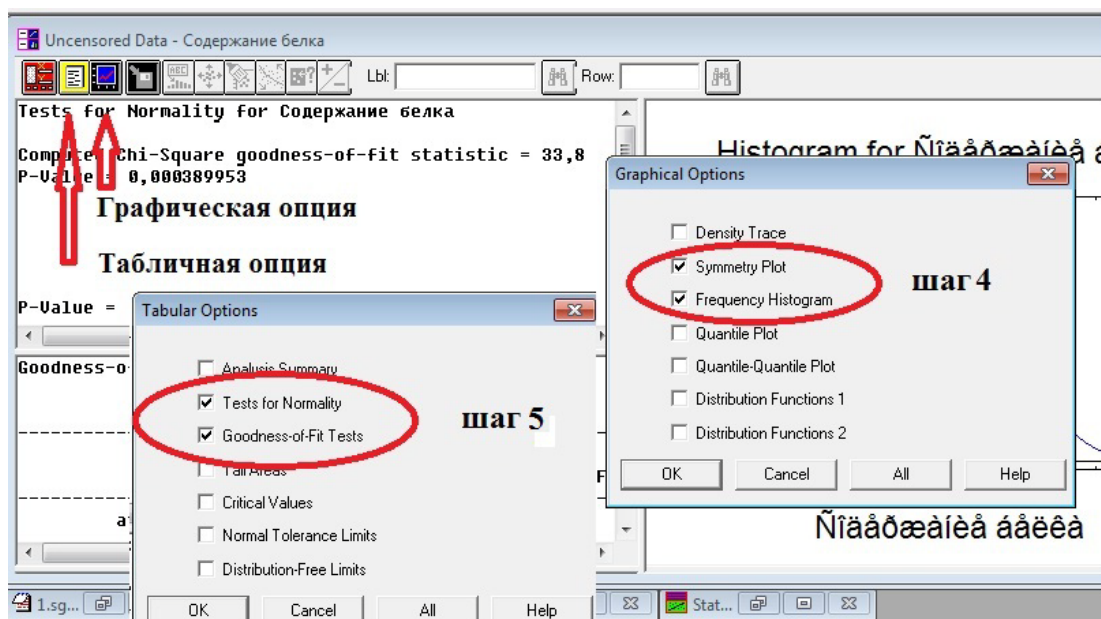


Рис. 7. Ввод данных в окно Graphic Options и Tabular Options

- Нажмите вторую кнопку Tabular Options (табличные опции) на панели инструментов окна. Установите флажки на Tests for Normality (тесты на нормальность) и Goodness – of Fit Tests (тесты согласия) (рис. 7, шаг 5) и нажмите ОК.
- На экране появится окно анализа Distribution Fitting с процедурами проверки нормальности распределения и тестами согласия. Двойным

нажатием левой клавиши мыши откройте окно проверки нормальности данных на весь экран (рис. 8).

Тесты на нормальность для Содержание белка

Хи-квадрат тест = 33.8

P-значение = 0.000389953

Шапиро-Уилкса W-статистика = 0.929838

P-значение = 0.090059

Тесты для малых выборок:

Для асимметрии = 0,518178

P-значение = 0.604331

Для эксцесса = -1,30811

P-значение = 0.190834

СтатКонсультант

Эта панель показывает результаты выполнения различных тестов, определяющих, адекватно ли моделируется нормальным распределением. Хи-квадрат тест показывает диапазон Содержания белка в 14 равно вероятных классах и сравнивает количество наблюдений в каждом классе к количеству ожидаемых. Тест Шапиро-Уилкса базируется на сравнении квантилей, установленных для нормального распределения, с квантилями данных. Стандартизированная асимметрия показывает отсутствие симметрии в данных.

Стандартизированный эксцесс показывает очертание графика распределения, то есть выровненный он или более острый, чем график нормального распределения.

Наименьшее p-значение среди выполненных тестов равно 0.000389953. Так как p-значение этого теста меньше 0.10, мы можем отвергнуть гипотезу о принадлежности Содержания белка к нормальному распределению с вероятностью 90% и более.

Рис. 8. Панель тестов на нормальность

Двойным нажатием левой клавиши мыши верните окно Distribution Fitting в исходное положение. Нажмите два раза на поле Тестов согласия, раскрывая их на весь экран (рис.9).

Тесты согласия для Содержания белка

хи-квадрат = 4,52008 Р-значение = 0.210503

Оцениваемая статистика Колмогорова DPLUS = 0.102779

Оцениваемая статистика Колмогорова DMINUS = 0.141168

Оцениваемая полная статистика DN = 0.141168

Аппроксимированное р-значение = 0.701493

Статконсультант

Эта панель показывает результаты выполнения различных тестов, определяющих, адекватно ли моделируется нормальным распределением. Тест Колмогорова-Смирнова считает максимальное расстояние между эмпирическим распределением Содержания белка и функцией нормального распределения. В этом случае максимальное расстояние 0.141168

Так как р-значение этого теста больше или равно 0.10, мы можем не отвергать гипотезу о принадлежности Содержания белка к нормальному распределению с вероятностью 90% и более.

Рис. 9. Панель тестов согласия

Если выборка принадлежит к нормальному закону распределения для анализа данных можно использовать параметрические методы. В противном случае используйте методы, свободные от распределения.

Вопросы к занятию:

1. Чем отличается эмпирическое распределение от теоретического?
2. Что такое кривая распределения и что она показывает?
3. Какими уровнями значимости можно пользоваться в практике агрономических исследований?
4. От чего зависит уровень значимости?
5. Чем отличаются наблюдаемые частоты от ожидаемых?
6. Какие тесты используют для проверки данных на нормальное распределение?
7. Какие выборки называются малыми, а какие большими?
8. В каких случаях применяют критерий Шапиро-Уилка?
9. В каких случаях применяют критерий согласия?
10. В каких случаях применяют критерий χ^2 ?

3. Параметрические методы статистического анализа

3.1. Корреляционный анализ

В агрономических исследованиях редко приходится иметь дело с точными и определенными функциональными связями, когда каждому значению одной величины соответствует строго определенное значение другой величины. Здесь чаще встречаются такие соотношения между переменными, когда каждому значению признака X соответствует не одно, а множество возможных значений признака Y , т.е. их распределение. Такие связи, обнаруживаемые лишь при массовом изучении признаков, в отличие от функциональных называются стохастическими (вероятностными) или корреляционными. Корреляционный анализ позволяет с помощью выборки делать выводы о степени статистической связи (мера связи) между признаками.

При изучении корреляционных связей возникают два основных вопроса – о тесноте связи и о форме связи. Для измерения тесноты и формы связи используют специальные статистические методы, называемые корреляцией и регрессией. По форме корреляция может быть линейной и криволинейной, по направлению прямой и обратной. Корреляцию и регрессию называют простой, если исследуется связь между двумя признаками, и множественной, когда изучается зависимость между тремя и более признаками.

Для оценки тесноты (силы) связи используют коэффициенты корреляции и корреляционное отношение, обозначаемые буквой r . Он является безразмерной величиной, изменяющейся в области $-1 < r < +1$.

При полных связях, когда корреляционная связь превращается в функциональную, значение коэффициента корреляции равно для положительных, или прямых, связей $+1,0$, для отрицательных или обратных связей $-1,0$. Считается, что при $r < 0,3$ корреляционная зависимость между признаками слабая, $r = 0,3-0,7$ – средняя, а при $r > 0,7$ – сильная.

Связь двух переменных X и Y можно символически представить следующим образом:

$X \rightarrow Y$ (X влияет на Y)

$X \leftrightarrow Y$ (X и Y взаимосвязаны)

$Z \begin{matrix} \nearrow X \\ \searrow Y \end{matrix}$ (X и Y непосредственно зависят от Z)

Выбор метода установления связи зависит от типа переменных. Если обе переменные измеряются количественными шкалами, то для оценки их связи используют корреляцию и регрессию.

Критерии выбора метода для определения связи между переменными

Тип переменной		Метод (примеры)
X	Y	
количественная	количественная	корреляция, регрессия
качественная	количественная	парный t-критерий, дисперсионный анализ
качественная	качественная	критерий χ^2

Например,

- нужно установить связь между сортом (качественные данные) и урожаем (количественные данные). Для этого в опыте испытывают пять сортов. Обработка: дисперсионный анализ;

- нужно установить связь между господствующим типом почвы и формой собственности сельскохозяйственных предприятий в регионе. Для этого рассматривается случайная выборка из предприятий. Предприятия разделяют на классы по формам собственности и классам почвы. Обработка: критерий χ^2 ;

- в некотором исследовании сопоставляли данные урожаев и количества дождевых осадков с апреля по июнь в течение 26 лет. Оба признака являются количественными. Обработка: корреляционный и регрессионный анализ.

Пример 3. Докажите, что имеется связь между относительной влажностью и липкости чернозема используя коэффициент корреляции Пирсона.

Относительная влажность, %: 19,9 20,9 26,1 29,4 30,5 40,3 44,8 48,8 55,6 58,3 64,5 76,6

Липкость, г/см³: 0,0 0,6 1,1 1,2 1,7 1,7 2,6 3,4 4,2 5,8 6,3 7,3

Также определить коэффициент корреляции (r), стандартную ошибку коэффициента корреляции (S_r), критерий существенности коэффициента корреляции (t_b).

- Создайте новую электронную таблицу, Введите имена и типы переменных, сохраните под номером 2.
- Выберите меню Describe→Numeric Data→Multiple-Variable Analysis (анализ множества переменных) (рис. 10, шаг 1). Выделите все переменные. Нажмите ОК (рис. 10, шаг 2). Появляется рабочее поле анализа множества переменных со сводкой, в которой подтверждается, что две переменные приняты к обработке.

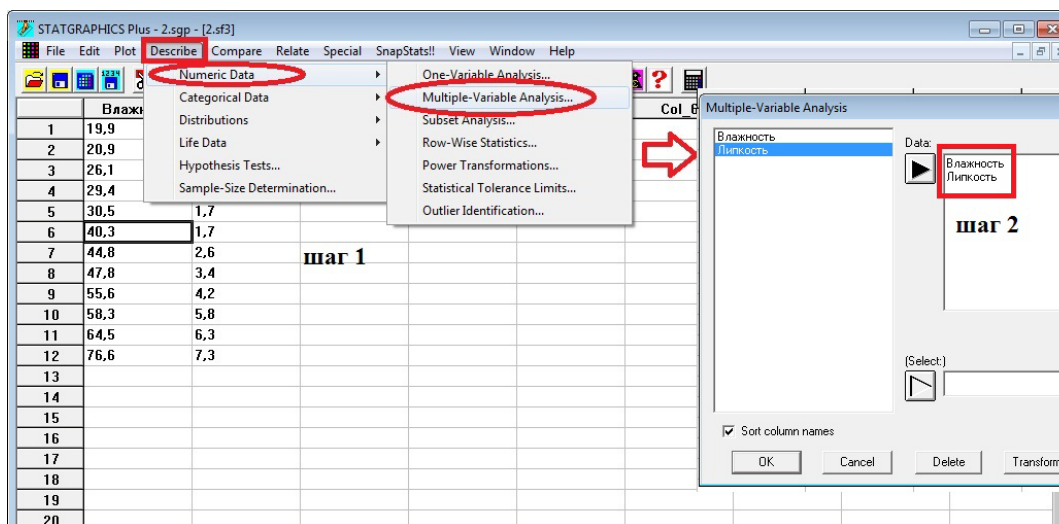


Рис. 10. Выбор данных

- Нажмите кнопку <Табличные опции> (рис. 11, шаг 3) и выберите Correlations (рис. 11, шаг 4).

