**РОССИЙСКИЙ УНИВЕРСИТЕТ ДРУЖБЫ НАРОДОВ**

**Факультет физико-математических и естественных наук**

**Кафедра прикладной информатики и теории вероятностей**

Практикум по математической статистике

Лабораторная работа №3

**Тема: «Дисперсионный анализ»**

Вариант 10

Выполнил

Студент: Феоктистов Владислав

Группа: НПМбд-01-19б

№ c/б: 1032192939

Преподаватель: Матюшенко Сергей Иванович

**МОСКВА**

2022 г.

**Цель работы:** приобрести практические навыки применения дисперсионного анализа для решения конкретных задач с использованием статистического пакета SPSS.

**Ход работы:**

1. Повторил теоретические основы дисперсионного анализа, используя материалы учебного пособия.
2. Разобрал примеры использования SPSS для реализации дисперсионного анализа.
3. В приведенной ниже таблице даны значения урожайности картофеля (тыс. тонн с гектара) в зависимости от сорта картофеля (фактор А) и типа примененного удобрения (фактор В). С помощью двухфакторного дисперсионного анализа без повторных измерений необходимо выяснить:
   1. значимы ли различия в средней урожайности различных сортов картофеля независимо от типа удобрения;
   2. значимо ли влияние типа применяемого удобрения на урожайность независимо от сорта.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **B1** | **B2** | **B3** | **B4** |
| **A1** | 2.609 | 3.055 | 5.013 | 4.53 |
| **A2** | 4.191 | 4.452 | 5.683 | 4.647 |
| **A3** | 6.067 | 5.446 | 3.665 | 5.621 |

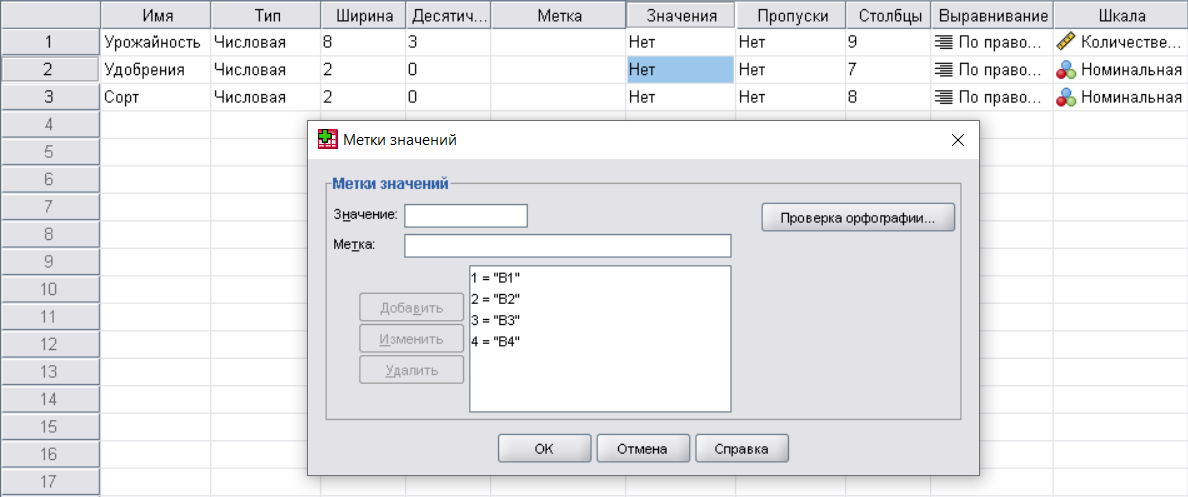
*Решение.* Согласно условию задачи исследуется влияние на урожайность (зависимую переменную) двух факторов – типа удобрений и сорта пшеницы.

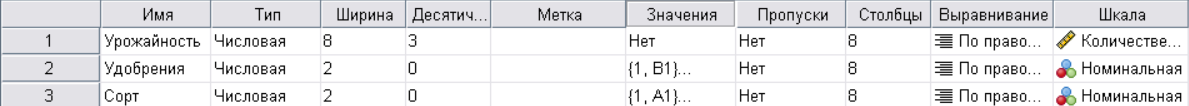
Выдвинем две нулевые гипотезы:

: различия в средней урожайности картофеля, вызванные влиянием типа удобрения (фактора A), выражены не более, чем различия, обусловленные случайными причинами.

