

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Тульский государственный университет»

Институт прикладной математики и компьютерных наук
Кафедра «Прикладная математика и информатика»

Утверждено на заседании кафедры
«Прикладная математика и информатика»
14 января 2020 г., протокол № 6

Заведующий кафедрой

 В.И. Иванов

МЕТОДИЧЕСКИЕ УКАЗАНИЯ
по выполнению лабораторных работ
по дисциплине (модулю)
«Теория вероятностей и математическая статистика»

основной профессиональной образовательной программы
высшего образования – программы бакалавриата

по направлению подготовки
01.03.03 Механика и математическое моделирование

с направленностью (профилем)
Механика деформируемого твёрдого тела

Форма обучения: очная

Идентификационный номер образовательной программы: 010303-01-20

Тула 2020 год

Разработчик методических указаний

Кочетыгов А.А., профессор каф. ПМиИ, к.т.н., доцент
(ФИО, должность, ученая степень, ученое звание)


(подпись)

Лабораторные работы по дисциплине «Теория вероятностей и математическая статистика» выполняются в 6–м семестре и имеют целью закрепление знаний по 3–му разделу (части) курса «Математическая статистика».

Предлагаются индивидуальные задания по обработке данных на ЭВМ с использованием пакета SPSS.

Для успешного выполнения ККР целесообразно использовать лекции и следующую литературу:

1. Кочетыгов А.А. Теория вероятностей: учеб. пособие / – Тула, Изд–во ТулГУ, 2016. – 234 с.

2. Кочетыгов А.А. Математическая статистика: учеб. пособие / – Тула, Изд–во ТулГУ, 2017. – 274 с.

3. Кочетыгов А.А. Математическая статистика. Решение задач с использованием пакета SPSS: учеб. пособие / Тула: Изд–во ТулГУ, 2011. – 156 с.

4. Кочетыгов А.А. Случайные процессы: учеб. пособие / – Тула, Изд–во ТулГУ, 2020. – 300 с.

СОДЕРЖАНИЕ

1. ВОЗМОЖНОСТИ И ОРГАНИЗАЦИЯ ПАКЕТА SPSS	4
1.1. ОСОБЕННОСТИ ПОЛЬЗОВАТЕЛЬСКОГО ИНТЕРФЕЙСА	4
1.2. БАЗОВЫЙ МОДУЛЬ (SPSS BASE).....	20
1.3. ДОПОЛНИТЕЛЬНЫЕ МОДУЛИ	20
1.4. ВВОД ДАННЫХ (DATA ENTRY)	23
1.5. ПРЕДСТАВЛЕНИЕ РЕЗУЛЬТАТОВ АНАЛИЗА	23
2. ПРАКТИКУМ ОБРАБОТКИ ДАННЫХ НА ЭВМ.....	25
2.1. ОПИСАНИЕ ПОКАЗАТЕЛЕЙ ПО ВЫБОРКЕ.....	25
2.2. ИНТЕРВАЛЬНЫЕ ОЦЕНКИ ПАРАМЕТРОВ.....	30
2.3. ПРОВЕРКА ГИПОТЕЗ ПО СТАТИСТИЧЕСКИМ ДАННЫМ.....	33
2.4. КОРРЕЛЯЦИОННЫЙ И РЕГРЕССИОННЫЙ АНАЛИЗЫ.....	36
2.5. ДИСПЕРСИОННЫЙ АНАЛИЗ	43
2.6. РЯДЫ ДИНАМИКИ	48
2.7. ИНДЕКСНЫЙ МЕТОД.....	53
2.8. КЛАСТЕРНЫЙ АНАЛИЗ.....	57
2.9. ДВУХФАКТОРНЫЙ ДИСПЕРСИОННЫЙ АНАЛИЗ.....	63
2.10. МНОГОМЕРНЫЙ АНАЛИЗ.....	67
2.11. МНОЖЕСТВЕННАЯ ЛИНЕЙНАЯ РЕГРЕССИЯ.....	69
2.12. ДИСКРИМИНАНТНЫЙ АНАЛИЗ	78
3. ИНДИВИДУАЛЬНЫЕ ЗАДАНИЯ ДЛЯ ИССЛЕДОВАНИЙ	82

1. ВОЗМОЖНОСТИ И ОРГАНИЗАЦИЯ ПАКЕТА SPSS

SPSS для Windows – пакет для статистического анализа данных и управления данными.

В отличие от других средств анализа данных, SPSS охватывает весь аналитический процесс, проще в использовании и имеет относительно невысокую общую стоимость владения.

Программный комплекс SPSS развивается уже на протяжении более 40 лет и предоставляет широкие возможности не только в сфере психологии, социологии, биологии, медицины, в области маркетинговых исследований, в управлении качеством продукции, но и в различных технических приложениях.

SPSS является самой распространённой программой для обработки статистической информации.

Наряду с различными специальными математическими процедурами анализа данных, пакет отличает удобство в работе и дружелюбный пользовательский интерфейс.

Версия 13.0 пакета ориентирована на работу в среде Windows.

1.1. Особенности пользовательского интерфейса

В этой главе представлен обзор использования SPSS для Windows, который в первую очередь будет касаться технических приемов работы с программой.

При каждом запуске программы появляется диалоговое окно (рис. 1.1.1), в котором можно выбрать необходимую операцию.

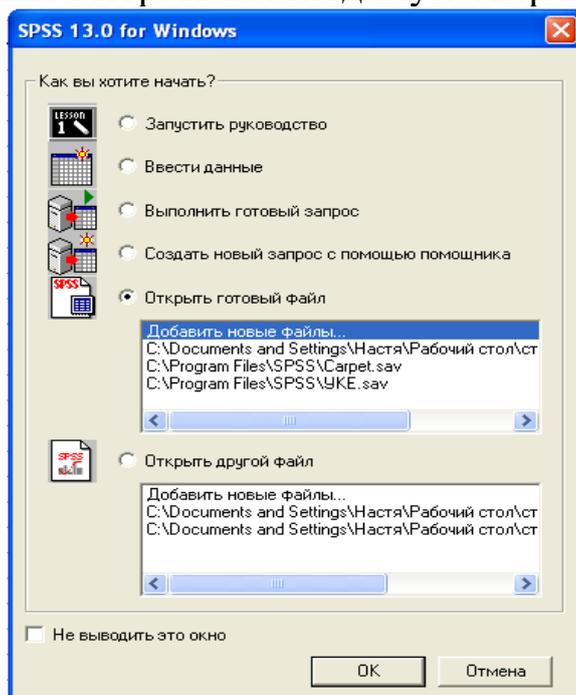


Рис. 1.1.1. Диалоговое окно *Open File*

Подтвердить выбор можно, щелкнув на кнопке ОК.

Подтвердить выбор файла, который необходимо открыть также можно, дважды щелкнув мышью на имя нужного файла (все файлы в SPSS сохраняются с расширением *.sav).

После этого содержимое файла *.sav отобразится в окне редактора данных, как показано на рис. 1.1.2.

	x	xlw	F	var	var	var	var
1	1,00	1,00	4,70				
2	1,00	2,00	9,30				
3	2,00	3,00	16,30				
4	2,00	4,00	30,20				
5	3,00	5,00	32,60				
6	3,00	6,00	53,50				
7	3,00	7,00	72,10				
8	4,00	8,00	79,10				
9	4,00	9,00	90,70				
10	4,00	10,00	95,30				
11	4,00	11,00	97,70				
12	4,00	12,00	100,00				
13	4,00	.	.				
14	5,00	.	.				
15	6,00	.	.				

Рис. 1.1.2. Редактор данных

Если был активизирован просмотр переменных, потребуется еще перейти на вкладку *Data View* (Просмотр данных).

1.1.1. Выбор статистической процедуры

Меню статистики, которое открывается по команде *Analyze* (Анализ), содержит список статистических методов.

После каждого пункта этого меню стоит стрелка. Она указывает на существование следующего уровня меню.

Доступный набор статистических методов зависит, в частности, от того, какие модули были установлены (рис. 1.1.3).