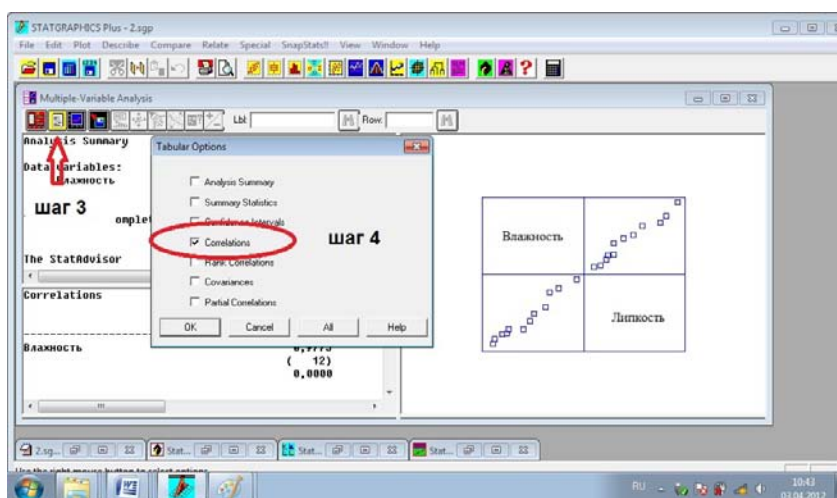


Рис. 11. Изменение выдаваемых статистик

- На экран выдается таблица корреляций (рис. 12)

Корреляции	Влажность	Липкость
Влажность		0,9773 ← r (12) 0,000
Липкость	0,9773 (12) 0,000	

Статконсультант

Эта панель показывает коэффициенты корреляции Пирсона между каждой парой переменных. Значения корреляционных коэффициентов между -1 и +1 и показывают меру линейной связи между переменными. В скобках указаны число пар данных, используемых при подсчете каждого коэффициента. Третье число – это р-значение, показывающее статистическую значимость корреляции. Р-значение меньше 0.05 показывает статистическую значимость ненулевой корреляции с 95% уровнем доверия. Следующие пары переменных имеют р-значение меньше 0.05: Влажность и липкость

Рис.12. Панель корреляционного анализа

Для оценки надежности выборочного коэффициента корреляции вычисляют его ошибку и критерий существенности. Стандартную ошибку коэффициента корреляции (S_r) определяют по формуле:

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

где r – коэффициент корреляции; n – число выборки.

Из формулы следует, что коэффициент корреляции, близкие к единице, оказываются всегда точнее коэффициентов корреляции, близких к нулю. С увеличением объектов исследования S_r также будет всегда уменьшаться, а точность в определении r – возрастать.

Критерий существенности коэффициента корреляции находим по формуле:

$$t_r = r/S_r$$

Если $t_{r \text{ факт}} \geq t_{r \text{ теор}}$, то корреляционная связь существенна, а когда $t_{r \text{ факт}} < t_{r \text{ теор}}$ не существенна. Теоретические значения t находят по таблице Стьюдента (приложение), принимая 5%-ный, а при более строгом подходе 1%-ный уровень значимости. Число степеней свободы принимают равным $n-2$.

В нашем примере $r = 0,977$ и $n = 12$ $S_r = \sqrt{\frac{1-0,977^2}{12-2}} = 0,067$; $t_r = 0,977/0,067 = 14,58$

Результат: Результат обработки показывает сильную положительную связь между переменными «относительная влажность» и «липкость» чернозема. Доверие к ним более 95% ($P < 0,05$). Коэффициент корреляции $r = 0,977$, $S_r = 0,067$, $t_r = 14,58$

Таким образом, можно сделать вывод о наличии сильной положительной связи между относительной влажностью почвы и липкостью.

3. 2. Регрессионный анализ

Для лучшего описания линейной зависимости часто бывает полезным провести прямую через точки корреляционного поля. Ее называют линией регрессии. Наряду с описанием зависимости, регрессию также можно использовать для целей прогноза. С одной стороны, линия регрессии помогает интерпретировать корреляционное поле за счет наглядного представления линейного тренда. Кроме того, ее можно использовать для прогноза. Например, если мы измеряли сумму осадков и урожай в течение 25 лет. По данным строим линию регрессии, и, зная сумму осадков в новом году, можем рассчитать предполагаемый урожай.

На примере 3 рассчитать:

1. коэффициент регрессии (b_{yx}),
2. уравнение регрессии,
3. стандартную ошибку коэффициента регрессии (S_b),
4. ошибку регрессии (S_{yx}),
5. критерий значимости коэффициента регрессии (t_b),
6. доверительный интервал коэффициенту регрессии.

- Откройте файл «2»

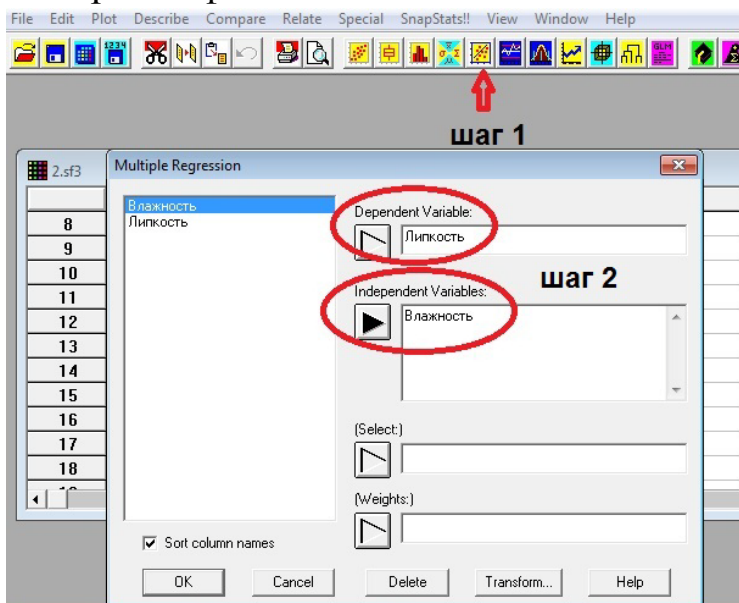


Рис. 13. Выбор данных

- Выберите значок в верхней части окна (рис. 13, шаг 1).
- В появившемся окне множественной регрессии выберете зависимую переменную (dependent) и независимую (independent). В нашем примере №3 липкость – зависимая величина, относительная - влажность не зависимая (рис. 13, шаг 2) и нажмите ОК.
- Появляется рабочее поле анализа. Нажмите кнопку <Табличные опции> (рис. 14, шаг 3) и выберите Analysis Summary и Confidence Intervals как показано на рис. 14, шаг 4 и ОК.

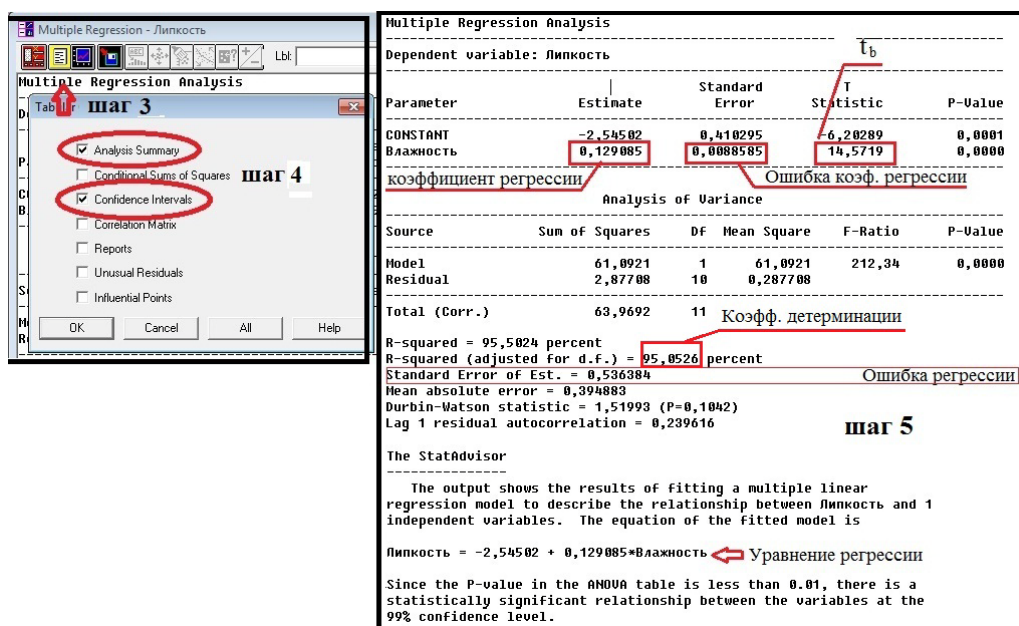


Рис. 14. Рабочее поле анализа

- На экране появиться окно множественного регрессионного анализа. Здесь указаны: коэффициент регрессии (b_{yx}), уравнение регрессии, стандартную ошибку коэффициента регрессии (S_b), ошибку регрессии (S_{yx}), критерий значимости коэффициента регрессии (t_b), коэффициент детерминации (d_{yx}) (рис. 14, шаг 5). Коэффициент регрессии (b_{yx}) показывает, как изменяется Y при изменении X на единицу измерения, и выражается в единицах Y. В нашем случае при изменении влажности на 1% липкость чернозема изменяется на 0,13 г/см². Судя по коэффициенту детерминации, примерно 95% изменений липкости обусловлено изменениям в содержании почвенного гумуса.

Результат: В нашем примере будет: $b_{yx} = 0,13$; $S_b = 0,009$; $S_{yx} = 0,54$; $t_b = 14,57$, $d_{yx} = 95,1$. Уравнение регрессии: $Y = 0,13X - 2,55$.

По уравнению регрессии можно рассчитать значения Y даже тех, которых нет в исходных данных. Однако нельзя использовать уравнение регрессии для интерполяции за пределы таблицы.

- Для нахождения доверительного интервала коэффициенту регрессии нужно двойным щелчком закрыть окно <Multiple Regression Analysis> и двойным щелчком открыть окно <95,0% confidence intervals for coefficient estimates>. На экране появится окно расчета доверительного интервала коэффициента регрессии (рис. 15).

Результат: В нашем примере доверительный интервал буде в диапазоне от 0,109347 до 0,148823. $b_{yx} = 0,13 \pm 0,04$ ($0,11 \div 0,14$)

95,0% confidence intervals for coefficient estimates				
Parameter	Estimate	Standard Error	Lower Limit	Upper Limit
CONSTANT	-2,54502	0,410295	-3,45921	-1,63082
Влажность	0,129085	0,0088585	0,109347	0,148823

↓
↓
коэффициент регрессии
Доверительные интервалы коэф регрессии

The StatAdvisor

This table shows 95,0% confidence intervals for the coefficients in the model. Confidence intervals show how precisely the coefficients can be estimated given the amount of available data and the noise which is present.

Рис. 15. Панель регрессионного анализа

Вопросы к занятию:

1. Почему в агрономических исследованиях используют корреляционный анализ?
2. В чем отличие простого корреляционного анализа от множественного?
3. О чем говорит полученный отрицательный коэффициент корреляции?
4. При каких значениях коэффициента корреляции наблюдается слабая, средняя и сильная корреляционная зависимость между признаками?
5. Что такое линия регрессии, и для каких целей она может использоваться?
6. О чем говорит коэффициент детерминации?
7. Зачем используют коэффициенты корреляции?
8. Какие формы корреляции вы знаете?

9. Если обе переменные измеряются количественно, какой метод установления связи вы используете?

10. Приведите пример количественных переменных?

3.3. Дисперсионный анализ (ДА)

Дисперсионный анализ разработан и введен в практику сельскохозяйственных и биологических исследований английским ученым Р.А. Фишером. Дисперсионный анализ широко используется для планирования эксперимента и статистической обработки его данных. Статистически обоснованный план эксперимента определяет и метод математического анализа результатов. Поэтому современный эксперимент нельзя правильно спланировать, не зная основ дисперсионного анализа.

Сущностью дисперсионного анализа является расчленение общей суммы квадратов отклонений и общего числа степеней свободы на части – компоненты, соответствующие структуре эксперимента, и оценка значимости действия и взаимодействия изучаемых факторов по F-критерию. Он дает возможность получить представление о степени, или доле влияния того или иного фактора в общей дисперсии признака, которую принимают за единицу или 100%.

Дисперсионный анализ быстро вошел в употребление при обработке экспериментальных данных благодаря следующим основным преимуществам его перед методом попарных сравнений по критерию Стьюдента:

1) вместо индивидуальных ошибок, средних по каждому варианту, в дисперсионном анализе используется обобщенная ошибка средних, которая опирается на большее число наблюдений и, следовательно, является более надежной базой для оценок;

2) методом дисперсионного анализа можно обрабатывать данные простых и сложных, однолетних и многолетних, однофакторных и многофакторных опытов.

3) дисперсионный анализ позволяет избежать громоздких вычислений при большом числе вариантов в опыте и позволяет компактно в виде существенных разностей представить итоги статистической обработки.

Современная теория планирования эксперимента и статистический анализ базируется на принципах рандомизации. Теория требует, чтобы все наблюдения были независимы. В этом случае дисперсионный анализ дает правильную, несмещенную оценку опыта.

Дисперсионный анализ невозможен для простых вегетационных или полевых опытов, проведенных без повторностей. Минимум повторностей –

две, однако в этом случае, особенно в полевых опытах, получаются большие ошибки и статистически не доказываются даже значительные эффекты вариантов. Поэтому однофакторные опыты проводят обычно в 4-6 кратной повторности.