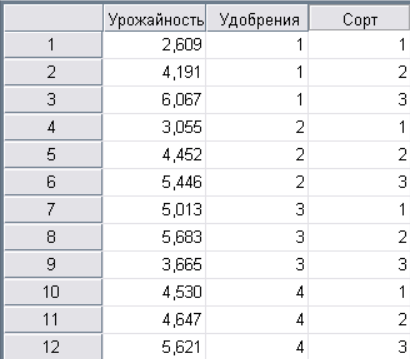
: различия в средней урожайности картофеля, вызванные влиянием сорта (фактора B), выражены не более, чем различия, обусловленные случайными причинами.

Для начала запустим программу SPSS и в окне «Переменные» опишем все исходные данные (для столбцов с номинальной шкалой можно добавить значения для соответствующих уровней).

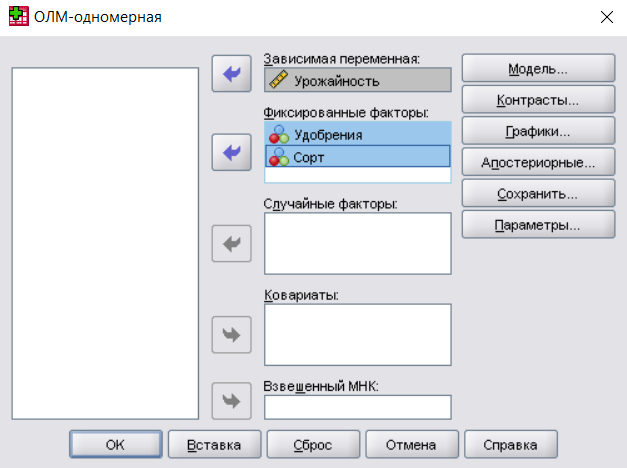


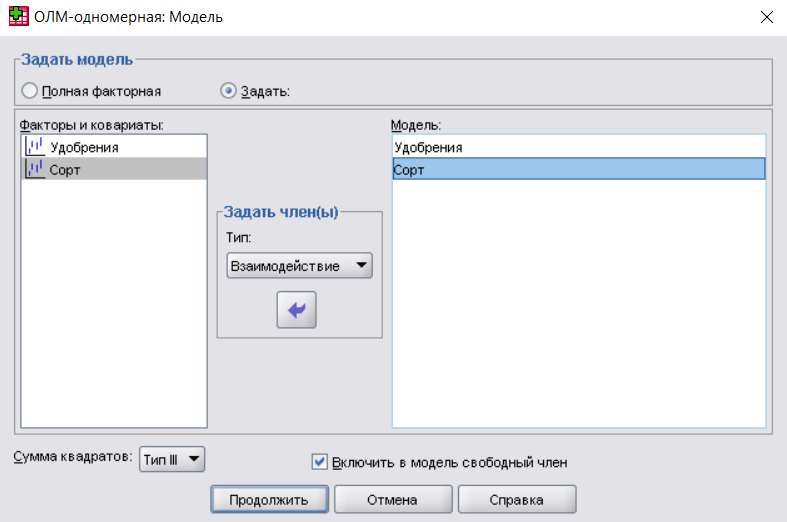


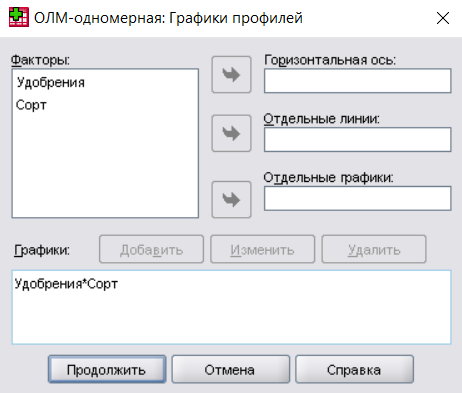
В окне «Данные» введем значения переменных согласно условию задачи, используя вышеприведенную таблицу. Первый столбец исходных данных отведем для значений урожайности картофеля, второй – тип удобрения, третий – для сорта картофеля.