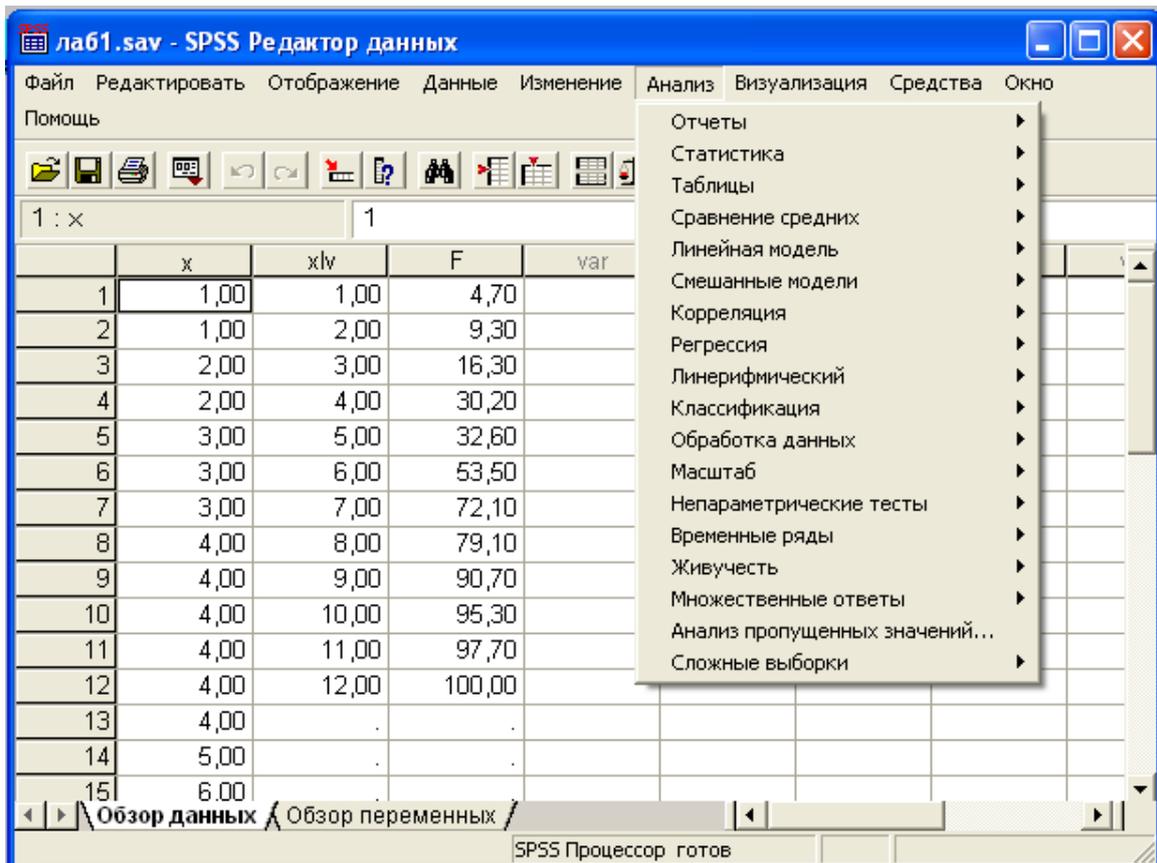


Рис. 1.1.3. Меню статистики

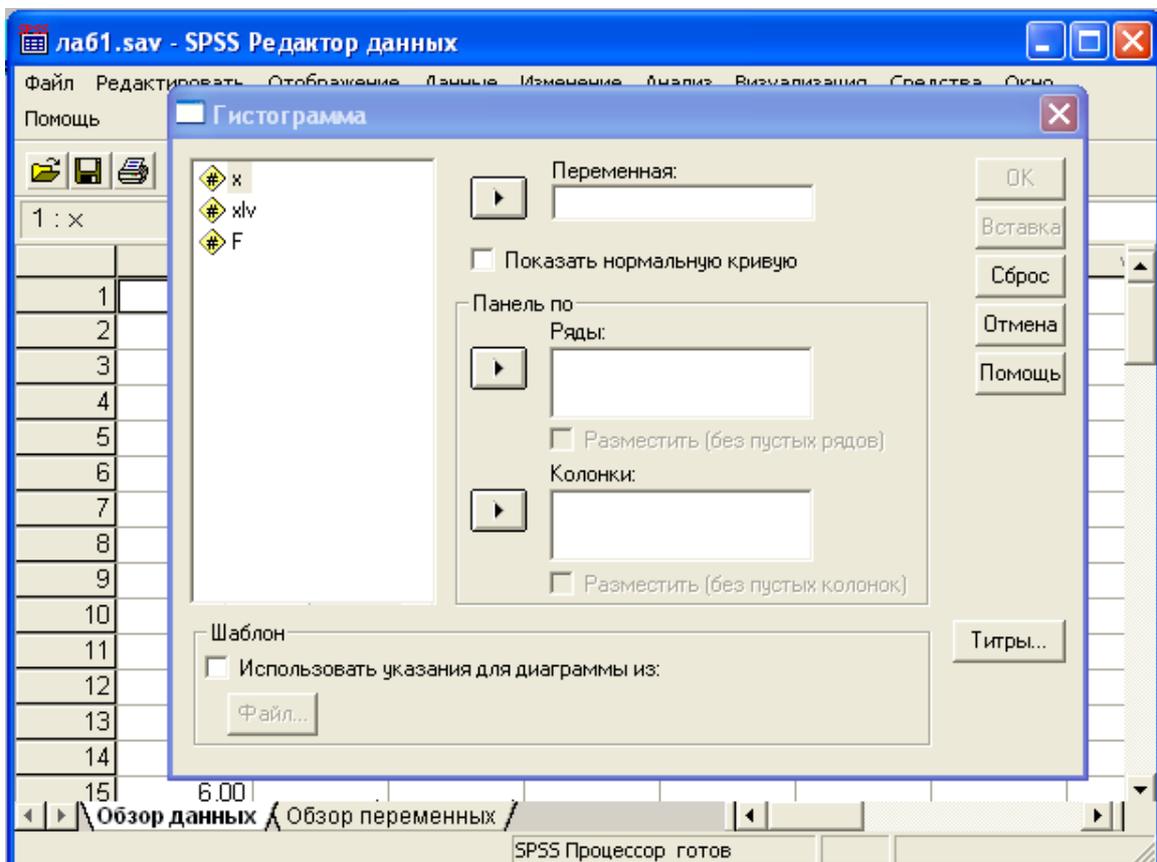


Рис. 1.1.4. Пример диалогового окна

При выборе какого–нибудь из пунктов обычно появляются диалоговые окна, которые, как правило, содержат следующие компоненты:

1. Список исходных переменных – список всех переменных в файле данных. В данный момент в списке исходных переменных присутствуют следующие переменные: *F*, *x*, *xlV*.

Перед именем каждой переменной стоит значок, по которому можно определить, является ли эта переменная численной или строковой.

2. Список выбранных переменных – список, содержащий переменные файла данных, которые были выбраны для анализа.

Список выбранных переменных также называют целевым списком или списком тестируемых переменных.

Этот список имеет заголовок *Variable(s)* (Переменная(ые)).

3. Командные кнопки – кнопки, при щелчке на которые выполняются определенные действия.

В этом диалоговом окне расположены кнопки *OK*, *Paste* (Вставить), *Reset* (Сброс или Отклонить), *Cancel* (Отмена) и *Help* (Справка), а также кнопки, открывающие вспомогательные диалоговые окна: *Statistics...* (Статистика), *Charts...* (Диаграммы или Графики) и *Format...* (Формат).

Кнопки вспомогательных диалоговых окон отличаются троеточием (...) после названия.

Пять стандартных командных кнопок в главном диалоговом окне имеют следующее назначение:

1. *OK* – кнопка *OK* запускает соответствующую процедуру. Одновременно она закрывает диалоговое окно.

2. *Paste* – эта кнопка переносит выбранный в диалоговом окне синтаксис команды в редактор синтаксиса.

Здесь можно отредактировать синтаксис команды и дополнить его другими опциями, недоступными в данном диалоговом окне.

3. *Reset* – эта кнопка отменяет перенос выделенной переменной в целевой список переменных.

4. *Cancel* – эта кнопка отменяет все изменения, сделанные с момента последнего открытия диалогового окна, и закрывает его.

5. *Help* – эта кнопка выводит контекстно–чувствительную справку.

При щелчке на ней открывается окно справки, содержащее сведения о текущем диалоговом окне.

Выбор переменных.

Для выбора переменной в списке исходных данных необходимо выделить нужную переменную, затем щелкнуть на кнопке, которая находится рядом со списком выбранных переменных.

Переменная будет перенесена из списка исходных переменных в список выбранных переменных.

Можно также дважды щелкнуть по нужной переменной, и она будет перенесена в список выбранных переменных.

Подтвердить операцию можно, щелкнув по кнопке ОК.

Результаты будут отображены в окне просмотра (*Viewer*).

Окно просмотра разделено на две части.

В левой отображается структура вывода, а в правой – собственно выводимые данные.

В разделе вывода отображаются как таблицы, так и графики.

Для возврата в редактор данных необходимо выбрать в меню команды *Window* (Окно) **.sav – SPSS Data Editor* или щелкнуть на панели инструментов на символе редактора данных .

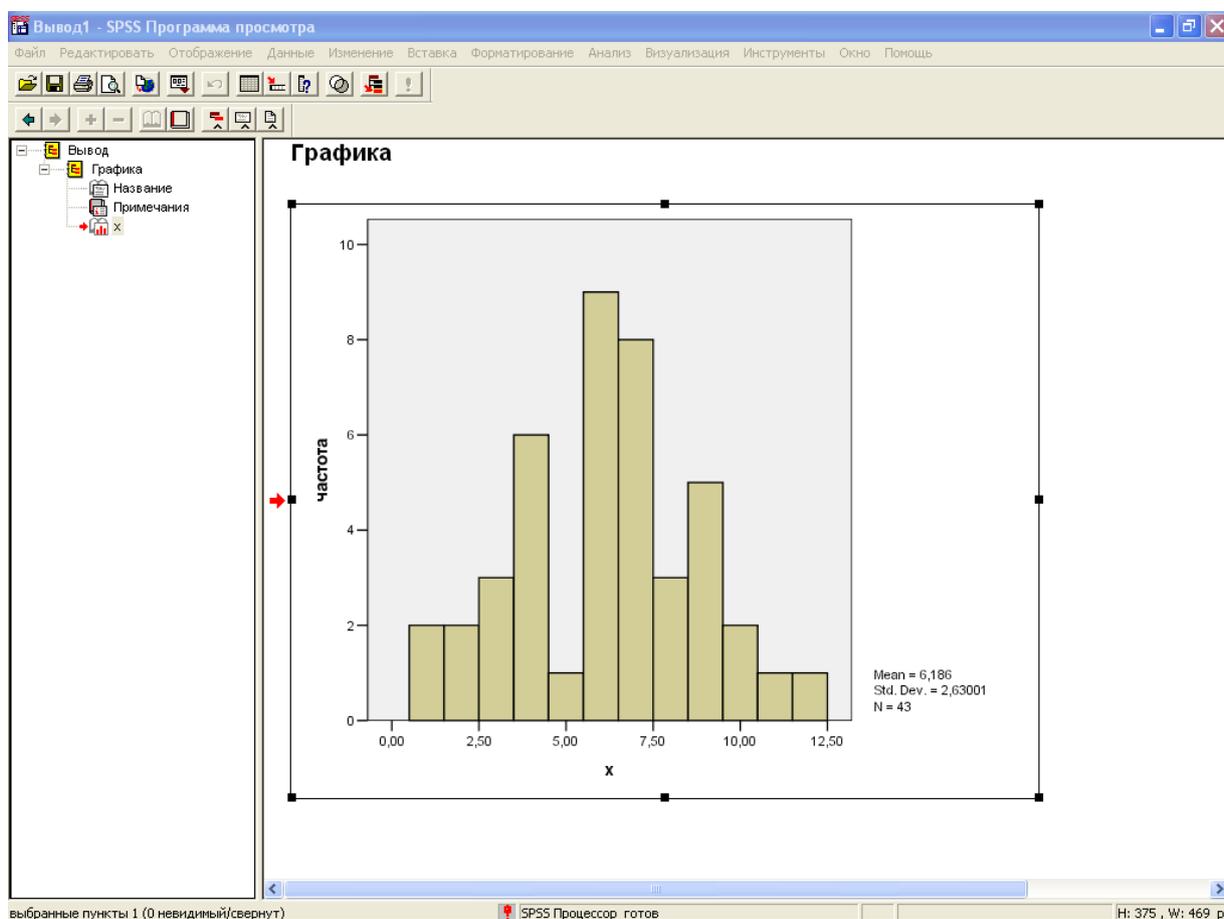


Рис. 1.1.5. Окно просмотра

1.1.2. Настройки редактора данных

Меню *View* (Вид) редактора данных содержит множество опций, с помощью которых можно произвести индивидуальную настройку редактора данных.

В частности, можно:

- показать или скрыть строку состояния. Команда: *Status bar* (Строка состояния);

- увеличить значки на панели символов и включить или отключить отображение кратких сведений. Команда: *Toolbars* (Панели символов);
- выбрать другой тип, начертание и размер шрифта. Команда: *Fonts* (Шрифты);
- включить или отключить отображение линий сетки. Команда: *Grid lines* (Линии сетки);
- отображать метки значений вместо фактических значений переменных. Команда: *Value labels* (Метки значений).

1.1.3. Панели символов

SPSS имеет следующие окна:

- Редактор данных (*Data Editor*).
- Окно просмотра (*Viewer*).
- Окно просмотра текста (*Text Viewer*).
- Редактор мобильных таблиц (*Pivot Table Editor*).
- Редактор диаграмм (*Diagram Editor*).
- Редактор текстового вывода (*Text Output Editor*).
- Редактор синтаксиса (*Syntax Editor*).
- Редактор скриптов (*Script Editor*).