Существует закономерность, обоснованная теоретически и экспериментально: эффекты взаимодействия факторов по мере возрастания их порядков убывают (эффект $AB > ABC > ABCD$ и т.д.). При этом парные взаимодействия, которые называют взаимодействия первого порядка, дают обычно значительные, а взаимодействия более высоких порядков между тремя и более факторами, как правило, незначительные и статистически несущественные эффекты (прибавки урожая). Следовательно, в двухфакторных опытах без повторностей теряется ценная информация о парных взаимодействиях и статистически не доказываются даже очень значительные разности по вариантам. Другая опасность постановки многофакторных опытов заключается в том, что случайное выпадение из учета хотя бы одной делянки лишает экспериментатора возможности статистически обработать полученные данные.

3.3.1. Дисперсионный анализ однофакторного опыта

Условия применения ДА:

- Предполагается, что все случайные величины имеют нормальное распределение;
- Дисперсии равны.

Чтобы применить на практике однофакторный ДА, нужно сначала проверить равенства дисперсий. В STATGRAPHICS для этого существуют специальные тесты, основанные на 4 критериях – критерии Кохрана, Бартлетта, Леверье и Хартли. Критерий Кохрана используют при сравнении трёх и более выборок одинакового объёма n . Он более грубый и менее чувствителен к нарушению нормальности.

Критерий Бартлетта является параметрическим и основан на дополнительном предположении о нормальности выборок данных. Поэтому перед применением критерия Бартлетта рекомендуется выполнить проверку нормальности. Критерий Бартлетта очень чувствителен к нарушению данного предположения. Плюсы: 1) объемы выборок могут быть различными (это его преимущество перед критерием Кохрана), 2) критерий Бартлетта выявляет отклонения, как в наибольшую, так и в наименьшую стороны. Минусы: 1) объем каждой выборки должен быть больше трех, 2) критерий очень

чувствителен к нарушению предположения о нормальности закона распределения исходных данных.

Проверка равенности дисперсий

Пример 4 Обработать данные вегетационного опыта с водными культурами по изучению действия соотношения $N : P_2O_5 : K_2O$ при питании рассады томатов на урожай плодов (табл. 2). В каждом варианте по 4 повторностей. Первый вариант контроль с соотношением $N : P_2O_5 : K_2O = 1:1:1$. Второй вариант имеет соотношение $N : P_2O_5 : K_2O = 1:2:1$, третий вариант 1:2:2, четвертый 2:1:1, пятый 2:2:1.

Таблица 2

Ранний урожай плодов (г на сосуд)

Варианты	Урожай			
1(st)	454	470	430	500
2	502	550	490	507
3	601	670	550	607
4	407	412	475	402
5	418	470	460	412

- Создайте новую электронную таблицу. Выберите меню **Compare** → **Multiple Samples** (Сравнение множества выборок). Появится окно анализа **Multiple-Samples Comparison**. Выберите все повторности урожая. По кнопке **<Табличные опции>** выберем **Variance Check** (Соответствие дисперсий) (рис. 16, шаг 1 и 2)

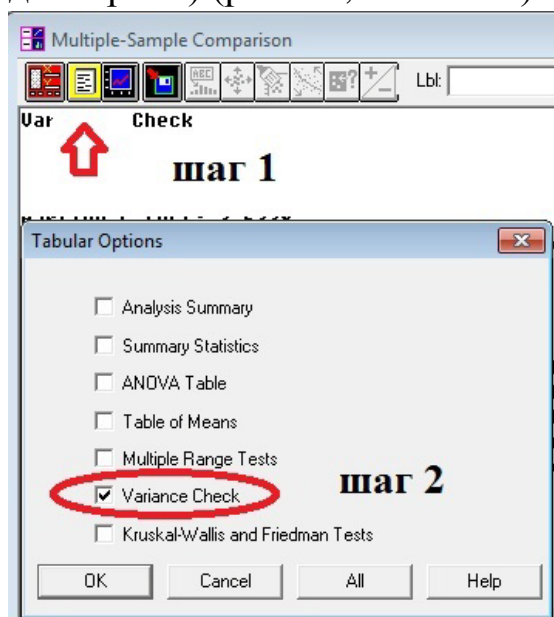


Рис. 16. Изменение выдаваемых статистик

- На экран будет выведена панель сравнения дисперсий по трем критериям (рис. 17).

<p>Проверка дисперсии</p> <p>Тест Кохрана: 0,403454 P-значение = 0,460731 Тест Барлетта: 1,11556 P-значение = 0,835919 Тест Хартли: 3,5338 Тест Леверье: 0,248334 P-значение = 0,906196</p> <p>СтатКонсультант</p> <p>Три статистики, показанные в этой таблице, служат для проверки нулевой гипотезы о том, что стандартные отклонения 5 столбцов одинаковы. Особый интерес имеют два P-значения. Поскольку P-значения больше или равны 0.05, значит, нет статистически значимых различий между стандартными отклонениями с 95% уровнем вероятности.</p>
Рис. 17. Панель проверки дисперсий

- Так как значимых различий между дисперсиями нет, приступаем к дисперсионному анализу.
- Создайте новую электронную таблицу. Введите имена переменных <Вариант> и <Урожай> и заполните электронную таблицу. Вводить в электронную таблицу нужно так, чтобы на каждый вариант соответствовала своя величина урожая. Например:

Варианты	Урожай
1	454
1	470
1	430
1	500
2	502
2	550
2	490 и тд.

- Сохраните под номером 3.
- Выберите меню Compare → Analysis of Variance → One-Way ANOVA... (Однофакторный дисперсионный анализ (ДА)) (Рис 18, шаг 1). В появившемся окне однофакторного ДА выбрать в качестве зависимой переменной (Dependent Variable) - показатели урожая, так как урожай

томатов зависит от соотношения NPK в опыте. В качестве фактора (Factor) выбрать варианты опыта (рис. 18, шаг 2) и ОК.

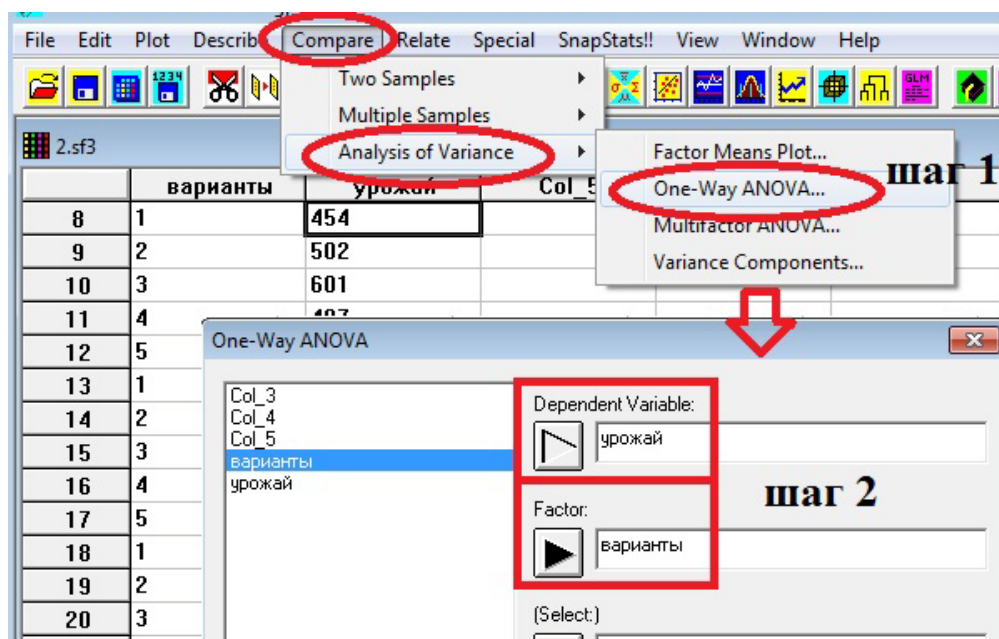


Рис. 18. Выбор данных

- Нажмите кнопку <Табличные опции> (рис. 19, шаг 3) и выберите и отметьте галочками ANOVA Table и Multiple Range Tests и нажмите ОК.

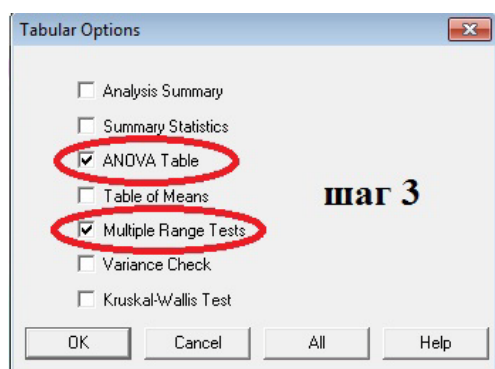


Рис. 19. Изменение выдаваемых статистик

- Двойным щелчком разворачиваем на весь экран результаты однофакторного ДА (рис.20).

ANOVA Table for урожай by варианты

Дисперсия вариантов

Analysis of Variance

Источник	Сумма квадратов	Степ. свободы	Средние квадраты	F-отношение	P-значение
Между группами	86960,8	4	21740,2	18,14	0,0000
Внутри групп	17979,8	15	1198,65		
Полная	104941,0	19			

Остаток (ошибки)

Общая дисперсия

СтатКонсультант

Таблица ДА показывает разложение дисперсии "Урожай" на две компоненты: на межгрупповую и внутригрупповую. F- отношение в этом случае равно 18.1372, это отношение межгрупповой оценки и внутригрупповой оценки. Так как р-значение F-теста меньше, чем 0.05, это говорит о статистической значимости различий между средними "Урожай" в зависимости от "Варианта" с 95% уровнем доверия

Рис. 20. Таблица дисперсионного анализа

- Итоги результатов опыта и статистической обработки данных записывают в таблицу 3

Таблица 3

Результаты дисперсионного анализа

Дисперсия	Сумма квадратов	Степени свободы	Средний квадрат	Fф (критерий существенности)
Общая	104941	19	-	-
Вариантов	86961	4	21740	18,3
Остаток (ошибки)	17980	15	1199	-

$F_f = 18,3$, F_{05} находим по таблице 2 приложений, исходя из 4 степеней свободы для дисперсии вариантов (числитель) и 15 степеней свободы для остатка (знаменатель). При $F_f > F_{05}$ в опыте есть существенные различия по вариантам на 5% уровне значимости.

Результат дисперсионного анализа нам показал на наличие значимых различий средних вариантов, однако неизвестно, какие именно средние различаются. Поэтому, после проведения дисперсионного анализа проводят попарные сравнения средних значений. Простейшим из них является оценка значимости разности между средними по наименьшей существенной разности НСР (сравнение на основе t-критерия). В англоязычной литературе LSD – Least Significant Difference. В опытной работе чаще всего проводят попарные сравнения средних по вариантам и вычисляют НСР. Но иногда, например,

когда в опыте нет контрольного варианта, возникает необходимость сравнить средние урожаи опытных вариантов со средним урожаем в опыте.

В системе государственного сортоиспытания сельскохозяйственных культур на основе $НСР_{05}$ все сорта распределяются на три группы:

I – отклонение от средних урожаев от стандарта (контроля) с положительным знаком больше $НСР_{05}$

II – отклонение с положительным или отрицательным знаком не выходят за пределы $НСР_{05}$;

III – отклонение с отрицательным знаком больше по абсолютной величине $НСР_{05}$.

- Двойным щелчком закрываем окно ДА и двойным щелчком открываем окно расчета $НСР_{05}$ (рис. 21).

Multiple Range Tests for урожай by вариант			
Method: 95,0 percent LSD			
вариант	Число наблюдений	Средние	Однородные группы
4	4	424,0	X
5	4	440,0	X
1	4	463,5	XX
2	4	512,25	X
3	4	607,0	X
Contrast		Разность	$НСР_{05}$
1 - 2		-48,75	52,1804
1 - 3		*-143,5	52,1804
1 - 4		39,5	52,1804
1 - 5		23,5	52,1804
2 - 3		*-94,75	52,1804
2 - 4		*88,25	52,1804
2 - 5		*72,25	52,1804
3 - 4		*183,0	52,1804
3 - 5		*167,0	52,1804
4 - 5		-16,0	52,1804

* Обозначает статистически значимые различия

СтатКонсультант

В этой таблице применяется несколько процедур сравнения, чтобы определить, какие средние существенно отличаются от других средних. В нижней части страницы показывается предполагаемое различие между каждой парой предних. Звездочками отмечаются 6 пар, указывая, что эти пары показывают статистически значимые различия в 95% уровне. В верхней части страницы, 3 однородные группы идентифицируются с помощью столбцов X. Внутри каждой колонки, уровни, содержащие X, образуют группу средних, в которых нет статистически значимых различий.