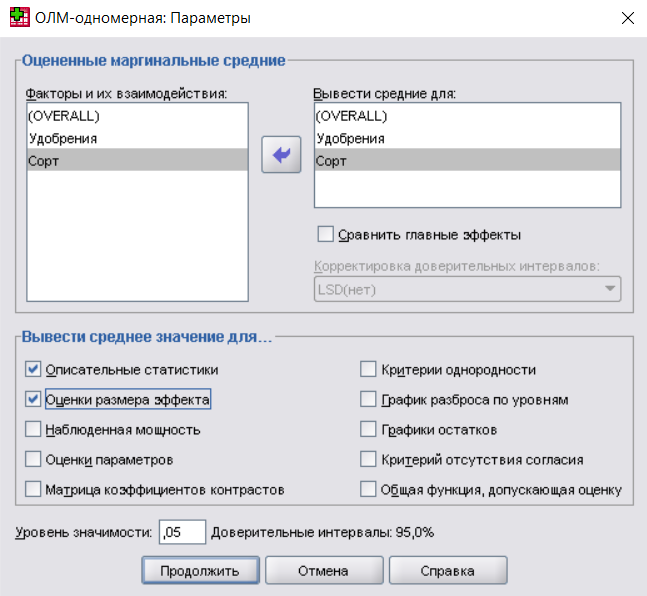


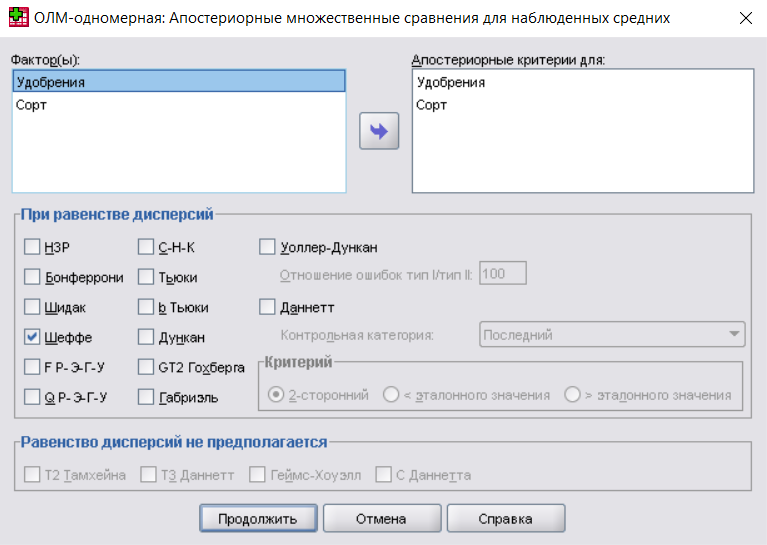
Теперь можно провести двухфакторный дисперсионный анализ без повторных изменений.











После всех предварительных настроек «ОЛМ-одномерная», как показано на рисунках выше, нажимаем на кнопку «ОК» и получаем результаты расчета в виде таблиц и графиков.

Вначале выводится сводная таблица «Межгрупповые факторы», в которой приведены общие сведения об изучаемых факторах, присвоенных метках и о количестве наблюдений (N) по каждому фактору.

**Одномерный дисперсионный анализ**



В таблице «Описательные статистики» содержатся средние значения и стандартные отклонения всех выборок, а также итоговые значения по всем данным.

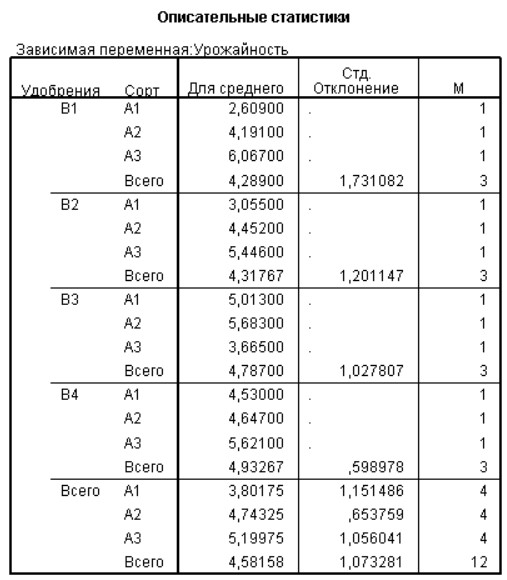


Таблица «Оценки эффектов межгрупповых факторов» содержит результаты проверки основных гипотез двухфакторного дисперсионного анализа.