Каждое окно, кроме редактора мобильных таблиц, имеет одну или две панели символов для вызова часто используемых команд.

Краткие сведения о каждом символе можно получить, если поместить на него указатель мыши.

Ниже представлены прежде всего те символы, которые встречаются в большинстве или во всех окнах.

 Открыть файл. Этот символ активизирует диалоговое окно открытия файла, причем по умолчанию предлагается открыть документ того же типа, который находится в активном окне. Следовательно, при помощи этого символа можно открыть файл данных, файл вывода или файл синтаксиса.

 Сохранить файл. Этот символ предназначен для сохранения рабочего файла. Если рабочему файлу еще не присвоено имя, этот символ активизирует диалоговое *Save As* (Сохранить как). Если вы находитесь не в окне редактора данных, активизируется диалоговое окно сохранения файла соответствующего типа – файла вывода или синтаксиса.

 Печать. Этот символ вызывает диалоговое окно вывода на печать в соответствии с типом активного окна. Он позволяет напечатать весь документ или только выделенную область.

 История вызова диалоговых окон. Этот символ выводит список 12 последних вызванных диалоговых окон. Это дает возможность быстро

перейти к одному из недавно вызванных диалоговых окон. Окно, вызванное в последнюю очередь, всегда находится в начале списка. Чтобы заново вызвать диалоговое окно, просто можно щелкнуть на соответствующем пункте списка.

 Перейти в редактор данных. Этот символ обеспечивает переход в редактор данных из любого окна.

 Перейти к наблюдению. Этот символ открывает диалоговое окно *Go to case* (Перейти к наблюдению). Его можно использовать для перехода к определенному наблюдению, так в SPSS называется набор значений переменных, набранных в строке редактора данных.

 Выбрать наблюдения. Этот символ открывает диалоговое окно *Select cases* (Выбрать наблюдения). Его можно использовать для отбора наблюдений, для которых выполняется определенное условие.

 Информация о переменных. Этот символ открывает диалоговое окно *Variables*, в котором отображаются описания выделенных переменных.

Из множества символов, которые возникают только в одном определенном окне, мы покажем лишь несколько. О назначении остальных легко можно узнать из кратких сведений (*Quick Info*) по данному символу.

В редакторе синтаксиса большое значение имеет символ *Syntax–Start* (Синтаксис–Начать), в случае если для вызова статистических процедур используется командный синтаксис SPSS.

 Синтаксис–Начать. В окне редакторе синтаксиса этот символ запускает на выполнение выделенные команды SPSS. Если не выделено ни одной команды, запускается команда, на которой находится курсор.

Три следующих символа могут быть задействованы в редакторе данных.

 Вставить наблюдение. В редакторе данных щелчок на этом символе вызывает вставку наблюдения над активной ячейкой.

 Вставить переменную. В редакторе данных щелчок на этом символе вызывает вставку новой переменной слева от активной переменной.

 Метки значений. Этот символ позволяет переключаться между отображением значений и меток значений.

1.1.4. Построение и редактирование графиков

Иногда возникает необходимость редактирования построенного графика в соответствии со своими требованиями.

Щёлкните дважды на какой-либо точке в пределах графика.

После этого он будет помещён в редактор диаграмм (рис. 1.1.6).

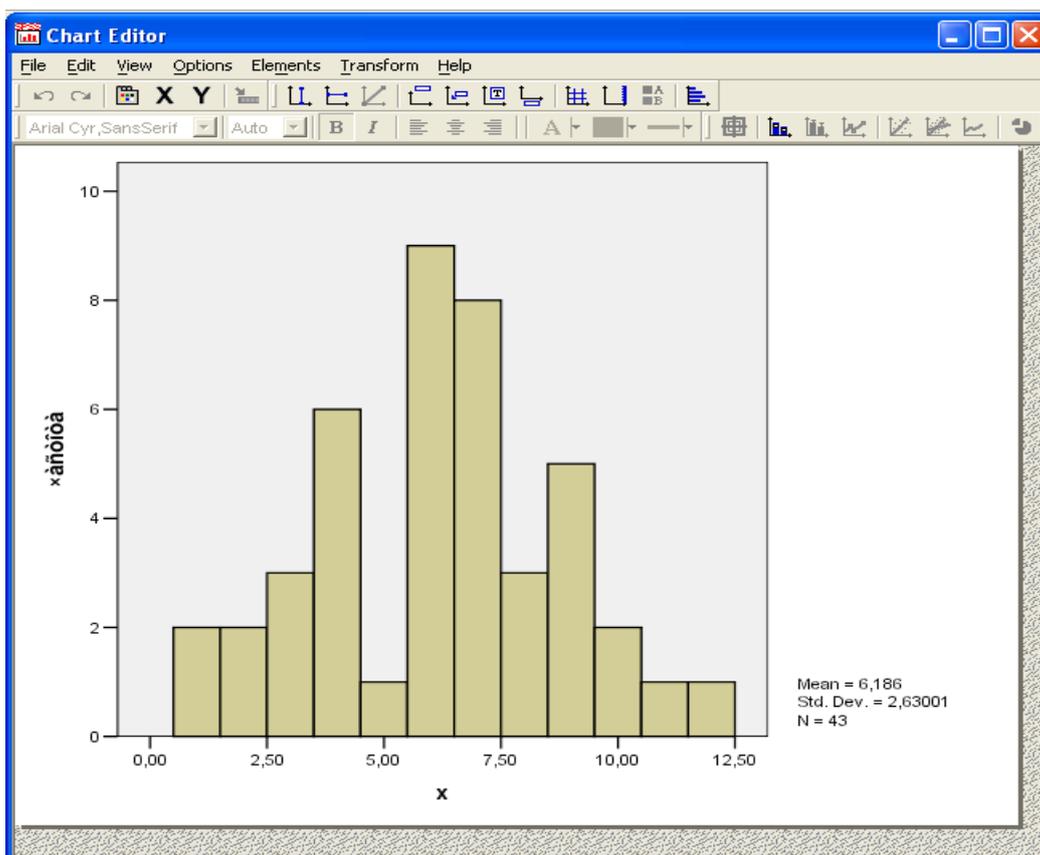


Рис. 1.1.6. Редактор гистограмм

Панель меню изменилась.

Теперь в меню присутствуют только опции, предназначенные для обработки графиков.

Также претерпели изменения и панели инструментов.

Измените сначала метод представления столбцов.

Столбцы должны быть представлены в трёхмерном виде.

В меню редактора можно выбрать вкладку *Elements – Data Value Label*, и таким образом вписать в каждую колонку гистограммы соответствующую частоту (рис. 1.1.7).

Теперь дадим графику название.

Выберите в меню *Option – Title* (Заголовок)

Откроется диалоговое окно *Titles* (Заголовки).

В поле *Title 1* (Заголовок 1) введите текст, например «Гистограмма».

Выберите для заголовка и подзаголовка центральное выравнивание – *Center* (Центр), также здесь можно задать необходимый размер, цвет, стиль и размер шрифта.

Подтвердите нажатием ОК.

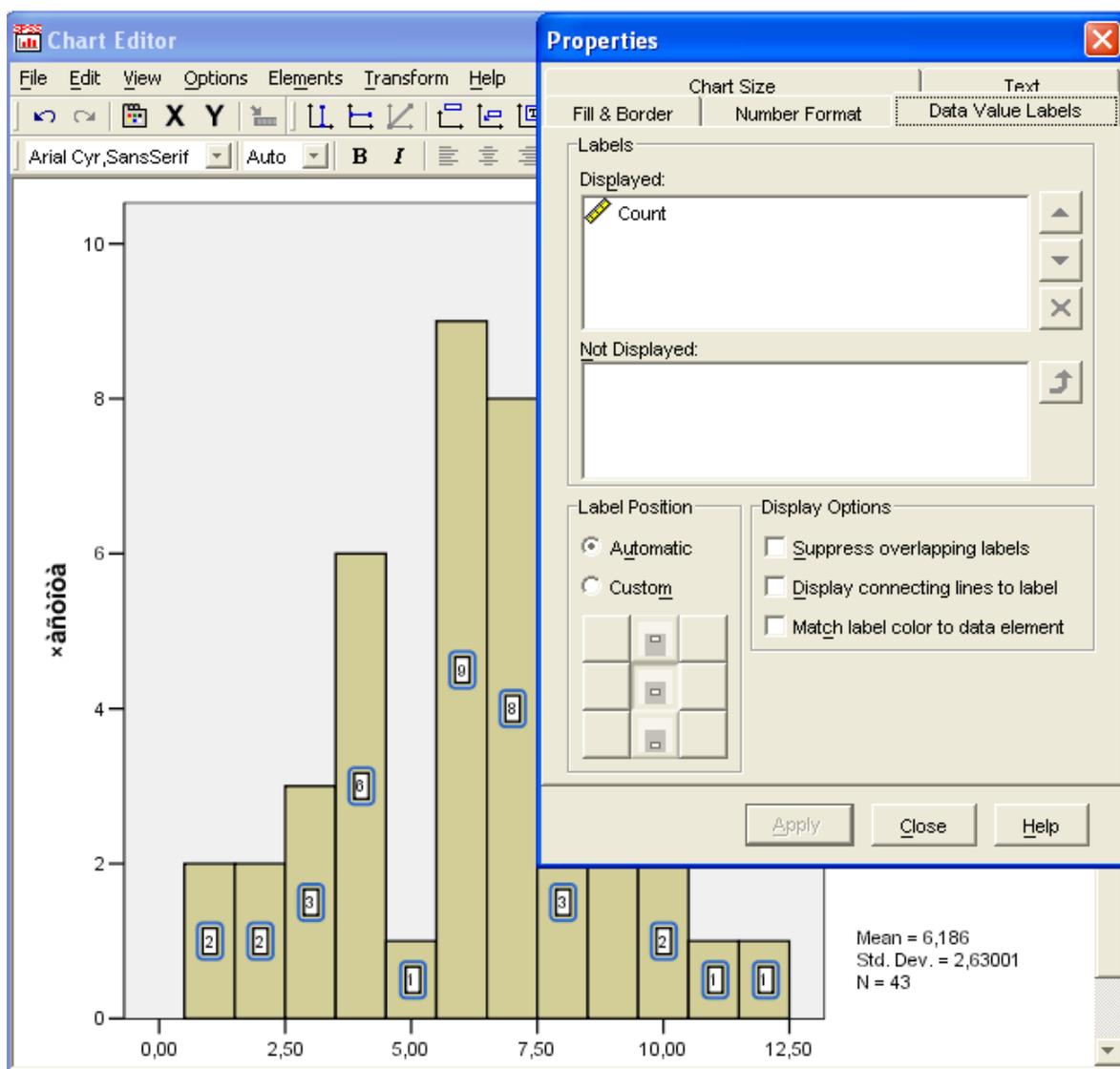


Рис. 1.1.7. Диалоговое меню *Data Value Labels*

Если Вы желаете сохранить построенный график, то поступите следующим образом:

- При помощи щелчка на значке **×** закройте редактор диаграмм.
- Отредактированный график останется в окне просмотра. Этот график (а в общем случае и любые другие результаты, выведенные в окно просмотра) можно сохранить в файле, который имеет формат *Viewer* (средства просмотра SPSS).
- Выберите в меню *File* (Файл) *Save As* (Сохранить как).
- Откроется диалоговое окно *Save As* (Сохранить как).
- Согласно предварительным установкам, SPSS обозначает файлы, которые имеют формат средства просмотра, присваивая им расширение **.spo*.
- Задайте подходящее имя файла и щёлкните на ОК.