Рис. 21. Оценка наименьшей существенной разности

Если количество повторений одинаково для всех вариантов, то $НСР_{05}$ будет одинакова для всех сравнений. Разность рассчитывается относительно к стандарту. В нашем примере для второго варианта разность будет составлять:

463,5-512,2 = -48,75 (рис. 21), но в таблицу 4 нужно данные вносить противоположным знаком, так как второй вариант дал урожай на 48,75 г на сосуд больше чем контрольный вариант.

По данным рис.21 заполняем таблицу 4 (разность со стандартом с противоположным знаком).

Таблица 4

Ранний урожай плодов томатов (г на сосуд)

Соотношение N : P ₂ O ₅ : K ₂ O	Урожай (среднее значение)	Разность со стандартом, г	Группа
1:1:1	463,5	-	st
1:2:1	512,2	48,7	II
1:2:2	607,0	143,5	I
2:1:1	424,0	-39,5	II
2:2:1	440,0	-23,5	II
НСР ₀₅	-	52,2	-

В нашем примере просят рассмотреть относительно к контрольному варианту, поэтому в рис. 21 смотрим разность 1-2, 1-3, 1-4, 1-5. Рядом указан НСР₀₅. Так как число наблюдений одинаково в каждом варианте НСР₀₅ будет иметь одинаковое значение.

В нашем примере из таблицы 4 и рис. 21 (однородные группы) видно, что третий вариант выделяется и относится I группе, где наблюдается наибольшее отклонение урожая с положительным знаком от контроля. Таким образом, усиленное питание рассады одновременно фосфором и калием обеспечивает получение более ранних и более высоких урожаев (соотношение 1:2:2).

В 4 и 5 вариантах отклонение с отрицательным знаком не выходят за пределы НСР₀₅, и показывают, что при усилении азотного питания наблюдается тенденция к снижению урожая.

3.3.2. Дисперсионный анализ многофакторного опыта.

Пример:

В двухфакторном опыте 2x3 с культурой ячменя изучено действие двух доз азота и трех доз фосфора (табл. 5). Провести дисперсионный анализ результатов опыта.

Урожай зерна ячменя в двухфакторном опыте 2x3

Азот - а	Фосфор - в	Урожай
а1 (без азота)	в 1	24,1 25,8 23,0 27,0
	в 2	28,4 29,7 30,1 27,4
	в 3	28,7 30,4 32,0 27,0
а 2	в 1	30,7 34,4 34,0 31,0
	в 2	46,7 45,4 47,1 46,3
	в 3	59,4 50,7 64,5 60,1

- В электронную таблицу вводим, так как показано на рис. 22. Каждая строчка должна быть заполнена.

	Азот	фосфор	урожай
1	1	1	24,1
2	1	2	28,4
3	1	3	28,7
4	2	1	30,7
5	2	2	46,7
6	2	3	59,4
7	1	1	25,8
8	1	2	29,7
9	1	3	30,4
10	2	1	34,4
11	2	2	45,4
12	2	3	50,7
13	1	1	23
14	1	2	30,1
15	1	3	32
16	2	1	34
17	2	2	47,1

Рис 22 Пример заполнения электронной таблицы

- Выберите меню Compare → Analysis of Variance → Multifactor ANOVA... (Многофакторный дисперсионный анализ (ДА)) (Рис 23 шаг 1).

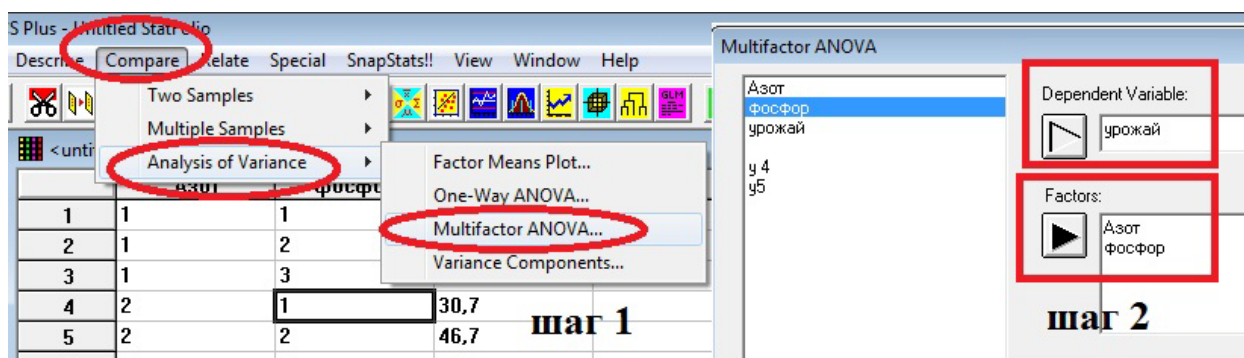


Рис. 23. Выбор дисперсионного анализа

- В появившемся окне многофакторного ДА выбрать в качестве зависимой переменной (Dependent Variable) - показатель урожая. В качестве фактора (Factor) выбрать азот и фосфор (рис. 23, шаг 2) и ОК.
- Нажмите кнопку <Табличные опции> (рис. 24, шаг 3) и выберите и отметьте галочками ANOVA Table, Table of Means и Multiple Range Tests и нажмите ОК.

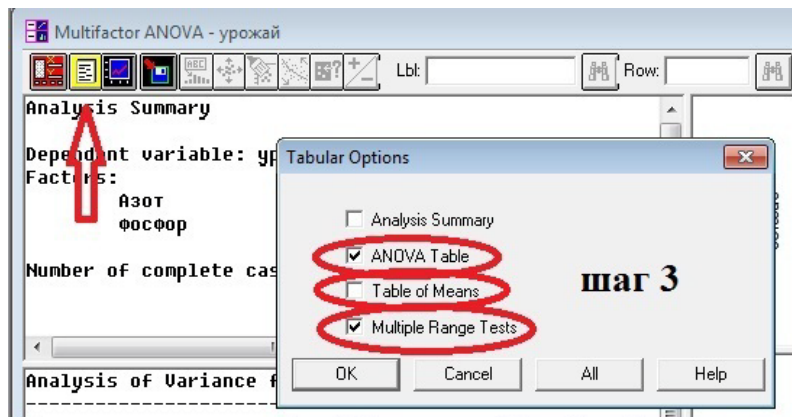


Рис. 24. Изменение выдаваемых статистик

- Двойным щелчком разворачиваем на весь экран результаты многофакторного ДА. В появившемся результате не учитывается взаимодействие фактора А и В. Чтобы данное взаимодействие вошло в статистику нужно щелкнуть правой кнопкой мыши и выбрать Analysis Options (рис. 25, шаг 4). Появится окно опций многофакторного ДА. В максимальное количество взаимодействий указать 2 (рис. 25, шаг 5) и нажать ОК.

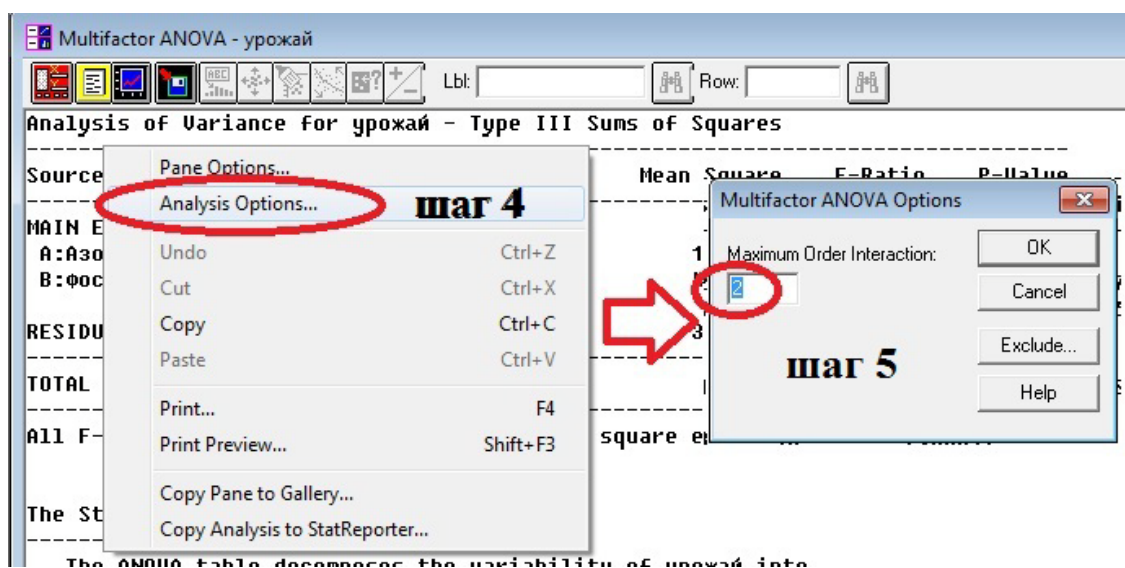


Рис. 25. Изменение выдаваемых статистик

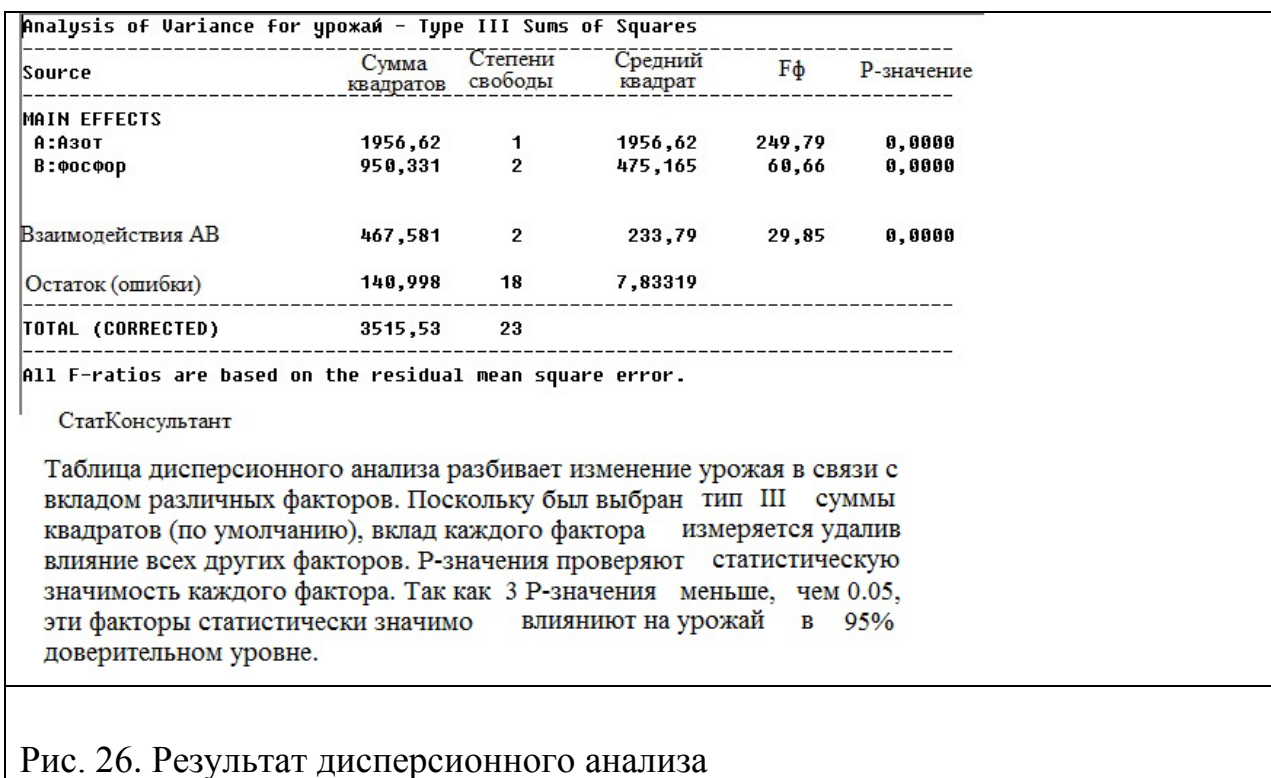


Рис. 26. Результат дисперсионного анализа

- Итоги результатов опыта и статистической обработки данных записывают в таблицу 6

Таблица 6

Результаты дисперсионного анализа

Дисперсия	Сумма квадратов	Степени свободы	Средний квадрат	Fф
Общая	3515,5	23	-	-
Азота, А	1956,6	1	1956,60	249,88
Фосфора, В	950,3	2	475,15	60,68
Взаимодействия, АВ	467,6	2	233,80	29,85
Остаток (ошибки)	141,0	18	7,83	-

- Двойным щелчком закрываем окно многофакторного ДА и двойным щелчком открываем окно таблицы средних (рис.27).