В данном случае имеем следующее:

1. Переменная «Удобрение» не оказывает статистически значимое влияние на распределение зависимой переменной «Урожайность», поскольку F = 0,252 при Знач. = 0,858 (средние значения урожайности по типам удобрений составили: 4,289; 4,318; 4,787 и 4,933).
2. Переменная «Сорт» также не оказывает статистически значимое влияние на распределение зависимой переменной «Урожайность», поскольку F =1,596, а Знач. = 0,278 (средние значения урожайности по сортам: 3,802; 4,743 и 5,200).

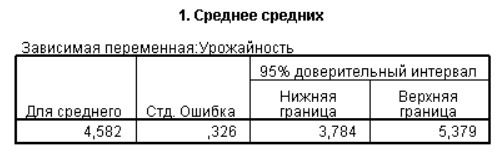
Таким образом, мы останавливаемся на гипотезах и и приходим к выводу, что различия в средней урожайности картофеля, вызванные влиянием типа удобрения (фактор A) и сорта (фактор B), выражены не более, чем различия, обусловленные случайными причинами. Однако, стоит отметить, что влияние фактора сорта сильнее, чем влияние фактора типа удобрения.

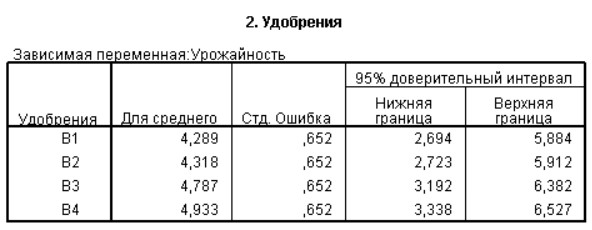
Кроме того, можно заметить, что коэффициент детерминации имеет значение , т.е. учтено влияние только 39,7% факторов, что достаточно мало, поэтому модель является плохой и ее нельзя использовать для дальнейших исследований.

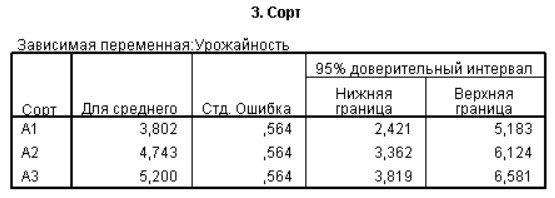


В следующих трех таблицах выводятся описательные статистики для совокупности всех данных и для каждого из исследуемых факторов.

**Оцененные маргинальные средние**



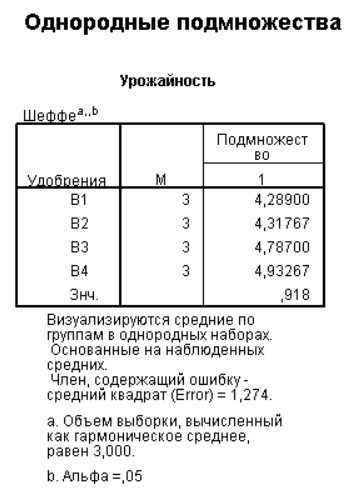
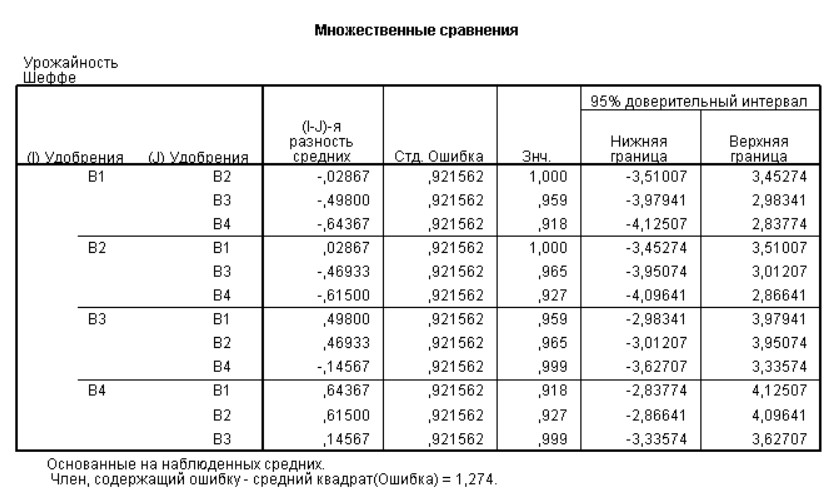