Теперь распечатаем график на принтере.

- Выберите в меню *File* (Файл) *Print* (Печать).
- Подтвердите установки диалогового окна *Print* (Печать) при помощи кнопки ОК.

1.1.5. Окно просмотра

Результаты производимых расчётов будут по очереди появляться в окне просмотра, согласно установкам, каждый последующий результат расчёта будет помещаться в конец окна.

Также можно просто в окне просмотра открыть файл **.spo*, в котором сохранены все рассчитанные данные (рис. 1.1.8).

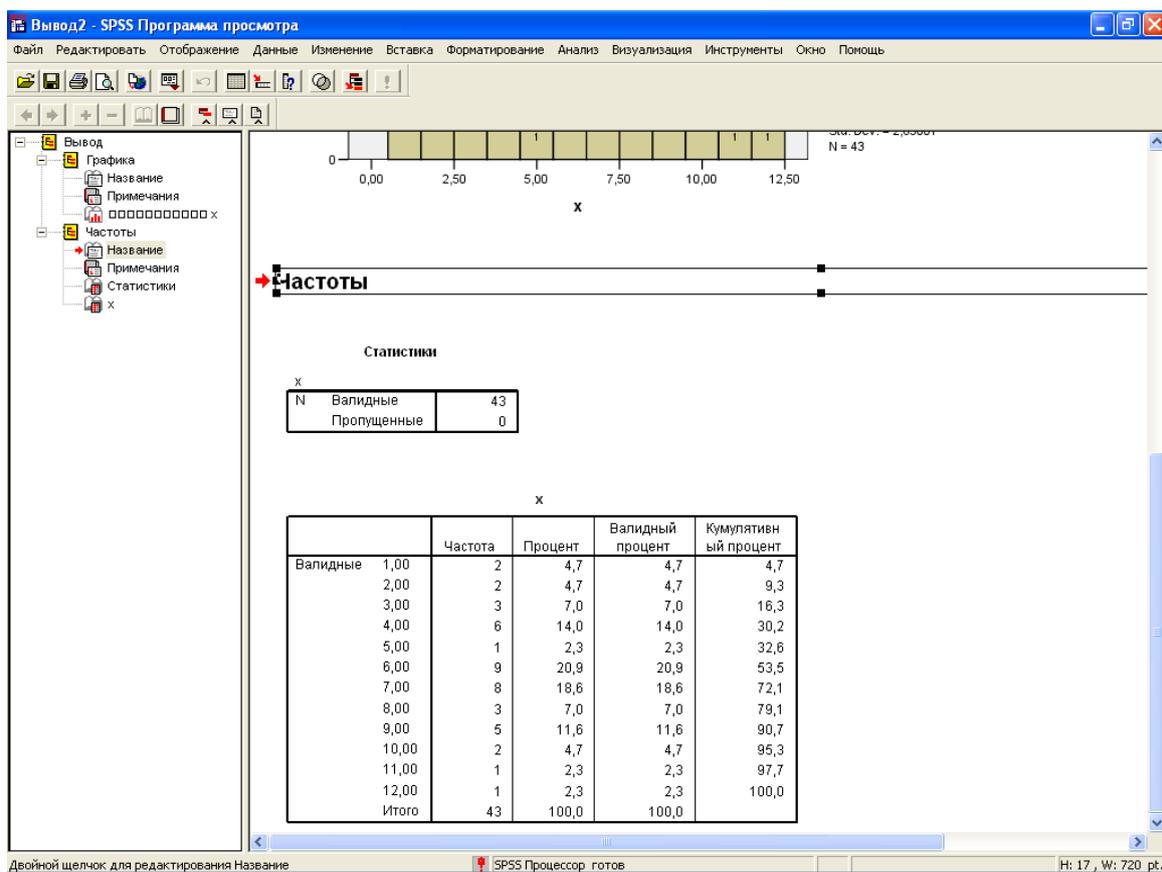


Рис. 1.1.8. Сохранённые результаты

Окно просмотра состоит из двух частей.

В левой части находится иерархия (обзор содержания) результатов; в правую часть помещаются таблицы с результатами расчётов и построенные графики.

Ширину этих частей окна можно изменять перетаскиванием разделительной границы при помощи мыши.

Результаты каждой выполненной статистической процедуры, а также графический вывод, отображаются в окне просмотра в виде блока, причём каждый блок является отдельным объектом.

В иерархии каждый блок озаглавливается соответствующим именем процедуры, перед которым устанавливается значок блока.

Этому значку предшествует небольшой четырёхугольник, в котором сначала указывается знак минус. Внутри каждого блока сначала можно видеть заголовок и примечания.

Далее идёт перечисление элементов блока, которым тоже предшествуют соответствующие символы.

Благодаря такой конструкции иерархии объектов, можно производить поиск необходимых элементов, переставлять их местами, копировать, удалять и т.д.

Поиск в окне просмотра.

Чтобы увидеть в области вывода необходимый объект или элемент, не нужно прокручивать всё окно просмотра. Чтобы попасть в нужное место, щёлкните на соответствующем символе в иерархии.

Удаление в окне просмотра.

Чтобы удалить некоторые элементы результатов расчётов, щёлкните на соответствующем символе и выберите в меню *Edit* (Правка) *Delete* (Удалить)

Вы можете также просто нажать на клавиатуре клавишу *<Delete>*.

Скрытый режим.

Вместо того, чтобы удалять части блоков, можете на некоторое время их скрыть. Они становятся невидимыми на экране и при печати.

Чтобы скрыть части результатов, щёлкните дважды на соответствующем символе в иерархии или выделите нужный элемент одним щелчком с последующим выбором меню *View* (Вид) *Hide* (Скрыть).

Если Вы вновь хотите сделать элемент видимым, повторно щёлкните дважды на значке или выделите его одним щелчком с последующим выбором меню *View* (Вид) *Show* (Показать).

Если же Вы хотите скрыть целый блок, содержащий весь вывод отдельной процедуры, щёлкните на маленьком квадратике слева от значка блока. При этом знак минус в квадратике превратится в знак плюс и данная процедура вместе со всем её содержимым исчезнет.

Вы можете также выделить значок блока и произвести следующий выбор меню *View* (Вид) *Collapse* (Свернуть).

Блок можно вновь сделать видимым при помощи повторного щелчка на квадратике; при этом знак плюс опять будет заменён знаком минус.

Можно также щелчком выделить значок блока и выбрать в меню *View* (Вид) *Expand* (Развернуть).

Перестановка в окне просмотра.

Если вы хотите переместить некоторую часть результатов на другое место, выделите соответствующий значок (если необходимо, то значок блока) и, удерживая нажатой левую кнопку мыши, переместите его к тому

элементу, после которого вы бы хотели расположить данные результаты или блок.

Альтернативная возможность перемещения элементов заключается в выделении значка, соответствующего необходимой части информации с последующим выбором меню *Edit* (Правка) *Cut* (Вырезать).

Затем выделите значок, позади которого вы бы хотели вставить вырезанный элемент и выберите в меню *Edit* (Правка) *Paste After* (Вставить после).

Копирование в окне просмотра.

Если вы хотите скопировать какую-либо часть информации в другое место (при этом сохранив её на прежнем месте), щёлкните на значке, соответствующем нужному элементу или блоку, не отпуская кнопку мыши, нажмите на клавиатуре клавишу <Ctrl> и перетащите значок к тому элементу, после которого должен быть вставлен копируемый элемент.

Вы можете также щёлкнуть на значке копируемого элемента и выбрать в меню следующие опции: *Edit* (Правка) *Copy* (Копировать).

Затем щёлкните на значке элемента, после которого должен быть вставлен копируемый элемент и выберите в меню *Edit* (Правка) *Paste After* (Вставить после).

Вывод примечаний.

При чтении результатов расчётов очень помогают примечания. В них содержится информация о соответствующем файле и общих установках программы.

По умолчанию эти примечания сначала являются скрытыми, но их можно сделать видимыми, если, к примеру, дважды щёлкнуть на значке примечания (*Notes*).

Изменение размера и типа шрифта иерархического списка.

Чтобы изменить размер знаков и тип шрифта в иерархическом списке, выберите в меню *View* (Вид) *Outline Size* (Размер знаков) и соответственно *View* (Вид) *Outline Font* (Шрифт знаков).

У вас появится возможность выбора среди трёх размеров (*Small* (Мелкий), *Medium* (Средний), *Large* (Крупный)) и большого количества шрифтов.

1.1.6. Редактор синтаксиса

Редактор синтаксиса представляет собой текстовое окно, применяемое для набора и запуска на исполнение команд SPSS.

Вы можете вводить команды непосредственно в окне набора или просто переносить установки диалоговых окон при помощи выключателя *Paste* (Вставить), находящегося в самих диалоговых окнах.

Этот перенос возможен благодаря тому, что все диалоговые окна написаны на командном языке SPSS.

С целью реализации дополнительных возможностей или каких-либо индивидуальных подходов к обработке данных, команды, помещённые в редактор синтаксиса, можно изменять.

1.1.7. Информация о файле

Для любого файла SPSS вы можете получить следующую информацию:

- список переменных с их описанием;
- полную информацию обо всех переменных;
- перечень наблюдений.

Откройте файл *.sav.

Если вы хотите просмотреть информацию о значениях переменных, их формате и метках, выберите в меню *Utilities* (Средства) *Variables* (Переменные).