Table of Least Squares Means For урожай with 95,0 Percent Confidence Intervals					
Level	Count	Mean	$S_{\bar{x}}$	Lower Limit	Upper Limit
GRAND MEAN	24	36,8292			
Азот					
1	12	27,8	0,807939	26,1026	29,4974
2	12	45,8583	0,807939	44,1609	47,5558
фосфор					
1	8	28,75	0,98952	26,6711	30,8289
2	8	37,6375	0,98952	35,5586	39,7164
3	8	44,1	0,98952	42,0211	46,1789
Азот by фосфор					
1 1	4	24,975	1,39939	22,035	27,915
1 2	4	28,9	1,39939	25,96	31,84
1 3	4	29,525	1,39939	26,585	32,465
2 1	4	32,525	1,39939	29,585	35,465
2 2	4	46,375	1,39939	43,435	49,315
2 3	4	58,675	1,39939	55,735	61,615

Рис. 27. Результаты дисперсионного анализа

- В рис. 27 рассчитаны средние значения отдельно для действия азотных, фосфорных удобрений и действие фосфорных удобрений на урожай ячменя в зависимости от обеспеченности азотом (азот by фосфор). Так же рассчитан $S_{\bar{x}}$ (обобщенная ошибка средней) равная 1,399.
- Совместное действие азота и фосфора можно представить в виде таблицы 7

Таблица 7

Действие азота и фосфора на урожай ячменя

Доза азота	Дозы фосфора		
	1	2	3
1 (без азота)	25,0	28,9	29,5
2	32,5	46,4	58,7

Чтобы построить график нужно нажать кнопку <Графические опции> (Graphical Options), которая находится рядом с табличными опциями (рис. 24) и отметить галкой <Interaction Plot>.

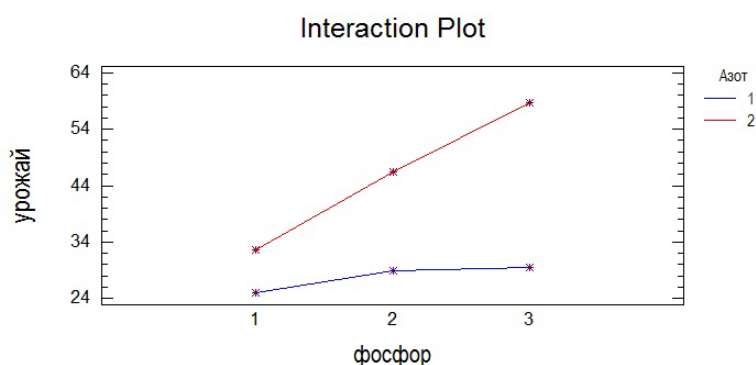


Рис. 28 Действие фосфорных удобрений на урожай ячменя в зависимости от обеспеченности азотом (1 без азота, 2 – по фону азота)

Вопросы к занятию

1. Кем был разработан дисперсионный анализ и введен в практику сельскохозяйственных и биологических исследований?
2. В чем суть дисперсионного анализа?
3. На каких этапах исследования используется дисперсионный анализ?
4. Перечислите преимущества дисперсионного анализа.
5. Какое оптимальное количество повторений необходимо для проведения ДА?
6. Почему нужно планировать опыты с повторностями?
7. Какая информация теряется в многофакторных опытах без повторностей, и какие опасности могут произойти?
8. В чем суть парных взаимодействий?
9. Что такое НСР и чем отличаются НСР₀₅ и НСР₀₁?
10. Условия применения ДА?

Приложение

Значения критерия t на 5% (0,05), 1% (0,01) и 0,1% - ном (0,001) уровни значимости.

Таблица 1

Число степеней свободы	Уровень значимости		
	0,05	0,01	0,001
1	12,71	63,66	-
2	4,3	9,93	31,60
3	3,18	5,84	12,94
4	2,78	4,60	8,61
5	2,57	4,03	6,86
6	2,45	3,71	5,96
7	2,37	3,50	5,41
8	2,31	3,36	5,04
9	2,26	3,25	4,78
10	2,23	3,17	4,59
11	2,20	3,11	4,44
12	2,18	3,06	4,32
13	2,16	3,01	4,22
14	2,15	2,98	4,14
15	2,13	2,95	4,07
16	2,12	2,92	4,02
17	2,11	2,90	3,97
18	2,10	2,88	3,92
19	2,09	2,86	3,88
20	2,09	2,85	3,85
21	2,08	2,853	3,82
22	2,07	2,82	3,79
23	2,07	2,81	3,77
24	2,06	2,80	3,75
25	2,06	2,79	3,73
26	2,06	2,78	3,71
27	2,05	2,77	3,69
28	2,05	2,76	3,67
29	2,05	2,76	3,66
30	2,04	2,75	3,65
50	2,01	2,68	3,50
100	1,98	2,63	3,39
∞	1,96	2,65	3,29

Значения критерия F на 5%-ном уровне значимости
(вероятность 95%)

Степень свободы для меньшей дисперсии (знаменателя)	Степень свободы для большей дисперсии (числителя)						
	1	2	3	4	5	6	7
1	161	200	216	225	230	234	237
2	18,51	19,00	19,16	19,25	19,30	19,33	19,36
3	10,13	9,55	9,28	9,12	9,01	8,94	8,88
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88
6	5,99	5,14	4,76	4,53	4,39	4,27	4,21
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01
12	4,75	3,88	3,49	3,26	3,11	3,00	2,92
13	4,64	3,80	3,41	3,18	3,02	2,92	2,84
14	4,60	3,74	3,34	3,11	2,96	2,85	2,77
15	4,54	3,60	3,29	3,06	2,90	2,79	2,70
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66
17	4,45	3,59	3,20	2,96	2,81	2,70	2,62
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58
19	4,38	3,52	3,13	2,90	2,74	2,63	2,55
20	4,35	3,49	3,10	2,87	2,71	2,60	2,52
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49
22	4,30	3,44	3,05	2,82	2,66	2,55	2,47
23	4,28	3,42	3,03	2,80	2,64	2,53	2,45
24	4,26	3,40	3,01	2,78	2,62	2,51	2,43
25	4,24	3,38	2,99	2,76	2,60	2,49	2,41

Степень свободы для меньшей дисперсии (знаменателя)	Степень свободы для большей дисперсии (числителя)						
	8	9	10	12	24	50	100
1	239	241	242	244	249	252	253
2	19,37	19,38	19,39	19,41	19,45	19,47	19,49
3	8,84	8,81	8,78	8,74	8,64	8,58	8,56
4	6,04	6,00	5,96	5,91	5,77	5,70	5,66
5	4,82	4,78	4,74	4,68	4,53	4,44	4,40
6	4,15	4,10	4,06	4,00	3,84	3,75	3,71
7	3,73	3,68	3,63	3,57	3,41	3,32	3,28
8	3,44	3,39	3,34	3,28	3,12	3,03	2,98
9	3,23	3,18	3,13	3,07	2,90	2,80	2,76
10	3,07	3,02	2,97	2,91	2,74	2,64	2,59
11	2,95	2,90	2,86	2,79	2,61	2,50	2,45
12	2,85	2,80	2,76	2,69	2,50	2,40	2,35
13	2,77	2,72	2,67	2,60	2,42	2,32	2,26
14	2,70	2,65	2,60	2,53	2,35	2,24	2,19
15	2,64	2,59	2,55	2,78	2,29	2,18	2,12
16	2,59	2,54	2,49	2,42	2,24	2,13	2,07
17	2,55	2,50	2,45	2,38	2,19	2,08	2,02
18	2,51	2,46	2,41	2,34	2,15	2,04	1,98
19	2,48	2,43	2,38	2,31	2,11	2,00	1,94
20	2,45	2,40	2,35	2,28	2,08	1,96	1,90
21	2,42	2,37	2,32	2,25	2,05	1,93	1,87
22	2,40	2,35	2,30	2,23	2,03	1,91	1,84
23	2,38	2,32	2,28	2,20	2,00	1,88	1,82
24	2,36	2,30	2,26	2,18	1,98	1,86	1,80
25	2,34	2,25	2,24	2,16	1,96	1,84	1,77

Сборник задач к курсу

«Статистическая обработка результатов полевых
агрохимических исследований
с помощью пакета
STATGRAPHICS PLUS for Windows»

Задания
Общие статистики (Summary statistics)

Опишите выборку в таблице 1, используя описательные статистики: средняя арифметическая, дисперсия, стандартное отклонение, ошибка средней арифметической, коэффициент вариации и относительная ошибка выборочной средней.

Таблица 1

Урожай, ц/га

Номера делянок	Урожай по номерам заданий ц/га									
	1	2	3	4	5	6	7	8	9	10
1	26,0	26,8	20,3	21,6	29,8	18,2	16,8	29,0	30,9	20,2
2	27,1	29,5	29,7	18,7	31,4	19,3	19,6	26,6	28,2	21,3
3	30,2	26,5	20,3	20,8	31,6	21,1	17,6	29,6	34,2	22,1
4	28,5	28,7	24,2	24,7	33,0	21,9	18,7	31,2	31,0	24,7
5	29,6	26,5	20,0	21,0	34,0	23,3	22,6	28,7	26,9	20,7
6	27,2	30,4	18,9	22,8	30,8	20,0	18,5	32,4	31,1	25,6
7	28,4	26,8	19,8	23,0	31,1	19,0	23,8	28,3	30,0	21,8
8	31,6	26,6	24,5	19,9	26,9	18,6	18,6	33,2	28,4	20,2
9	30,5	24,1	20,7	22,5	30,8	23,6	20,4	29,6	32,1	21,6
10	30,0	27,2	23,4	24,3	28,9	15,2	19,0	30,6	29,0	20,2
11	26,2	27,0	21,5	21,8	29,0	21,4	24,0	29,2	31,1	18,4
12	26,3	26,7	22,9	19,9	30,6	18,5	18,8	27,8	28,4	21,5
13	30,4	30,7	28,5	26,0	37,8	21,3	16,8	29,2	31,4	22,3
14	30,7	26,9	21,4	22,9	34,2	20,1	18,9	30,4	27,2	21,9
15	29,8	31,6	26,2	27,2	31,2	19,5	17,8	28,9	30,1	23,9
16	27,4	26,6	20,1	20,0	33,0	23,2	18,7	26,6	31,3	21,8
17	28,6	27,0	20,0	21,2	30,3	20,1	19,0	28,5	24,2	22,0
18	31,8	26,8	23,6	19,1	28,1	18,8	20,8	26,4	28,6	20,4
19	32,7	29,3	22,9	22,7	31,0	19,8	20,6	29,8	32,3	26,8
20	34,2	27,4	21,6	20,5	29,1	18,4	19,2	24,8	29,2	20,4

- | | |
|--------------------------|----------------------------|
| 1. номера делянок 1 и 2 | 10. номера делянок 10 и 11 |
| 2. номера делянок 2 и 3 | 11. номера делянок 11 и 12 |
| 3. номера делянок 3 и 4 | 12. номера делянок 12 и 13 |
| 4. номера делянок 4 и 5 | 13. номера делянок 13 и 14 |
| 5. номера делянок 5 и 6 | 14. номера делянок 14 и 15 |
| 6. номера делянок 6 и 7 | 15. номера делянок 15 и 16 |
| 7. номера делянок 7 и 8 | 16. номера делянок 16 и 17 |
| 8. номера делянок 8 и 9 | 17. номера делянок 16 и 17 |
| 9. номера делянок 9 и 10 | 18. номера делянок 17 и 18 |
| | 19. номера делянок 18 и 19 |
| | 20. номера делянок 19 и 20 |

Проверка данных на нормальное распределение

Проведите проверку нормальности распределения:

1. В выборке растений кукурузы измеряли длину:

175 172 171 179 168 171 163 177 176 186 193 172 162 161 182

2. Стекловидность семян пшеницы составила:

91 92 90 97 98 93 95 91 94 93 95 97 94 91 98 95 96 93 92 93

3. Техническая длина стебля (см) у 30 растений льна: 90 110 99 100 68

70 72 73 70 76 82 80 68 69 74 72 69 80 59 79 81 84 108 83 81 99 98
109 101 88

4. При определении содержания фосфора в растительном материале получены следующие результаты (в мг P_2O_5 на 100 г сухого вещества):

