

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: Феокистов Владислав

Группа: НПМбд-01-196

№ с/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 A) и типа примененного удобрения (фактор B). С помощью двухфакторного дисперсионного анализа без повторных измерений необходимо выяснить:
 - а. значимы ли различия в средней урожайности различных сортов картофеля независимо от типа удобрения;
 - б. значимо ли влияние типа применяемого удобрения на урожайность независимо от сорта.

	B ₁	B ₂	B ₃	B ₄
A ₁	2.609	3.055	5.013	4.53
A ₂	4.191	4.452	5.683	4.647
A ₃	6.067	5.446	3.665	5.621

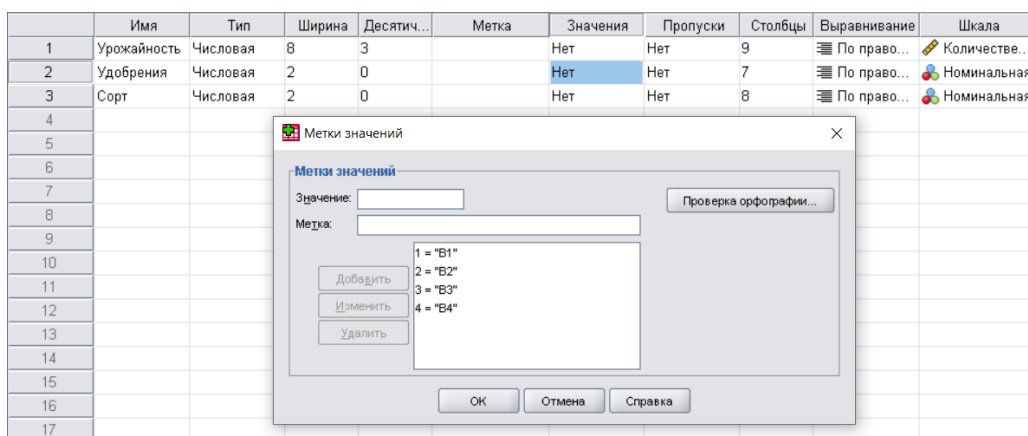
Решение. Согласно условию задачи исследуется влияние на урожайность (зависимую переменную) двух факторов – типа удобрений и сорта пшеницы.

Выдвинем две нулевые гипотезы:

$H_{0,A}$: различия в средней урожайности картофеля, вызванные влиянием типа удобрения (фактора A), выражены не более, чем различия, обусловленные случайными причинами.

$H_{0,B}$: 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).



	Имя	Тип	Ширина	Десятич...	Метка	Значения	Пропуски	Столбцы	Выравнивание	Шкала
1	Урожайность	Числовая	8	3		Нет	Нет	8	По право...	Количество...
2	Удобрения	Числовая	2	0		{1, B1}...	Нет	8	По право...	Номинальная
3	Сорт	Числовая	2	0		{1, A1}...	Нет	8	По право...	Номинальная

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.

	Урожайность	Удобрения	Сорт
1	2,609	1	1
2	4,191	1	2
3	6,067	1	3
4	3,055	2	1
5	4,452	2	2
6	5,446	2	3
7	5,013	3	1
8	5,683	3	2
9	3,665	3	3
10	4,530	4	1
11	4,647	4	2
12	5,621	4	3

Теперь можно провести двухфакторный дисперсионный анализ без повторных изменений.

ОЛМ-одномерная

Зависимая переменная: Урожайность

Фиксированные факторы: Удобрения, Сорт

Случайные факторы:

Ковариаты:

Взвешенный МНК:

Модель...
Контрасты...
Графики...
Апостериорные...
Сохранить...
Параметры...

OK Вставка Сброс Отмена Справка

ОЛМ-одномерная: Модель

Задать модель

☐ Полная факторная ☒ Задать:

Факторы и ковариаты: Удобрения, Сорт

Модель: Удобрения, Сорт

Задать член(ы)

Тип: Взаимодействие

Сумма квадратов: Тип II ☒ Включить в модель свободный член

Продолжить Отмена Справка

ОЛМ-одномерная: Графики профилей

Факторы:
Удобрения
Сорт

Горизонтальная ось:

Отдельные линии:

Отдельные графики:

Графики: Добавить Изменить Удалить

Удобрения*Сорт

Продолжить Отмена Справка

ОЛМ-одномерная: Параметры

Оцененные маргинальные средние

Факторы и их взаимодействия:
(OVERALL)
Удобрения
Сорт

Вывести средние для:
(OVERALL)
Удобрения
Сорт

☐ Сравнить главные эффекты

Корректировка доверительных интервалов:
LSD(нет)

Вывести среднее значение для...