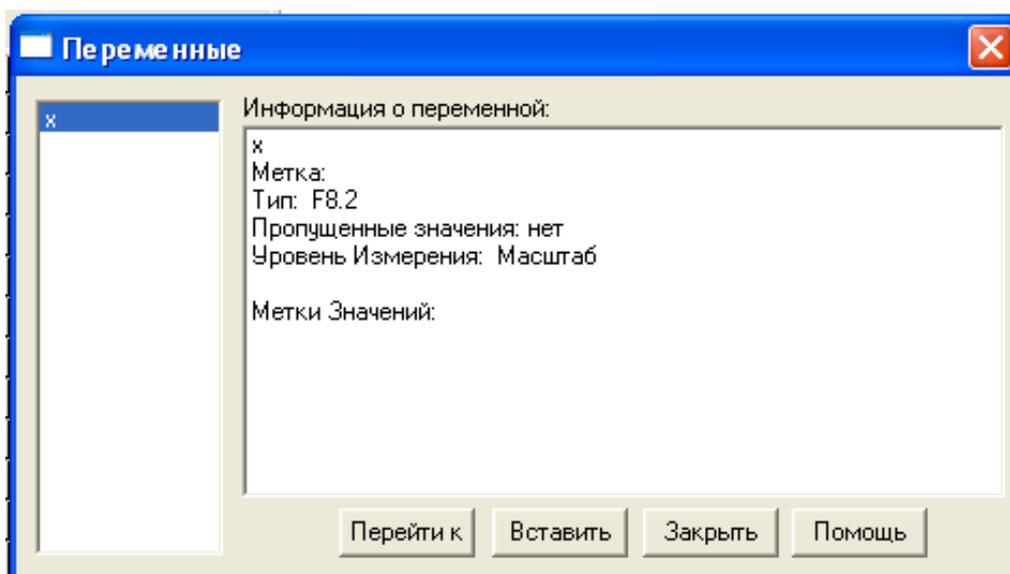


Рис. 1.1.9. Диалоговое окно *Variables* (Переменные)

В информационном окне выводится имя переменной, значения и метки переменной, тип переменной, а также указывается количество пропущенных значений.

Из диалогового окна *Variables* (Переменные) можно сразу перейти к рассматриваемой переменной в окне данных.

Щёлкните для этого на выключателе *Go To* (Перейти к).

Окно данных прокручивается горизонтально таким образом, что переменная, отмеченная нами в диалоговом окне *Variables* (Переменные), оказывается в окне данных на первой позиции.

Выключатель *Paste* (Вставить) копирует имена всех выделенных переменных в редактор синтаксиса.

Некоторую информацию о переменной можно также получить и в любой момент, находясь в диалоговом окне какой-либо статистической процедуры. Чтобы закрыть информационное окно, просто щёлкните на любой точке за его пределами.

Если же вы хотите получить полную информацию обо всех переменных текущего (рабочего) файла, выберите в меню *Utilities* (Дополнительные возможности) *File Info* (Информация о файле).

Если вы хотите получить такую информацию о файле, который не является в данный момент рабочим, то выберите в меню *File* (Файл) *Display Data Info* (Показать информацию о файле).

Выделите необходимый файл, к примеру, *wahl.sav* и подтвердите выбор нажатием кнопки *Open* (Открыть).

Информация о выбранном файле появится в окне просмотра.

1.1.8. Справочная система

Справку в SPSS можно вызвать несколькими способами.

Нажать в любой момент работы функциональную клавишу <F1>. Откроется диалоговое окно *Help: SPSS for Windows* (Справочная система: SPSS для Windows).

Выбрать в главном меню опцию *Help* (Справка).

Находясь в любом диалоговом окне, нажать переключатель с названием *Help* (Справка) и вы получите справку по текущей теме.

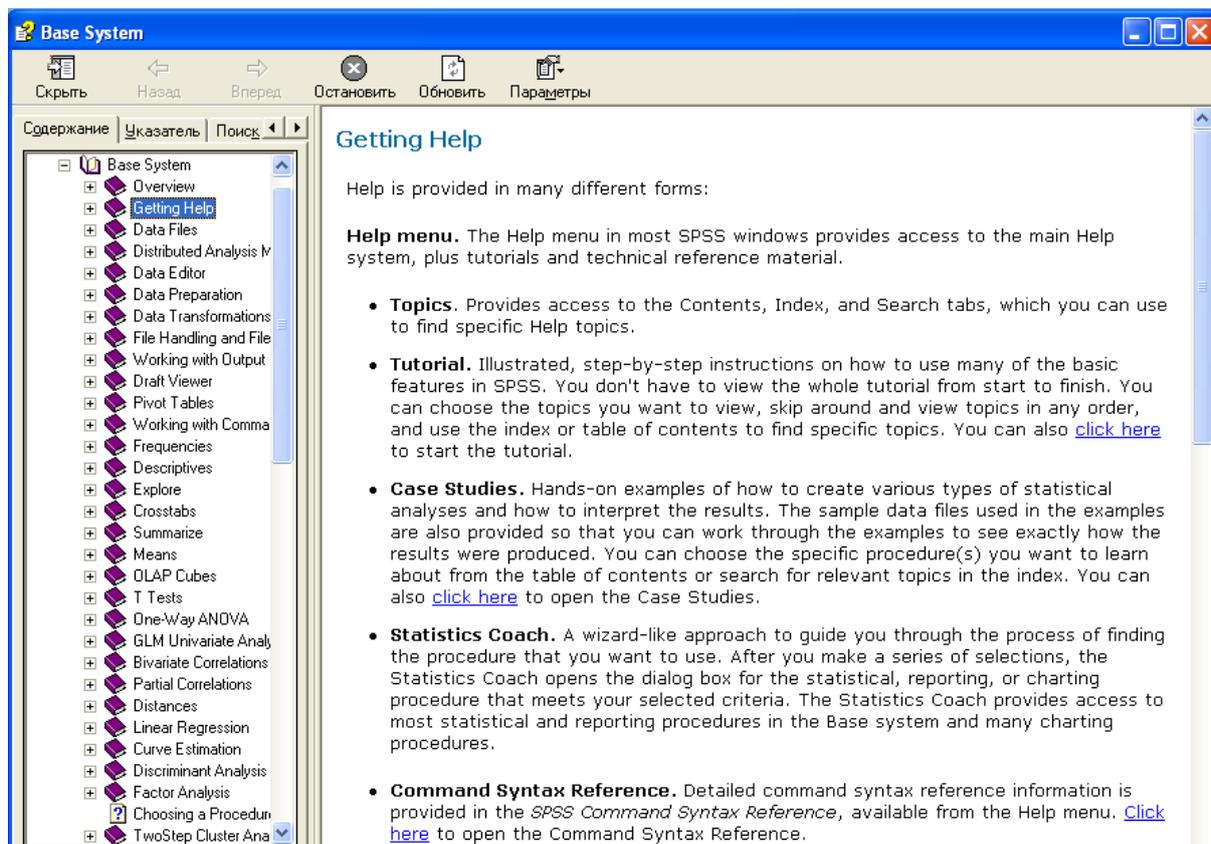


Рис. 1.1.10. Справка SPSS

Чтобы, находясь в справочной системе, параллельно иметь возможность работать в редакторе синтаксиса вы можете уменьшить окно справки до любого необходимого размера и расположить его в удобном месте.

Любую информацию, находящуюся в диалоговом окне справки, можно напечатать с помощью принтера.

Для этого выберите команду *Print* (Печать).

1.1.9. Настройки

Для того, чтобы изменить системные настройки SPSS, выберите в меню *Edit* (Правка) *Options* (Параметры).

Откроется диалоговое окно *Options* (Параметры).

В этом диалоговом окне находятся десять регистрационных карт. Названия отдельных параметров говорят сами за себя, поэтому остановимся только на описании самих регистрационных карт.

1. *General* (Основные).

Здесь вы можете задать тип сортировки списков переменных.

Сортировка в алфавитном порядке, установленная по умолчанию, может быть изменена на порядок, в котором переменные были расположены в рабочем файле.

Вы можете также задать, что указывать во всех диалоговых окнах — метки значений или имена переменных.

2. *Viewer* (Браузер).

Здесь можно установить тип и размер шрифта заголовков и текста, отображаемых в окне просмотра, а также задать размеры страницы.

3. *DraftViewer* (Браузер эскиза).

На этой карте присутствуют различные установки внешнего вида таблиц и текста.

4. *Output Labels* (Показ меток).

Вы можете выбрать, будут ли для обозначения переменных указываться их имена или соответствующие метки (установка по умолчанию) или и то и другое одновременно.

Для обозначения категорий переменной вы можете выбрать значение переменной или метку значения (установка по умолчанию) или оба варианта одновременно.

5. *Charts* (Графики).

Наряду с установками шрифта вы можете также задать, будут ли различные столбцы, линии, области и т.д. отображаться разными цветами (установка по умолчанию) или же при помощи разных штриховок и соответственно типов линий.

Вы также можете управлять компоновкой рамки (рамка снаружи или внутри) и организовывать отображение координатной сетки.

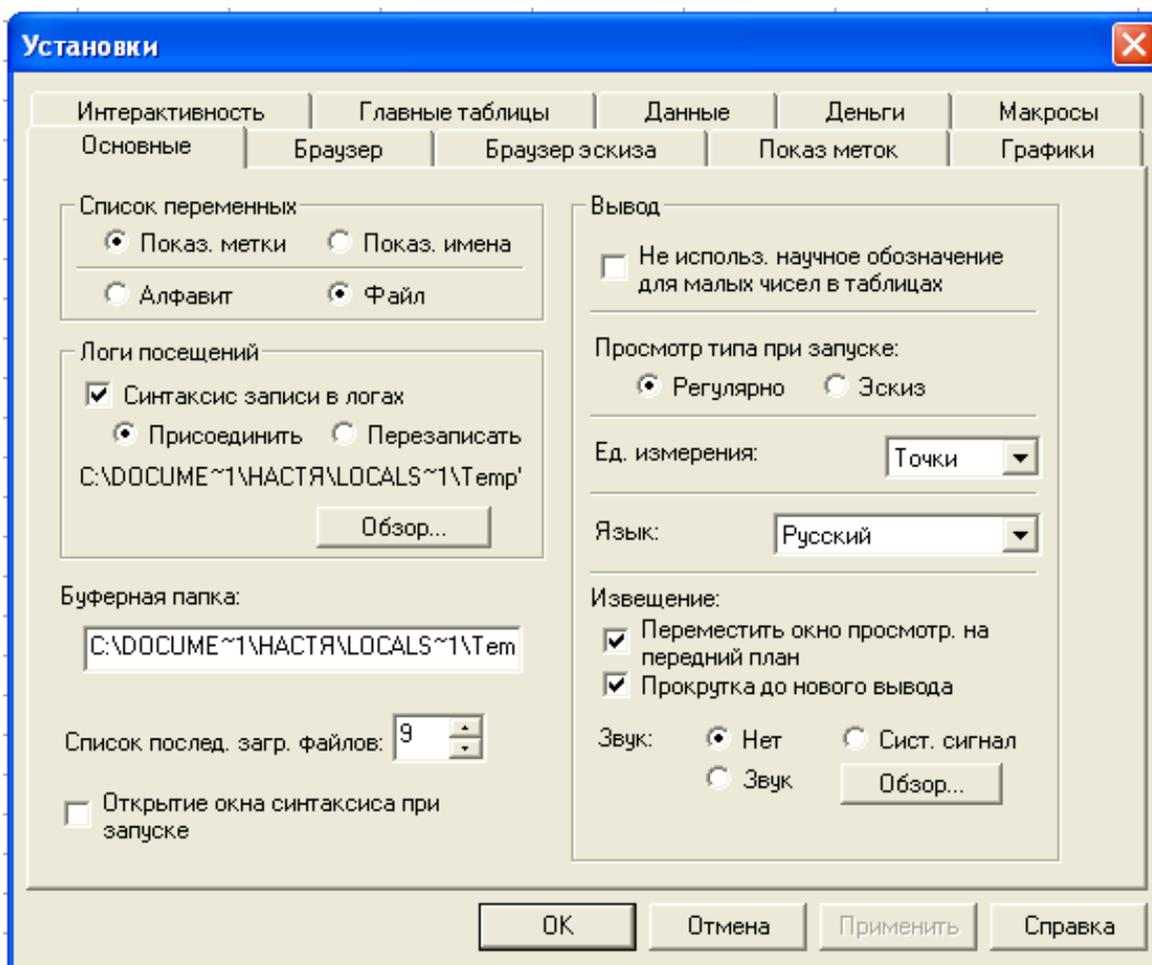


Рис. 1.1.11. Диалоговое окно «Параметры SPSS»

6. *Interactive* (Интерактивный режим).