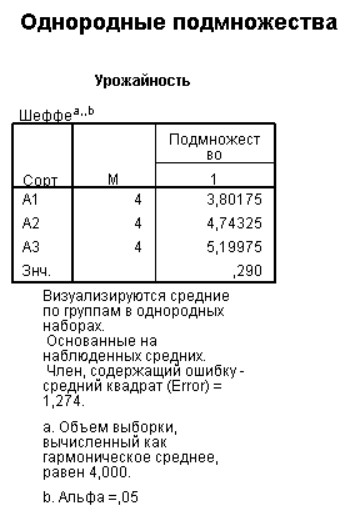
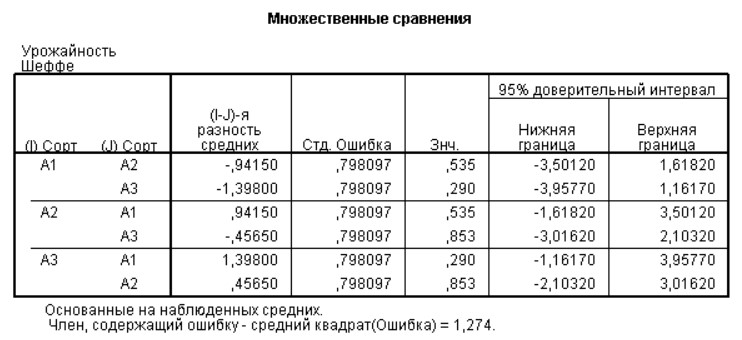
Далее SPSS вывел результаты расчета апостериорных критериев, в частности теста Шеффе, по сравнению отдельных типов удобрений и сортов картофеля. По результатам расчета можно сказать, что удобрения B4 и B3 наиболее значимы по своему воздействию на урожайность картофеля и отличаются от других типов удобрений, и что сорт картофеля A3 является наиболее урожайным, однако, как было сказано ранее, в целом сорт и удобрения не оказывают значимого влияния на урожайность.

**Апостериорные критерии**

**Удобрения**



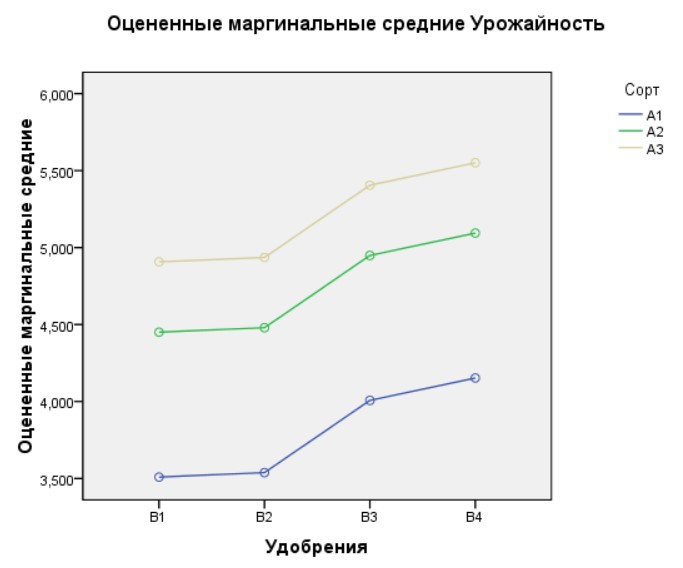
**Сорт**



Вывод результатов заканчивается графиками профилей, на которых представлены зависимости средней урожайности картофеля от типа удобрений и от сорта.

Из данного графика видно, что наиболее эффективно на среднюю урожайность оказывает влияние удобрения B4 и наименее эффективно – удобрение B1. Если рассматривать влияние сорта, можно сказать, что наиболее урожайным является сорт A3 и наименее урожайным – сорт A1. Но, повторяясь, в целом сорт и удобрения не оказывают значимого влияния на урожайность.

**Графики профилей**



Таким образом, в результате решения данной задачи методом дисперсионного анализа, реализованного в алгоритме одномерной линейной модели, нулевые гипотезы и не отвергаются, что говорит об отсутствии существенного влияния названных факторов на урожайность (различия в средней урожайности картофеля, вызванные влиянием типа удобрения (фактор A) и сорта (фактор B), выражены не более, чем различия, обусловленные случайными причинами).

**Вывод:** приобрёл практические навыки применения дисперсионного анализа для решения конкретных задач с использованием статистического пакета SPSS.