☒ Описательные статистики
☒ Оценки размера эффекта
☐ Наблюдённая мощность
☐ Оценки параметров
☐ Матрица коэффициентов контрастов

☐ Критерии однородности
☐ График разброса по уровням
☐ Графики остатков
☐ Критерий отсутствия согласия
☐ Общая функция, допускающая оценку

Уровень значимости: ,05 Доверительные интервалы: 95,0%

Продолжить Отмена Справка

ОЛМ-одномерная: Апостериорные множественные сравнения для наблюдаемых средних

Фактор(ы):
Удобрения
Сорт

Апостериорные критерии для:
Удобрения
Сорт

При равенстве дисперсий

☐ НЗР ☐ С-Н-К ☐ Уоллер-Дункан
☐ Бонферрони ☐ Тьюки ☐ Отношение ошибок тип I/тип II: 100
☐ Шидак ☐ b Тьюки ☐ Даннетт
☒ Шеффе ☐ Дункан ☐ Контрольная категория: Последний
☐ F P-3-G-Y ☐ GT2 Гохберга
☐ Q P-3-G-Y ☐ Габриэль Критерий
☒ 2-сторонний ☐ < эталонного значения ☐ > эталонного значения

Равенство дисперсий не предполагается

☐ T2 Дамхейна ☐ T3 Даннетт ☐ Геймс-Хоуэлл ☐ С Даннетта

Продолжить Отмена Справка

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

Межгрупповые факторы

		Метка значения	М
Удобрения	1	B1	3
	2	B2	3
	3	B3	3
	4	B4	3
Сорт	1	A1	4
	2	A2	4
	3	A3	4

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

Описательные статистики

Зависимая переменная: Урожайность

Удобрения	Сорт	Для среднего	Стд. Отклонение	М
B1	A1	2,60900	.	1
	A2	4,19100	.	1
	A3	6,06700	.	1
	Всего	4,28900	1,731082	3
B2	A1	3,05500	.	1
	A2	4,45200	.	1
	A3	5,44600	.	1
	Всего	4,31767	1,201147	3
B3	A1	5,01300	.	1
	A2	5,68300	.	1
	A3	3,66500	.	1
	Всего	4,78700	1,027807	3
B4	A1	4,53000	.	1
	A2	4,64700	.	1
	A3	5,62100	.	1
	Всего	4,93267	,598978	3
Всего	A1	3,80175	1,151486	4
	A2	4,74325	,653759	4
	A3	5,19975	1,056041	4
	Всего	4,58158	1,073281	12

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 $\text{Знач.} = 0,858$ (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 $\text{Знач.} = 0,278$ (средние значения урожайности по сортам: 3,802; 4,743 и 5,200).

Таким образом, мы останавливаемся на гипотезах $H_{0,A}$ и $H_{0,B}$ и приходим к выводу, что различия в средней урожайности картофеля, вызванные влиянием типа удобрения (фактор А) и сорта (фактор В), выражены не более, чем различия,

обусловленные случайными причинами, поэтому говорить о том какие сорта и удобрения лучше не имеет смысла.

Кроме того, можно заметить, что коэффициент детерминации имеет значение $R^2 = 0,397$, т.е. учтено влияние только 39,7% факторов, что достаточно мало, поэтому модель является плохой и ее нельзя использовать для дальнейших исследований.

Оценка эффектов межгрупповых факторов

Зависимая переменная: Урожайность

Источник	Сумма квадратов типа III	ст. св.	Средний квадрат	Щ	Знач.	Частная Эта в Квадрате
Скорректированная модель	5,028 ^a	5	1,006	,789	,593	,397
Свободный член	251,891	1	251,891	197,730	,000	,971
Удобрения	,962	3	,321	,252	,858	,112
Сорт	4,066	2	2,033	1,596	,278	,347
Ошибка	7,643	6	1,274			
Всего	264,562	12				
Скорректированный итог	12,671	11				

a. R квадрат = ,397 (Скорректированный R квадрат = -,106)

Так как факторы «сорт» и «удобрения» не оказывают влияние на «урожайность», то выведенные в SPSS далее таблицы и графики различных оценок для анализа не имеют смысла, поэтому их вывод был исключен.

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