Вы можете выбрать параметры интерактивных графиков задав, к примеру, некоторый образец.

Если из соображений последующей обработки и вывода на печать вы хотите, чтобы диаграмма была построена в чёрно–белом виде, активируйте для этого образец *Grayscale do* (Оттенки серого).

7. *Pivot Tables* (Главные таблицы).

Здесь можно выбрать внешний вид (компоновку) мобильных таблиц.

8. *Data* (Данные).

В этой карте может быть изменён формат представления рассчитанных переменных (установка по умолчанию: восемь позиций, причём две из них приходится десятичные знаки).

Для отображения года двумя последними цифрами Вы можете дополнительно указать столетие. Если вы активируете автоматическую опцию, столетие будет отсчитываться будет в пределах от 1931 до 2030.

9. *Currency* (Деньги).

Здесь можно указать денежный формат.

10. *Scripts* (Макросы).

Вы можете активировать автоматические сценарии.

1.2. Базовый модуль (SPSS Base)

SPSS Base – это ключевой элемент пакета SPSS для *Windows*.

Он включает все процедуры ввода, отбора и корректировки данных, а также большинство предлагаемых в SPSS статистических методов.

Наряду с простыми методиками статистического анализа, такими как частотный анализ, расчет статистических характеристик, таблиц сопряженности, корреляций, построения графиков, этот модуль включает *t*-тесты и большое количество других непараметрических тестов, а также усложненные методы, такие как многомерный линейный регрессионный анализ, дискриминантный анализ, факторный анализ, кластерный анализ, дисперсионный анализ, анализ пригодности (анализ надежности) и многомерное шкалирование.

С *SPSS Base* можно интегрировать дополнительные модули и другое программное обеспечение SPSS для обеспечения работы по планированию сбора данных, сбору данных, а также внедрению и распространению полученных результатов.

Кроме того, дополнительные модули и программное обеспечение расширяют возможности анализа данных, создания отчетов, а также управления данными и подготовки данных к анализу.

Ниже перечислены вспомогательные средства пакета SPSS.

1.3. Дополнительные модули

Tables.

Модуль *SPSS Tables* позволяет быстро и без лишних усилий создавать совершенные, готовые для презентации таблицы для наглядного и эффективного представления результатов анализа.

Среди возможностей – предварительный просмотр таблиц в процессе их создания, расчет тестовых статистик, а также улучшенные возможности управления данными.

Regression Models.

Построение моделей и предсказание результатов в случаях, когда линейная регрессия оказывается неприменимой.

В *SPSS Regression Models* заложены такие методы анализа данных, как логистическая регрессия (мультиномиальная и бинарная), нелинейная регрессия, регрессия методом взвешенных наименьших квадратов, регрессия двухэтапным методом наименьших квадратов, а также пробит-анализ.

Advanced Models.

Среди мощных многомерных методов *SPSS Advanced Models* – смешанные модели, процедура общего линейного моделирования (ОЛМ), анализ компонент дисперсии, многофакторный дисперсионный анализ,

оценивание Каплана–Мейера, регрессия Кокса, иерархические логлинейные модели, логлинейный анализ, анализ выживаемости и новая процедура универсальные политомические логит–модели (для порядковых исходов).

Categories.

Анализ категориальных данных. *SPSS Categories* позволяет проводить оптимальное шкалирование, включающее анализ соответствий и процедуру *CATPCA* (Анализ главных компонент методом наименьших квадратов с чередованием).

Для поиска неявных зависимостей в данных в *SPSS Categories* есть процедура многомерного шкалирования *PROXSCAL*.

Classification Trees.

SPSS Classification Trees позволяет непосредственно в *SPSS* для *Windows* строить деревья классификаций и решений, идентифицировать группы, находить взаимосвязи в данных и предсказывать будущие события.

Maps.

Мощные геоинформационные возможности, встроенные в *SPSS* для *Windows*. Создание высококачественных карт и проведение демографического анализа на основе данных *SPSS*.

Trends.

Прогнозирование временных рядов: модели экспоненциального сглаживания, а также методы оценивания авторегрессионных моделей.

Exact Tests.

SPSS Exact Tests подсчитывает точные *p*–значения, независимо от структуры данных, даже если количество наблюдений мало, наблюдения разделены на небольшие группы или если в некоторых переменных количество наблюдений в одной категории превышает 80%.

Missing Value Analysis.

Если в данных есть пропущенные значения, этот модуль поможет обнаружить взаимосвязи между пропущенными значениями и другими переменными.

Кроме того, *SPSS Missing Value Analysis* может оценить, какие значения должны были бы стоять на местах пропущенных.

Можно также оценивать среднее, ковариационную и корреляционную матрицы при помощи регрессионного алгоритма или алгоритма максимизации ожидания.

Complex Samples.

Позволяет учитывать сложные планы выборок при проведении анализа данных опросов и обследований, а также планировать сложные выборки и производить отбор. *SPSS Complex Samples* позволяет существенно

снизить вероятность сделать ошибочные выводы для стратифицированных, кластеризованных и многоэтапных выборок.

Conjoint.

Совместный анализ применяется при исследовании рынка для изучения потребительских свойств продуктов на предмет их привлекательности.

SPSS *Conjoint* помогает маркетологам успешно выводить на рынок новые товары.

Conjoint – анализ выявляет наиболее важные для покупателей атрибуты товаров и наиболее предпочтительные параметры атрибутов этих товаров, а также позволяет изучать цены и брэнды.

1.4. Ввод данных (Data Entry)

SPSS *Data Entry* – это эффективный и надежный инструмент, предназначенный для ввода и чистки данных.

При помощи *Data Entry* можно без лишних усилий разрабатывать профессиональные формы, как для печати, так и для ввода данных, организовать быстрый и качественный ввод данных, создать правила, проверяющие логическую согласованность введенных данных.

В отличие от других программ *Data Entry* специально разработана для ввода данных с форм и вопросников и позволяет существенно экономить время и ресурсы.

Разрабатывая формы для ввода, Вы не должны думать о структуре файла данных и о переменных (полях), а можете сконцентрироваться на создании вопросов и форм, а SPSS *Data Entry* автоматически создаст переменные, соответствующие вопросам.

Можно создавать любые типы вопросов, включая вопросы с множественными ответами и матрицы вопросов.

Чистка введенных данных – еще одна черта *Data Entry*, которая делает эту программу уникальной. Формы SPSS *Data Entry* можно дополнять правилами проверки, как отдельных переменных, так и логического соответствия нескольких переменных.

Кроме того, для ускорения и автоматизации ввода можно создавать правила перехода и автозаполнения.

Для организации скоростного ввода данных удобно пользоваться табличной формой ввода, позволяющей вводить данные в табличную сетку, подобную электронной таблице.

При организации ввода с помощью формы данные вводятся путем проставления отметок в элементах управления формой или путем ввода значений в текстовые окна.

Форма запоминает данные и сохраняет их в файле данных.

При вводе можно использовать клавиатуру, мышь и панель Навигатора, чтобы передвигаться по элементам управления и наблюдениям.

Программное обеспечение SPSS *Data Entry* состоит из следующих элементов:

- *DataEntry Builder* – создание форм/вопросников и правил чистки;
- *DataEntry Station* – ввод данных на локальных компьютерах.

1.5. Представление результатов анализа

Визуализация результатов при помощи диаграмм.

SPSS обладает целым рядом графических возможностей для визуализации результатов анализа.

Многочисленные типы диаграмм позволяют представлять результаты в наглядной форме:

- Категориальные диаграммы (включая несколько типов столбиков, линий, областей, кругов и ящиков).
- Диаграммы для контроля качества (включая диаграммы Парето, \bar{X} -среднего и Сигма).
- Гистограммы и диаграммы рассеяния (включая перекрывающиеся, матричные и трехмерные).
- Диагностические и исследовательские графики (включая графики по наблюдениям и графики временных рядов).
- Вероятностные графики (включая графики наблюдаемых и ожидаемых значений).
- Графики автокорреляционной и частной автокорреляционной функции (включая преобразование натурального логарифма и сезонное и несезонное дифференцирование).
- Графики кросс-корреляционной функции (включая преобразование при помощи натурального логарифмирования, сезонное и несезонное дифференцирование).

Система презентационной графики SPSS позволяет без лишних усилий создавать диаграммы, наилучшим образом описывающие результаты анализа, а также редактировать созданные диаграммы для их более тонкой настройки.

Системой презентационной графики также легко пользоваться в случае работы в производственном режиме.

Для этого достаточно однажды создать диаграмму, а после этого применить параметры диаграммы к новым создаваемым диаграммам.

Представление результатов в виде таблиц.

Представить результаты в виде таблиц помогает дополнительный модуль *Tables*.

Интерактивный интерфейс построения таблиц обновляется в режиме реального времени, так что можно видеть, как будет выглядеть таблица, и изменять ее в процессе построения.

Такие возможности, как объединение несколько категорий в одну, вставка итогов и подитогов сверху, снизу, справа или слева в таблице, добавление подкатегорий, изменение типов переменных и исключение категорий позволяют быстро и эффективно управлять внешним видом таблиц.

Кроме того, вместе с таблицами можно рассчитывать статистические критерии, что позволяет устанавливать и подчеркивать достоверность полученных результатов.

Скорость работы и производственные возможности *Tables* существенно облегчают работу тех, кому приходится регулярно создавать большие отчеты, и обеспечивают быстрое и эффективное представление информации, заложенной в огромных массивах данных, в удобной и понятной форме.

SPSS Tables обладает целым рядом возможностей, обеспечивающих удобную и быструю доставку получаемых табличных отчетов тем, для кого они предназначены.

Интерактивные мобильные таблицы, создаваемые в *SPSS Tables*, можно экспортировать в *Word* и *Excel*.

Дополнительного форматирования таблиц не требуется, однако, при необходимости в таблицы можно вставлять содержательную и описательную информацию.