0,56 0,52 0,56 0,53 0,49 0,49 0,41 0,57 0,48 0,47 0,51 0,54 0,56 0,53 0,49
0,49 0,49 0,41 0,57 0,48 0,47 0,53 0,49 0,54 0,56

5. В вегетационном опыте получены урожаи томатов по параллельным сосудам (г/сосуд):

587 564 539 541 561 604 551 468 470 482 452 530 555 601 534 541
543 560 561 456 568

6. Содержание крахмала в картофеле сорта Даренка, %:

15,4 14,8 17,1 16,8 14,2 15,8 16,4 17,2 14,5 14,9 14,8 16,7 18,1 17,5 13,2
14,5 16,2 17,1 17,5 13,2

7. Лежкость картофеля сорта Голубизна составляет, %:

86 85 84 86 90 87 86 82 81 92 73 86 88 85 89 61 88 85 84 92 93 85
88 67 75 84 88 86 85 87 82

8. Урожай огурца сорта Архангельск, выращенный теплице, в июне месяце составил:

23,3 22,3 16,3 18,8 15,3 18,6 18,2 16,8 18,7 15,2 14,6 12,2 20,0 17,5 21,4
27,8 25,0 25,0 22,8 22,5 16,0 22,5 25,0 16,1

9. Урожай огурца сорта Архангельск, выращенный теплице, в марте
месяце составил:

2,1 4,1 6,1 7,5 9,5 10,0 9,8 9,1 14,0 17,4 17,8 20,3 12,5 12,5 17,1 22,2
20,0 20,0 15,9 20,0 14,0 20,0 20,0 14,3

10. Урожай огурца сорта Архангельск, выращенный теплице, в апреле:

14,8 14,2 30,3 18,8 21,0 16,7 16,4 15,2 23,3 21,8 20,7 20,3 20,0 17,5 21,4
27,8 25,0 25,0 22,8 22,5 16,0 22,5 25,0 16,1

11. Деловой выход рассады (шт/м²) капусты белокочанной:

210 220 223 215 265 254 254 210 264 275 200 215 220 240 245 244 235
255 254 210 211 213 264 250 257 248 231 220 210 250

12. Деловой выход рассады (шт/м²) томатов:

100 110 120 125 125 126 127 105 110 115 120 117 116 121 123 124 110
112 118 125 120 119т 116 110 111 110 116 121 123 124 110 112 126
127 105 110

13. Деловой выход рассады перца (шт/м²):

155 171 156 175 165 145 148 168 157 150 175 165 145 148 155 164 170
171 174 150 167 162 163 156

165 148 155 164 170 171 174 150 167 145 148 155 164 170 171 174 150

14. При влажности почвы 90% получили следующий урожай огурцов
(масса плодов на одном растении, г):

591 568 546 598 579 564 587 588 592 584 581 580 581 591 579 564 587
588 592 584 587 584 564 598 581 564

15. При влажности почвы 65% получили следующий урожай томатов
(масса плодов на одном растении, г):

146 154 155 152 164 148 147 139 138 137 145 144 143 143 143 141 140
154 155 149 148 156 155 156 156 156 147 146 140

16. При влажности почвы 40% получили следующий урожай томатов
(масса плодов на одном растении, г):

39 40 45 38 36 29 28 38 37 31 34 45 41 40 41 39 38 36 35 38 37 31
34 45 41 40 41 45 38 36 29 28 38

17. Урожайность лука на перу составил (кг/ м²):

12 12 11 15 16 12 12 14 15 16 17 12 12 13 14 17 14 14 13 12 11 10
11 12 14 16 16 14 15 12 12 13 14 17 14 14

18. Содержание сырой клейковины в тритикале составляет (%):

12,9 16,4 9,5 11,2 10,2 12,9 12,1 12,0 11,5 11,8 9,8 15,4 10,2 12,9 12,1
12,0 9,8 15,8 11,8 11,7 14,8 14,7 14,5 12,0 11,5 11,8 9,8 15,4 10,2 12,9

19. Содержание сырой клейковины в тритикале сорта Юбилейная составляет (%):

21,7 21,5 21,6 21,0 19,8 21,0 21,4 21,7 21,0 20,9 20,5 21,1 21,0 21,0 19,8
21,0 21,4 21,7

20. Высота соломины (см) у зерновых сортов тритикале составляет:

81 115 110 84 90 95 96 100 110 81 85 86 92 105 104 103 100 95 96
100 110

Корреляционный и регрессионный анализ

Провести корреляционный и регрессионный анализ данных. Определить коэффициент корреляции (r), стандартную ошибку коэффициента корреляции (S_r), критерий значимости коэффициенту корреляции, коэффициент регрессии (b_{yx}), уравнение регрессии, стандартную ошибку коэффициента регрессии (S_b), ошибку регрессии (S_{yx}), критерий значимости коэффициента регрессии (t_b), доверительный интервал коэффициенту регрессии, коэффициент детерминации (d_{yx}). Варианты:

1. Определить пораженность льна-долгунца фузариозом (ряд Y) в зависимости от интервала между посевом на одном и том же поле восприимчивых к грибным патогенам сортов льна-долгунца (ряд X в табл. 2). По уравнению регрессии рассчитайте пораженность льна-долгунца фузариозом в интервале между посевом 11 лет. Сделать выводы.

Таблица 2

Таблица показателей опыта

№ пары	Значение признаков	
	X , годы	Y , %
1	1	88
2	2	76
3	3	70
4	4	5
5	5	12
6	6	28
7	7	45
8	8	45
9	9	9
10	10	62

2. Провести корреляционный и регрессионный анализ для выборочной совокупности таблицы 3, в которой представлены результаты определения содержания гумуса и подвижных форм фосфатов в пахотном слое

легкосуглинистой дерново-подзолистой почвы. По уравнению регрессии рассчитайте, какое будет содержание подвижных форм фосфатов в пахотном слое, при содержании гумуса 1,56% и 2,10%. Сделать выводы.

Таблица 3

Таблица показателей опыта

Номер пары	X	Y	Номер пары	X	Y
1	1,57	30	17	1,35	17
2	1,58	28	18	1,31	17
3	1,21	25	19	1,29	16
4	1,21	27	20	1,38	17
5	1,41	25	21	1,38	16
6	1,47	24	22	1,36	14
7	1,45	25	23	1,36	16
8	1,49	27	24	1,20	17
9	1,38	24	25	1,36	16
10	1,41	25	26	1,29	14
11	1,55	25	27	1,30	12
12	1,45	25	28	1,32	12
13	1,30	22	29	1,17	11
14	1,30	22	30	1,22	11
15	1,39	20	31	1,09	9
16	1,46	22	32	1,13	9

3. Провести корреляционный и регрессионный анализ для выборочной совокупности таблицы 4, в которой представлены результаты определения содержания гумуса и подвижных форм фосфатов в пахотном слое легкосуглинистой дерново-подзолистой почвы. По уравнению регрессии рассчитайте, какое будет содержание подвижных форм фосфатов в пахотном слое, при содержании гумуса 0,97% и 2,30%. Сделать выводы.

Таблица 4

Таблица показателей опыта

Номер пары	X	Y	Номер пары	X	Y
1	0,96	16	17	1,42	27
2	1,08	9	18	1,36	25
3	1,16	19	19	1,55	24
4	1,12	17	20	1,36	22
5	1,01	11	21	1,46	28
6	1,07	11	22	1,39	28

7	1,10	16	23	1,63	36
8	1,22	17	24	1,57	36
9	1,22	16	25	1,37	27
10	1,12	19	26	1,48	25
11	0,86	20	27	1,61	28
12	0,79	19	28	1,61	30
13	1,19	23	29	1,70	28
14	1,15	22	30	1,62	28
15	1,13	18	31	1,04	13
16	1,34	20	32	1,22	10

4. В некотором исследовании сопоставляли данные урожаев (y) и количества дождевых осадков с апреля по июнь (x) в течение 26 лет (табл. 5). По уравнению регрессии рассчитайте какой был бы урожай в следующем году выпадет 185 и 300 мм осадков. Сделать выводы.

Таблица 5

Таблица показателей опыта

Год	Осадки (X), мм	Урожай (Y) ц/га	Год	Осадки (X), мм	Урожай (Y) ц/га
1900	177	26,2	1913	108	19,4
1901	96	25,0	1914	132	30,6
1902	144	32,6	1915	89	16,4
1903	105	26,6	1916	147	30,4
1904	111	19,6	1917	98	19,2
1905	135	20,4	1918	106	20,2
1906	209	29,2	1919	123	25,6
1907	161	33,8	1920	156	31,0
1908	246	26,6	1921	191	35,8
1909	108	22,6	1922	162	31,6
1910	137	24,2	1923	235	33,6
1911	71	16,6	1924	147	30,4
1912	119	29,8	1925	110	30,2

5. Провести корреляционный и регрессионный анализ для выборочной совокупности таблицы 6, в которой представлены результаты определения плотности сложения почвы г/см^3 и общей пористости %. По уравнению регрессии рассчитайте общую порозность при плотности сложения равной 0,98 и 1,41 г/см^3 . Сделать выводы.

Таблица показателей опыта

Номер пары	Плотность сложения, г/см ³	Общая порозность, %	Номер пары	Плотность сложения, г/см ³	Общая порозность, %
1	1,07	59,8	13	1,49	46,6
2	1,07	59,3	14	1,36	51,4
3	1,33	51,5	15	1,11	57,8
4	1,38	50,0	16	1,11	58,1
5	1,42	48,7	17	1,49	46,6
6	1,29	53,1	18	1,58	42,8
7	1,21	54,0	19	1,58	42,1
8	1,21	53,5	20	1,56	45,1
9	1,21	54,5	21	1,56	45,1
10	1,21	54,3	22	1,50	45,1
11	1,40	49,6	23	1,50	45,1
12	1,45	47,7	24	1,22	53,8

6. Провести корреляционный и регрессионный анализ для выборочной совокупности таблицы 7, в которой представлены результаты определения плотности сложения почвы г/см³ и общей пористости. По уравнению регрессии рассчитайте общую порозность при плотности сложения равной 0,98 и 1,80 г/см³. Сделать выводы.

Таблица 7

Таблица показателей опыта

Номер пары	Плотность сложения, г/см ³	Общая порозность, %	Номер пары	Плотность сложения, г/см ³	Общая порозность, %
1	1,22	53,8	13	1,20	54,7
2	1,22	53,8	14	1,37	49,8
3	1,22	53,8	15	1,37	49,8
4	1,53	44,4	16	1,61	42,3
5	1,53	44,4	17	1,79	35,1
6	1,59	43,4	18	1,75	37,1
7	1,59	42,8	19	1,20	55,4
8	1,42	49,3	20	1,20	54,4
9	1,51	46,6	21	1,20	54,7
10	0,88	65,6	22	1,20	54,7
11	0,88	65,6	23	1,20	54,7
12	1,20	55,4	24	1,71	38,5

7. Провести корреляционный и регрессионный анализ для выборочной совокупности таблицы 8, в которой представлены результаты определения влажности грунта (% к абсолютно сухой почве) и осмотической давлению (атм) почвенного раствора. По уравнению регрессии рассчитайте осмотическое давление почвенного раствора при влажности грунта 3,0 и 15,0%. Сделать выводы.

Таблица 8

Таблица показателей опыта

Влажность грунта (% к абсолютно сухой почве)	Осмотической давлению (атм)
5,8	965
8,7	418
11,9	138
13,2	72
14,9	38
17,1	15,2
18	7,6
18,9	3,8

8. Провести корреляционный и регрессионный анализ для выборочной совокупности таблицы 9, в которой представлены результаты определения плотности сложения почвы (г/см³) и полной влагоемкости (%) почв. По уравнению регрессии рассчитайте полной влагоемкости (%) при плотности сложения почвы 0,81 и 1,05 (г/см³). Сделать выводы.

Таблица 9

Таблица показателей опыта

Номер пары	Плотность сложения, г/см ³	Полная влагоемкость, %	Номер пары	Плотность сложения, г/см ³	Полная влагоемкость, %
1	1,07	56,3	9	1,49	31,6
2	1,33	39,9	10	1,36	39,4
3	1,38	30,6	11	1,11	59,0
4	1,42	33,7	12	1,49	30,5
5	1,29	42,3	13	1,58	28,2
6	1,21	43,8	14	1,56	27,3
7	1,40	32,8	15	1,50	30,1
8	1,45	29,3	16	1,49	31,6

9. Провести корреляционный и регрессионный анализ для выборочной совокупности таблицы 10, в которой представлены результаты определения плотности сложения почвы (г/см^3) и полной влагоемкости (%) почв. По уравнению регрессии рассчитайте полной влагоемкости (%) при плотности сложения почвы 0,65 и $1,50(\text{г/см}^3)$. Сделать выводы.