Представление результатов анализа на географических картах.

Если в данных есть географические переменные, например город или регион, то результаты проведенного анализа можно представить на географических картах в удобном и простом для восприятия виде при помощи *Maples*.

Привязка данных к картам производится автоматически, что позволяет сосредоточиться на создании отчета.

Полученные карты с нанесенной статистической информацией можно редактировать для придания им требуемого внешнего вида.

Карты также можно экспортировать в стандартные графические форматы для распространения, в том числе и через Интернет.

2. ПРАКТИКУМ ОБРАБОТКИ ДАННЫХ НА ЭВМ

2.1. Описание показателей по выборке

Задание 2.1.

Исследовать количество опозданий работников цеха за отчетные месяцы: 8, 7, 6, 7, 6, 2, 1, 4, 6, 8, 10, 12, 9, 8, 7, 6, 7, 7, 6, 9, 4, 10, 11, 7, 6, 9, 4, 6, 2, 3, 7, 5, 6, 4, 4, 3, 9, 1, 7, 6, 9, 4, 3.

Выполнение.

Для начала представим графически выборку.

В меню выберем *Graphs* (Визуализация) *Line* (Линия) выберем *Simple* (Простой), в поле *Data in Chart Are* (Данные для графика) выберем *Values of individual cases* (Значения отдельных наблюдений) *Define* (Определить).

В появившемся окне *Define Simple* из левого списка переместим кнопкой–стрелкой *x* в поле *Line Represent*. Нажмем ОК.

Получили следующий график (рис. 2.1.1), который представляет линию выборки.

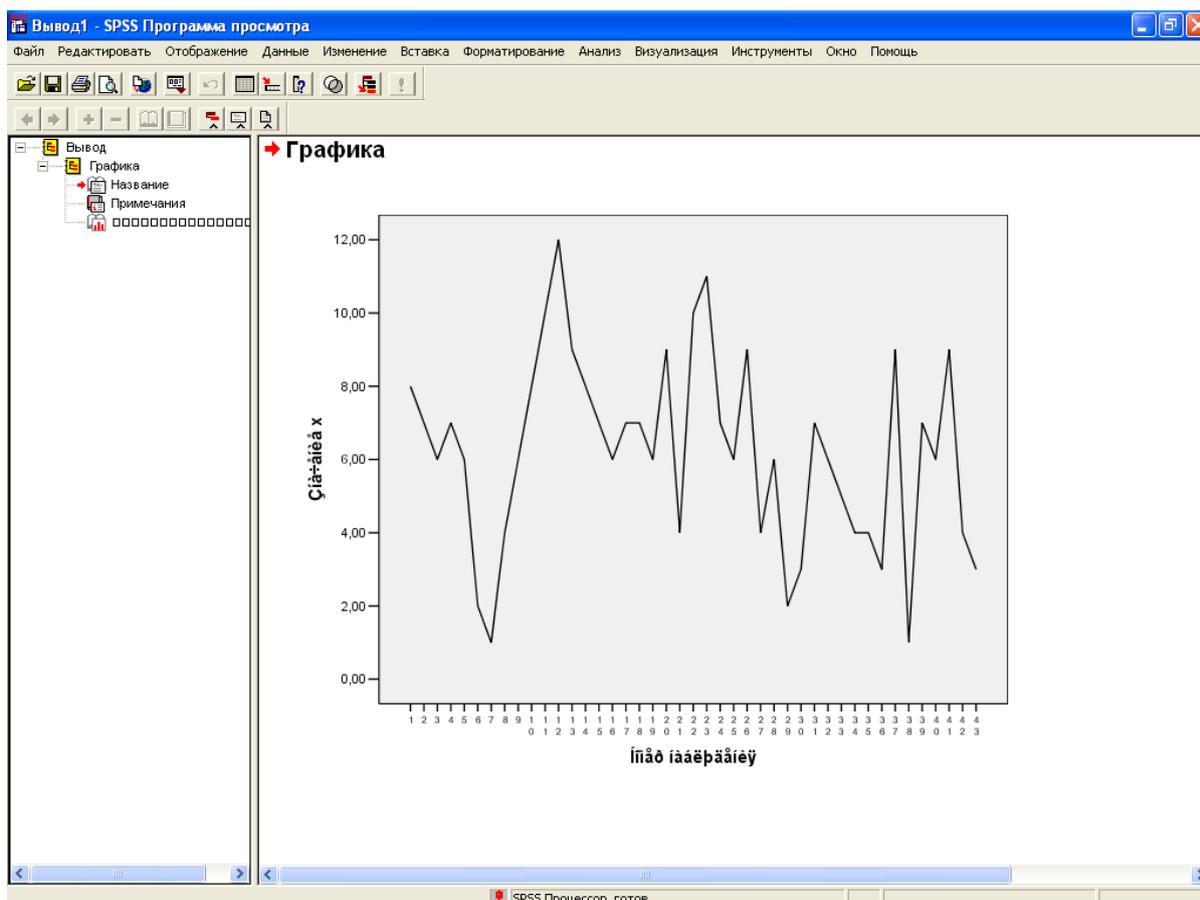


Рис. 2.1.1. Линия выборки

Теперь пострим гистограмму значений.

В меню выберем *Graphs* (Визуализация) *Histogram* (Гистограмма) и занесем в строку *Variable* (Переменная) *x* (рис. 2.1.2).

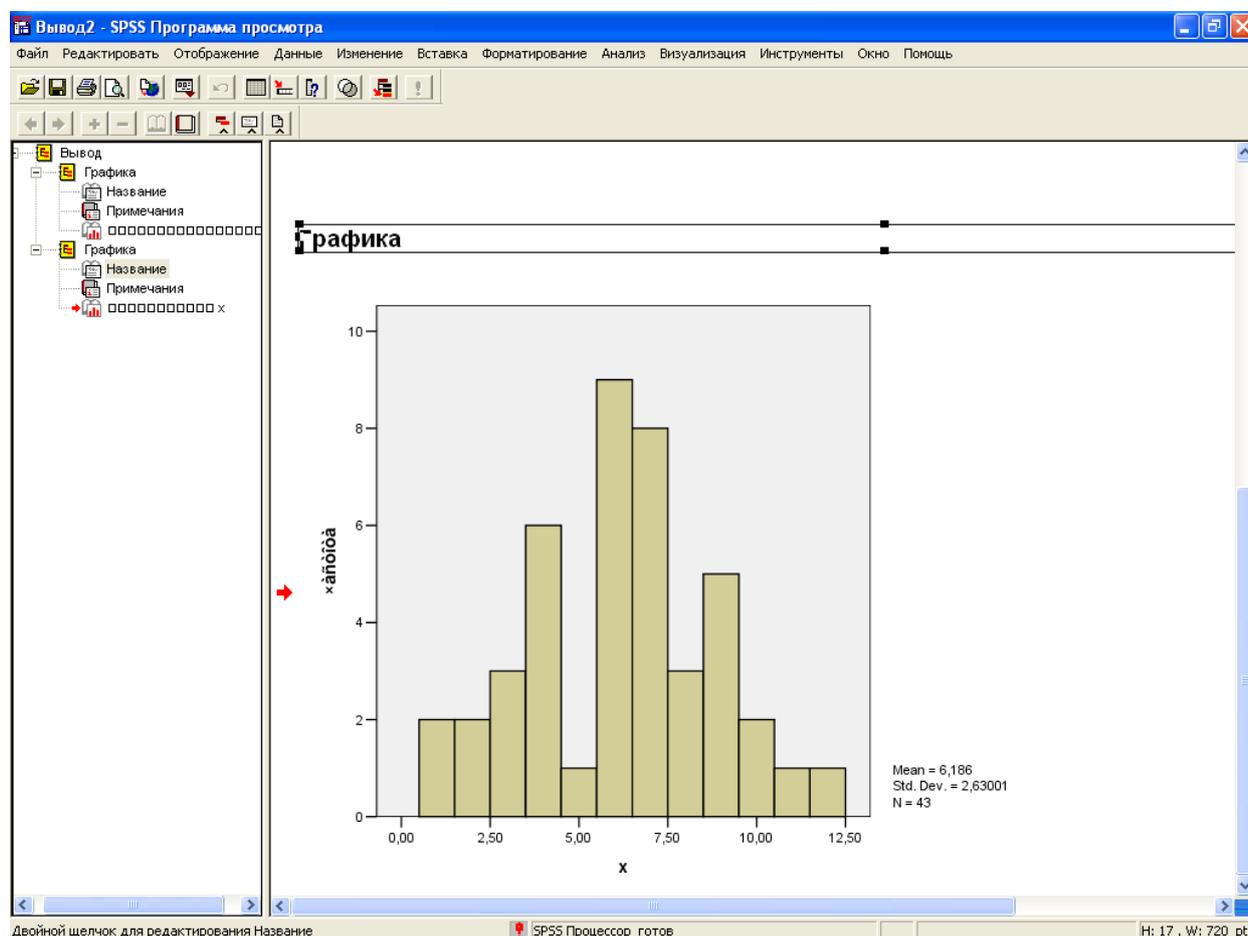


Рис. 2.1.2. Гистограмма значений

Первым этапом статистического анализа данных, как правило, является частотный анализ и получение основных параметров исследуемой величины.

Предварительный анализ данных включает в себя вычисление основных параметров, характеризующих переменные (математическое ожидание, среднее квадратичное отклонение, дисперсия, минимум, максимум и т.д.), составление таблиц частот для одной или нескольких переменных.

Выберем в меню *Analyze* (Анализ) *Statistics* (Статистика) *Frequencies* (Частоты), появится диалоговое окно *Frequencies* (Частоты) (рис. 2.1.4).

	x	var											
1	8,00												
2	7,00												
3	6,00												
4	7,00												
5	6,00												
6	2,00												
7	1,00												
8	4,00												
9	6,00												
10	8,00												
11	10,00												
12	12,00												
13	9,00												
14	8,00												
15	7,00												
16	6,00												
17	7,00												
18	7,00												
19	6,00												
20	9,00												
21	4,00												
22	10,00												
23	11,00												
24	7,00												
25	6,00												
26	9,00												
27	4,00												
28	6,00												
29	2,00												
30	3,00												
31	7,00												
32	6,00												

Рис. 2.1.3. Фрагмент файла данных

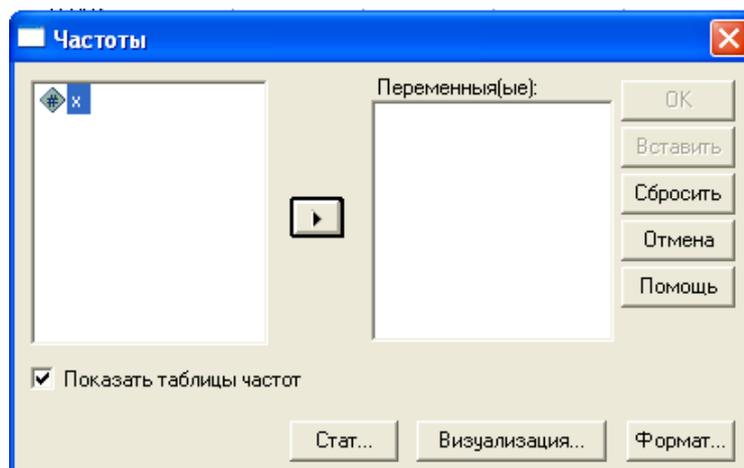


Рис. 2.1.4. Вид диалогового окна *Frequencies* (Частоты)

Кнопкой с треугольником перенесем переменную x в список выходных переменных.