Таблица 10

Таблица показателей опыта

Номер пары	Плотность сложения, г/см^3	Полная влагоемкость, %	Номер пары	Плотность сложения, г/см^3	Полная влагоемкость, %
1	1,22	45,9	9	1,37	30,4
2	1,53	27,5	10	1,59	30,1
3	1,59	28,2	11	1,79	18,0
4	1,42	36,8	12	1,75	16,9
5	1,51	35,2	13	1,20	40,7
6	0,88	72,4	14	1,71	21,4
7	1,37	30,4	15	1,73	19,5
8	1,59	30,1	16	1,79	17,7

10. Провести корреляционный и регрессионный анализ для выборочной совокупности таблицы 11, в которой представлены результаты определения плотности сложения почвы (г/см^3) и капиллярной влагоемкости (%) почв. По уравнению регрессии рассчитайте полной влагоемкости (%) при плотности сложения почвы 0,75 и $1,51(\text{г/см}^3)$. Сделать выводы.

Таблица 11

Таблица показателей опыта

Номер пары	Плотность сложения, г/см^3	Капиллярная влагоемкость, %	Номер пары	Плотность сложения, г/см^3	Капиллярная влагоемкость, %
1	1,07	45,3	9	1,49	27,1
2	1,33	31,8	10	1,36	33,2
3	1,38	27,5	11	1,11	51,4
4	1,42	29,0	12	1,49	27,2
5	1,29	36,2	13	1,58	25,5
6	1,21	39,3	14	1,56	22,7
7	1,40	28,7	15	1,50	25,4
8	1,45	25,2	16	1,22	42,3

11. Провести корреляционный и регрессионный анализ для выборочной совокупности таблицы 12, в которой представлены результаты определения плотности сложения почвы (г/см^3) и капиллярной влагоемкости (%) почв. По

уравнению регрессии рассчитайте полной влагоемкости (%) при плотности сложения почвы 0,75 и 1,50(г/см³). Сделать выводы.

Таблица 12

Таблица показателей опыта

Номер пары	Плотность сложения, г/см ³	Капиллярная влагоемкость, %	Номер пары	Плотность сложения, г/см ³	Капиллярная влагоемкость, %
1	1,22	42,3	9	1,37	28,2
2	1,53	25,6	10	1,59	27,0
3	1,59	24,5	11	1,79	16,6
4	1,42	35,8	12	1,75	15,6
5	1,51	32,1	13	1,20	35,5
6	0,88	64,4	14	1,71	19,6
7	1,37	28,2	15	1,73	18,0
8	1,59	27,0	16	1,79	15,9

12. Провести корреляционный и регрессионный анализ для выборочной совокупности таблицы 13, в которой представлены результаты определения плотности твердой фазы почвы г/см³ и общей порозности %. По уравнению регрессии рассчитайте общую порозность при плотности твердой фазы почвы равной 2,69 и 2,61 г/см³. Сделать выводы.

Таблица 13

Таблица показателей опыта

Номер пары	Плотность тв. фазы почвы, г/см ³	Общая порозность, %
1	2,65	59,6
2	2,74	51,5
3	2,76	50,0
4	2,77	48,7
5	2,75	53,1
6	2,78	46,0
7	2,64	54,1
8	2,78	49,6
9	2,77	47,7
10	2,79	46,6
11	2,80	51,4

13. Провести корреляционный и регрессионный анализ для выборочной совокупности таблицы 14, в которой представлены результаты определения

плотности твердой фазы почвы г/см^3 и общей порозности %. По уравнению регрессии рассчитайте общую порозность при плотности твердой фазы почвы равной 2,85 и 2,81 г/см^3 . Сделать выводы.

Таблица 14

Таблица показателей опыта

Номер пары	Плотность тв.фазы почвы, г/см^3	Общая порозность, %
1	2,64	58,0
2	2,79	46,6
3	2,74	42,3
4	2,84	45,1
5	2,73	45,1
6	2,76	46,7
7	2,64	53,8
8	2,75	44,4
9	2,80	43,2
10	2,80	49,3
11	2,83	46,6

14. Провести корреляционный и регрессионный анализ для выборочной совокупности таблицы 15, в которой представлены результаты определения плотности твердой фазы почвы г/см^3 и общей порозности %.

Таблица 15

Таблица показателей опыта

Номер пары	Плотность тв.фазы почвы, г/см^3	Общая порозность, %
1	2,56	65,6
2	2,67	55,1
3	2,73	49,8
4	2,79	43,0
5	2,76	35,1
6	2,78	37,1
7	2,75	35,6
8	2,65	54,7
9	2,78	38,5
10	2,76	37,3
11	2,79	35,8

По уравнению регрессии рассчитайте общую порозность при плотности твердой фазы почвы равной 2,45 и 2,68 г/см³. Сделать выводы.

15. Провести корреляционный и регрессионный анализ для выборочной совокупности таблицы 16, в которой представлены результаты определения содержания физической глины (%) и содержания водопрочных агрегатов (%) в нижних иллювиальных горизонтах серых лесных почв РТ. По уравнению регрессии рассчитайте содержание водопрочных агрегатов если содержание физической глины будет 30,1 и 45,2%. Сделать выводы.

Таблица 16

Таблица показателей опыта

Номер пары	Содержание фракции менее 0,01мм, %	Содержание водопрочных агрегатов, %	Номер пары	Содержание фракции менее 0,01мм, %	Содержание водопрочных агрегатов, %
1	43,1	49,2	6	50,6	63,2
2	47,0	53,4	7	50,9	78,0
3	37,4	19,6	8	49,3	68,4
4	39,1	13,7	9	36,9	28,6
5	37,8	8,0	10	34,8	30,9

16. Провести корреляционный и регрессионный анализ для выборочной совокупности таблицы 17, в которой представлены результаты определения содержания физической глины (%) и содержания водопрочных агрегатов (%) в нижних иллювиальных горизонтах серых лесных почв РТ. По уравнению регрессии рассчитайте содержание водопрочных агрегатов если содержание физической глины будет 20,0 и 38,5%. Сделать выводы

Таблица 17

Таблица показателей опыта

Номер пары	Содержание фракции менее 0,01мм, %	Содержание водопрочных агрегатов, %	Номер пары	Содержание фракции менее 0,01мм, %	Содержание водопрочных агрегатов, %
1	39,8	50,4	6	42,3	59,8
2	44,5	43,1	7	48,0	64,9
3	44,1	56,0	8	53,5	74,2
4	32,4	10,0	9	46,0	26,7
5	32,5	8,6	10	37,8	10,4

17. Провести корреляционный и регрессионный анализ для выборочной совокупности таблицы 18, в которой представлены результаты определения содержания физической глины (%) и содержания водопрочных агрегатов (%) в нижних иллювиальных горизонтах серых лесных почв РТ. По уравнению регрессии рассчитайте содержание водопрочных агрегатов если содержание физической глины будет 46,9 и 60,3%. Сделать выводы

Таблица 18

Таблица показателей опыта

Номер пары	Содержание фракции менее 0,01мм, %	Содержание водопрочных агрегатов, %
1	44,6	62,2
2	42,8	57,5
3	43,1	53,2
4	35,8	59,0
5	34,0	52,2
6	26,0	48,1

18. Провести корреляционный и регрессионный анализ для выборочной совокупности таблицы 19, в которой представлены результаты определения содержания илистой фракции (%) и содержания водопрочных агрегатов (%) в нижних иллювиальных горизонтах серых лесных почв РТ. По уравнению регрессии рассчитайте содержание водопрочных агрегатов если содержание ила будет 26,1 и 35,6%. Сделать выводы

Таблица 19

Таблица показателей опыта

Номер пары	Содержание фракции менее 0,001мм, %	Содержание водопрочных агрегатов, %	Номер пары	Содержание фракции менее 0,001мм, %	Содержание водопрочных агрегатов, %
1	28,2	49,2	6	32,8	63,2
2	31,9	53,4	7	38,0	78,0
3	29,5	19,6	8	34,4	68,4
4	28,9	13,7	9	24,4	28,6
5	26,6	8,0	10	22,3	30,9

19. Провести корреляционный и регрессионный анализ для выборочной совокупности таблицы 20, в которой представлены результаты определения содержания илистой фракции (%) и содержания водопрочных агрегатов (%) в нижних иллювиальных горизонтах серых лесных почв РТ. По уравнению

регрессии рассчитайте содержание водопрочных агрегатов если содержание ила будет 30,5 и 20,8%. Сделать выводы

Таблица 20

Таблица показателей опыта

Номер пары	Содержание фракции менее 0,001мм, %	Содержание водопрочных агрегатов, %	Номер пары	Содержание фракции менее 0,001мм, %	Содержание водопрочных агрегатов, %
1	23,7	50,4	6	25,1	59,8
2	29,6	43,1	7	34,5	64,9
3	31,9	56,0	8	35,5	74,2
4	23,8	10,0	9	28,2	26,7
5	22,8	8,6	10	23,1	10,4

20. Провести корреляционный и регрессионный анализ для выборочной совокупности таблицы 21, в которой представлены результаты определения содержания илистой фракции (%) и содержания водопрочных агрегатов (%) в нижних иллювиальных горизонтах серых лесных почв РТ. По уравнению регрессии рассчитайте содержание водопрочных агрегатов если содержание ила будет 41,8 и 21,3%. Сделать выводы

Таблица 21

Содержание фракции менее 0,001мм, %	Содержание водопрочных агрегатов, %
30,6	62,2
28,8	57,5
30,4	53,2
22,6	59,0
24,2	52,2
18,6	48,1

Дисперсионный анализ

1. На основании дисперсионного анализа ответьте на вопрос, значительно ли различие в действии форм азотных удобрений на урожай овсяницы луговой (таблица 22) ($\alpha=5\%$)?

Урожай овсяницы (г на сосуд)

Варианты (формы азота)	Урожай, X						Число наблюдений n
1 (контроль)	15.0	16.2	13.4	14.8	-	-	4
2 (сульфат аммония)	28.4	29.4	29.3	27.1	-	-	4
3 (аммиачная селитра)	25.0	28.2	27.7	26.1	25.0	27.1	6
4 (мочевина)	24.3	23.8	25.1	22.2	24.7	23.0	6

2. Провести дисперсионный анализ данных опыта (табл. 23), определить $НСР_{05}$ и сгруппировать сорта по отношению к контролю. На основании дисперсионного анализа ответьте на вопрос, существуют ли значимые различия урожайности сортов озимой пшеницы ($\alpha=5\%$)?

Таблица 23

Урожай озимой пшеницы (ц с 1 га)

Варианты (сорта)	Повторения, X		
	1	2	3
1 (st)	47.8	46.9	45.4
2	53.7	50.3	50.6
3	46.7	42.0	43.4
4	48.0	47.0	45.9
5	41.8	40.0	43.0

3. Провести дисперсионный анализ данных опыта (табл. 24), определить $НСР_{05}$ и сгруппировать сорта по отношению к контролю. На основании дисперсионного анализа ответьте на вопрос, существуют ли значимые различия урожайности сортов сахарной свеклы ($\alpha=5\%$)?

Таблица 24

Урожай корней сахарной свеклы (ц с 1 га).

Варианты (сорта)	Повторения, X		
	1	2	3
1 (st)	360	370	375
2	410	426	423
3	421	422	432
4	216	290	296

4. Провести дисперсионный анализ данных опыта (табл. 25), определить $НСР_{05}$ и сгруппировать сорта по отношению к контролю. На основании дисперсионного анализа ответьте на вопрос, существуют ли значимые различия урожайности сортов сахарной свеклы ($\alpha=5\%$)?

Урожай корней сахарной свеклы (ц с 1 га).

Варианты (сорта)	Повторения, X		
	1	2	3
1 (st)	356	378	396
2	246	290	310
3	369	363	360
4	220	232	260
5	318	350	348

5. Провести дисперсионный анализ данных опыта (табл. 26), определить $НСР_{05}$ и сгруппировать сорта по отношению к контролю. На основании дисперсионного анализа ответьте на вопрос, существуют ли значимые различия урожайности от применения различных доз фосфорных удобрений ($\alpha=5\%$)?

Таблица 26

Действие различных доз фосфорных удобрений на урожай ячменя

Фосфор	Урожай (г на сосуд)
1	24,1 25,8 23,0 27,0
2	28,4 29,7 30,1 27,4
3	28,7 30,4 32,0 27,0
4	30,7 34,4 34,0 31,0
5	46,7 45,4 47,1 46,3
6	59,4 50,7 64,5 60,1

6. В совхозе «Комсомольский» заложили опыт на каштановых почвах влияния различной способа обработки на урожай яровой пшеницы (табл. 27). На основании дисперсионного анализа ответьте на вопрос, существуют ли значимые различия урожайности от способов обработки почв ($\alpha=5\%$)? В качестве контроля – нормальная вспашка.

Таблица 27

Урожай яровой пшеницы на каштановых почвах

Способ обработки; глубина, см	Урожай зерна, ц/га
1 Нормальная вспашка, 0-20 (st)	11,5 11,0 10,8 10,7
2 Дискование, 0-10	4,3 4,8 4,0 4,2
3 Безотвальная вспашка, 0-40	5,8 5,7 5,6 5,9
4 Глубокая вспашка, 0-40	11,4 11,2 11,8 11,3

7. Провести дисперсионный анализ данных опыта (табл. 28). На основании дисперсионного анализа ответьте на вопрос, существуют ли

значимые различия урожайности хлопчатника при различной заделке пласта смеси люцерны и ежи сборной ($\alpha=5\%$)?

Таблица 28

Влияние глубины заделки пласта на урожайность хлопчатника

Глубина заделки пласта трав, см	Урожай, ц/га
25 (1)	38,4 37,5 38,1 38,3
35 (2)	42,3 42,0 42,1 41,9
45 (3)	44,5 44,9 45,2 45,2

8. Провести дисперсионный анализ данных опыта (табл. 29). На основании дисперсионного анализа ответьте на вопрос, влияют ли применение удобрений в песчаной почве на урожай яровой пшеницы ($\alpha=5\%$)?

Таблица 29

Влияние разных удобрений на урожай яровой пшеницы, ц/га

Удобрения	Количество, тыс. г
1 Контроль	2,7 2,1 2,2 2,8
2 NPK	1,4 1,1 1,0 1,6
3 Известь	5,4 5,6 5,2 5,0

9. Провести дисперсионный анализ данных опыта (табл. 30). На основании дисперсионного анализа ответьте на вопрос, существуют ли значимые различия урожайности с/х культур при использовании различных удобрений ($\alpha=5\%$)?

Таблица 30

Действие разных удобрений на урожай сельскохозяйственных культур

Удобрения	Относительная величина урожая
1 Контроль	100,0 110,0 105,1 99,2
2 Vac. megaterium	110,5 119,8 105,4 105,0
3 P ₂ O ₅	144,8 139,5 156,8 142,3
4 P ₂ O ₅ + Vac. Megaterium	163,0 165,5 171,3 150,1

10. Провести дисперсионный анализ данных опыта (табл.31). На основании дисперсионного анализа ответьте на вопрос, существуют ли значимые различия урожайности с/х культур при использовании различных удобрений ($\alpha=5\%$)?

Таблица 31

Действие разных удобрений на урожай сельскохозяйственных культур

Удобрения	Относительная величина урожая
1 Контроль	100,0 110,0 105,1 99,2
2 P ₂ O ₅	144,8 139,5 156,8 142,3
3 Навоз	189,5 180,1 175,3 185,5

11. В двухфакторном опыте, поставленном в четырех рендомизированных повторениях, изучено действие трех градаций орошения (0 - без орошения, 1- умеренное орошение и 2 – обильное орошение) и трех доз азота (0- без азота, 1 – 60, 2 - 120 функтов на акр) на урожай семян хлопчатника (табл. 32). Провести дисперсионный анализ данных.

Таблица 32

Влияние орошения и доз азота на урожай семян хлопчатника
(ц с 1 акра)

Орошение А	Дозы азота В	Повторения, X			
		1	2	3	4
0	0	19	20	15	15
	1	20	20	20	18
	2	18	20	18	18
1	0	32	29	18	21
	1	40	39	33	34
	2	39	38	40	37
2	0	30	31	21	17
	1	42	35	28	33
	2	38	38	36	35

12. В двухфакторном опыте, поставленном в четырех рендомизированных повторениях, изучено действие трех градаций орошения (0 - без орошения, 1- умеренное орошение и 2 – обильное орошение) и трех доз азота (0- без азота, 1 – 120, 2 - 240 функтов на акр) на урожай семян хлопчатника (табл. 33). Провести дисперсионный анализ данных.

Влияние орошения и доз азота на урожай семян хлопчатника
(ц с 1 акра)

Орошение А	Дозы азота В	Повторения, X			
		1	2	3	4
0	0	19	20	15	15
	1	18	20	18	18
	2	20	19	18	19
1	0	32	29	18	21
	1	39	38	40	38
	2	44	42	40	39
2	0	30	31	21	17
	1	38	38	36	35
	2	48	51	50	48

13. В двухфакторном опыте с почвенной культурой ячменя изучено действие двух доз азота и двух доз фосфора (табл. 34). Провести дисперсионный анализ результатов опыта.

Таблица 34

Урожай зерна ячменя в двухфакторном опыте (г на сосуд)

Азот А	Фосфор В	Урожай			
0	0	24,1	25,8	23,0	27,0
	1	28,4	29,7	30,1	27,4
1	0	30,7	34,4	34,0	31,0
	1	46,7	45,4	47,1	46,3

14. В двухфакторном опыте с почвенной культурой ячменя изучено действие двух доз азота и двух доз фосфора (табл. 35). Провести дисперсионный анализ результатов опыта.

Таблица 35

Урожай зерна ячменя в двухфакторном опыте (г на сосуд)

Азот А	Фосфор В	Урожай			
0	0	24,1	25,8	23,0	27,0
	1	28,7	30,4	32,0	27,0
1	0	30,7	34,4	34,0	31,0
	1	59,4	50,7	64,5	60,1

15. В двухфакторном опыте, поставленном в четырех рендомизированных повторениях, изучено действие трех градаций орошения (о - без орошения, 1- умеренное орошение) и трех доз азота (0- без азота, 1 – 60, 2 – 120, 3 - 240

функтов на акр) на урожай семян хлопчатника (табл. 36). Провести дисперсионный анализ данных.

Таблица 36

Влияние орошения и доз азота на урожай семян хлопчатника
(ц с 1 акра)

Орошение А	Дозы азота В	Повторения, Х			
		1	2	3	4
0	0	19	20	15	15
	1	20	20	20	18
	2	18	20	18	18
	3	20	19	18	19
1	0	32	29	18	21
	1	40	39	33	34
	2	39	38	40	38
	3	44	42	40	39

16. В двухфакторном полевом опыте исследовали урожай яровой пшеницы на черноземах Курганской области при разной обработке почвы (табл. 37). На основании ДА ответьте на вопрос, существуют ли значимые различия урожайности в курганской опытной станции и колхоза «Заветы Ильича» при различных способах обработки почвы.

Таблица 37

Урожай яровой пшеницы, ц/га

Способ обработки почвы	Колхозы	Урожай, ц/га
Обычная обработка, 0-25 (1)	Колхоз «Заветы Ильича». (1)	19,6 20,8 21,1 20,5
	Курганская опытная станция (2)	22,5 22,8 21,9 22,1
Безотвальная вспашка, 0-40 (2)	Колхоз «Заветы Ильича». (1)	21,7 21,8 20,9 21,5
	Курганская опытная станция (2)	25,4 23,7 25,3 24,5

17. В двухфакторном полевом опыте исследовали количество актиномицетов в зависимости от способов возделывания и применения удобрений (табл. 38). Для микробиологического исследования брались образцы из делянок: 1 – бессменный пар, 2 – бессменная рожь, 3 – севооборот. На основании ДА ответьте на вопрос, существуют ли значимые различия количества актиномицетов, % на дерново-подзолистой почве.

Таблица 38

Влияния различных удобрений на микрофлору дерново-подзолистой почвы
(средние данные за лето).

Севооборот	Внесенные удобрения	Актиномицеты, %
1 Бессменный пар	1- не удобренный	19,5 19,7 19,9 19,3
	2- NPK	4,9 4,6 5,0 5,2
	3- Навоз	10,9 10,7 11,0 10,8
2 Бессменная рожь	1 - не удобренный	50,8 51,2 50,7 51,3
	2- NPK	40,5 40,6 40,8 41,0
	3-Навоз	52,0 52,4 52,1 51,9
3 семипольный севооборот	1 - не удобренный	25,4 25,6 25,5 25,3
	2- NPK	20,1 20,3 20,0 20,8
	3-Навоз	51,9 52,0 51,4 51,5

18. В двухфакторном полевом опыте исследовали среднюю урожайность озимой пшеницы при монокультуре и севообороте (табл. 39). На основании ДА ответьте на вопрос, существуют ли значимые различия урожайности озимой пшеницы?

Таблица 39

Урожайность озимой пшеницы на черноземе при монокультуре и севообороте
(ц/га)

Севооборот	Внесенные удобрения	Урожай, ц/га
1 Севооборот	1 без удобрений	16,0 15,5 17,2 16,9
	2 NPK	26,8 27,9 25,8 27,1
2 Монокультура	1 без удобрений	6,9 5,1 7,8 6,8
	2 NPK	15,4 16,9 17,1 14,0

19. В двухфакторном полевом опыте исследовали влияние люцерны и однолетних злаков как предшественников озимой пшеницы при различном фоне азотного питания (табл. 40). На основании ДА ответьте на вопрос, существуют ли значимые различия урожайности озимой пшеницы?

Урожай озимой пшеницы, ц/га

Доза азота под пшеницу, кг/га	Предшественники	Урожай, ц/га
1 без азота	1 люцерна	56,9 55,7 54,6 58,1
	2 злаки	35,3 33,2 36,8 35,8
2 50	1 люцерна	68,8 62,3 69,1 70,1
	2 злаки	49,1 47,2 45,1 51,2
3 100,4	1 люцерна	70,0 72,3 71,0 69,2
	2 злаки	60,5 61,3 62,2 60,1
4 150,6	1 люцерна	68,8 68,9 66,2 69,2
	2 злаки	65,0 62,3 66,2 67,1

20. В двухфакторном полевом опыте исследовали число грибов в зависимости от способов возделывания и применения удобрений (табл. 41). Для микробиологического исследования брались образцы из делянок: 1 – бессменный пар, 2 – бессменная рожь, 3 – семипольный севооборот. На основании ДА ответьте на вопрос, существуют ли значимые различия числа грибов (тыс. на 1 га) на дерново-подзолистой почве.

Таблица 41

Влияния различных удобрений на микрофлору дерново-подзолистой почвы (средние данные за лето).

Севооборот	Внесенные удобрения	Число грибов, тыс на 1 га
1 Пар занятой	1 не удобренный	15,0 14,8 13,2 16,9
	2 NPK	23,6 20,3 21,2 24,9
	3 Навоз	30,0 33,3 31,0 28,5
2 Бессменная рожь	1 не удобренный	29,0 25,4 31,0 30,0
	2 NPK	57,2 55,5 56,9 58,9
	3 Навоз	28,1 26,4 30,1 28,9
3 семипольный севооборот	1 не удобренный	4,0 4,4 5,2 5,5
	2 NPK	44,6 48,9 51,0 45,5
	3 Навоз	72,0 77,7 71,2 70,3

Литература

1. ГОСТ Р ИСО 5479-2002 Статистические методы. Проверка отклонения распределения вероятностей от нормального распределения. – М.: Изд-во стандартов. 2002. 30с.
2. Доспехов Б.А. Методика полевого опыта (с основами статистической обработки результатов исследований). -5-е изд., доп. и перераб. – М.: Агропромиздат, 1985. – 351 с.
3. Компьютерная биология. Ч1. Статистический пакет STATGRAPHICS Plus for Windows для генетиков: Учебно-методическое пособие для студентов биолого-почвенного факультета – Казань: Унипресс, 1999. 40с.
4. Новиков Д.А., Новочадов В.В Статистические методы в медико-биологическом эксперименте (типовые случаи). Волгоград: Изд-во ВолМГУ, 2005. -84с.
5. Хан Г., Шапиро С. Статистические модели в инженерных задачах. М.: Мир, 1969. 395с.
6. Ханс-Петер Пифо. Статистика для бакалавров по специальностям: Агробиология (АБ), Аграрные науки (АН) и Возобновляемые природные ресурсы (ВПР) в Университете Хоэнхайм. 1. семестр (АБ, АН, ВПР), 3. Семестр (ВПР). – М.: Изд-вл ВНИИА, 2011. -288с.