Для получения основных параметров, характеризующих переменную щелкнем в диалоговом окне *Frequencies* (Частоты) на кнопке *Statistics* (Статистика).

Откроется диалоговое окно *Frequencies: Statistics* (Частоты: Статистика). В текущем диалоговом окне выберем интересующие нас пункты.

В группе *Central Tendency* (Средние) выберем следующие характеристики: *Mean* (Среднее значение), *Median* (Медиана).

В группе *Dispersion* (Разброс) выберем следующие меры разброса: *Std. deviation* (Стандартное отклонение), *Variance* (Дисперсия), *Minimum* (Минимум), *Maximum* (Максимум).

Для продолжения расчетов нажмем кнопку *Continue* (Продолжить). Для графического представления имеющихся данных щелкнем на кнопке *Charts* (Диаграммы). Откроется диалоговое окно *Frequencies: Charts* (Частоты: Диаграммы).

Выберем в группе *Chart Type* (Тип диаграммы) пункт *Bar Charts* (Столбчатая диаграмма), а в группе *Chart Values* (Значения диаграммы) – пункт *Frequencies* (Частота).

Подтвердим выбор кнопкой *Continue* и вернемся в диалоговое окно *Frequencies*.

Подтвердим операцию нажатием кнопки ОК в диалоговом окне *Frequencies* (Частоты). Вывод основных результатов выглядит следующим образом (рис. 2.1.5).

Статистики

x		
N	Валидные	43
	Пропущенные	0
Среднее		6,1860
Медиана		6,0000
Мода		6,00
Стд.отклонение		2,63001
Дисперсия		6,917
Асимметрия		-,047
Стд. ошибка асимметрии		,361
Эксцесс		-,379
Стд. ошибка эксцесса		,709
Размах		11,00
Минимум		1,00
Максимум		12,00
Сумма		266,00

x					
		Частота	Процент	Валидный процент	Кумулятивный процент
Валидные	1,00	2	4,7	4,7	4,7
	2,00	2	4,7	4,7	9,3
	3,00	3	7,0	7,0	16,3
	4,00	6	14,0	14,0	30,2
	5,00	1	2,3	2,3	32,6
	6,00	9	20,9	20,9	53,5
	7,00	8	18,6	18,6	72,1
	8,00	3	7,0	7,0	79,1
	9,00	5	11,6	11,6	90,7
	10,00	2	4,7	4,7	95,3
	11,00	1	2,3	2,3	97,7
	12,00	1	2,3	2,3	100,0
	Итого	43	100,0	100,0	

Рис. 2.1.5. Результаты выполнения процедуры

В первой из таблиц представлены основные статистические параметры, характеризующие исследуемый показатель урожайности x .

Видно, что математическое ожидание $m_x = 6,1860$, медиана $M_e = 6,0000$, среднее квадратическое отклонение $\sigma_x = 2,63001$, дисперсия $D_x = \sigma_x^2 = 6,917$, также приведены максимальное и минимальное значение исследуемого показателя.

Во второй таблице можно увидеть частотное распределение исследуемой величины, т.е. каждая строка частотной таблицы описывает одно возможное значение.

Теперь возможно построение полигона.

Полигон – это ломаная, соединяющая точки $(x_i; n_i)$.

Для этого скопируем из окна просмотра результатов в редактор данных 1 и 2 столбцы в таблице x .

Для этого необходимо дважды кликнуть левой кнопкой мыши в нужном столбце, выделив соответствующие значения, скопируем их.

Вернемся в редактор данных, поставим указатель мыши в любой из свободных столбцов, расположенных правее столбца значений x , и вставим скопированные значения. Аналогично сделаем с частотами.

Для простоты переименуем VAR00001 в x_i и VAR00002 – n_i .

Для чего необходимо дважды кликнуть левой кнопкой мыши по VAR00001, или перейти на вкладку «Обзор переменных», и в соответствующих строках ввести соответствующее значение. Теперь редактор переменных выглядит, как показано на рисунке 2.1.6.

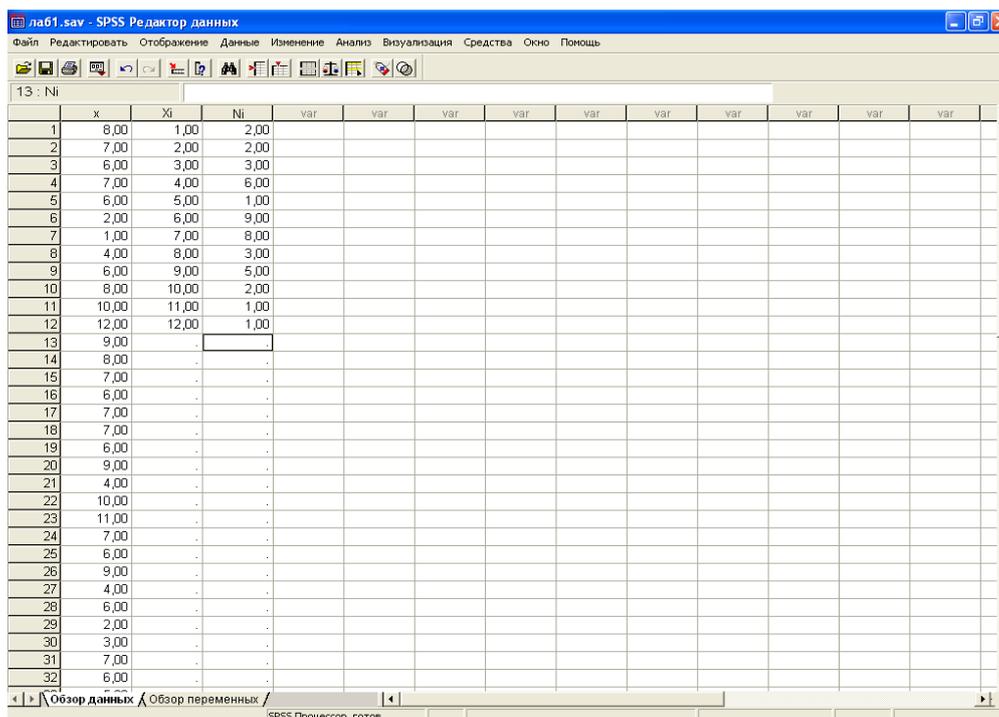


Рис. 2.1.6. Редактор данных

Для построения полигона частот необходимо выбрать *Graphs* (Визуализация) *Scatter* (Разброс/точка) *Simple* (Простой) *Define* (Определить).

В определении переменных выбираем за ось $Y - n_i$ за ось $X - x_i$ – ОК.

В редакторе графиков изменяем значения метки \circ на $*$, дважды кликнув по любой из метки и установив соответствующее значение метки.

Теперь график выглядит как на рис. 2.1.7.

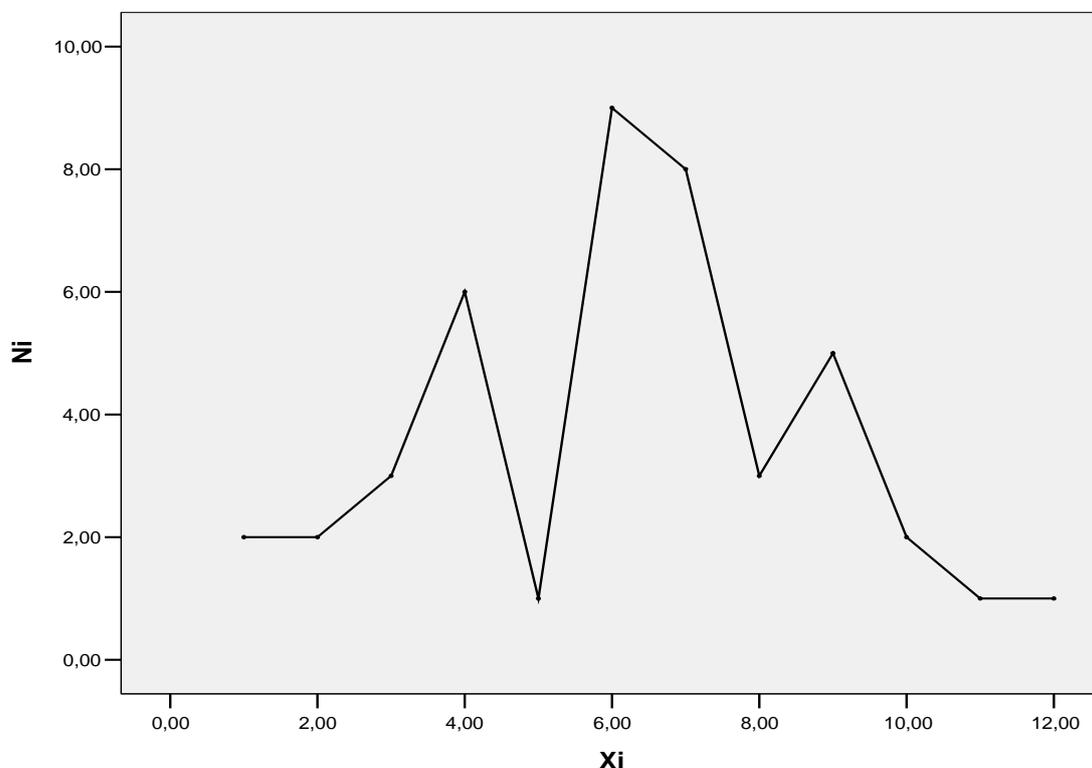


Рис. 2.1.7. Полигон частот

2.2. Интервальные оценки параметров

Задание 2.2.

По имеющимся данным и результатам расчетов, выполненных в работе 2.1, построить доверительные интервалы для математического ожидания для различных значений доверительной вероятности (0,9; 0,95; 0,99)

Выполнение.

SPSS предоставляет возможность нахождения доверительных интервалов для математического ожидания для конкретной выборки с необходимым значением доверительной вероятности.

Чтобы понять, что может предложить SPSS для решения этой задачи, возьмем переменную X (количество опозданий работников цеха за отчетные месяцы) из предыдущей работы.

Перейдем к исследованию данных, выбрав команды меню: *Analyze* (Анализ) *Statistics* (Статистика) *Analyze* (Анализ).

Откроется диалоговое окно *Analyze* (Анализ) (см. рис. 2.2.2).

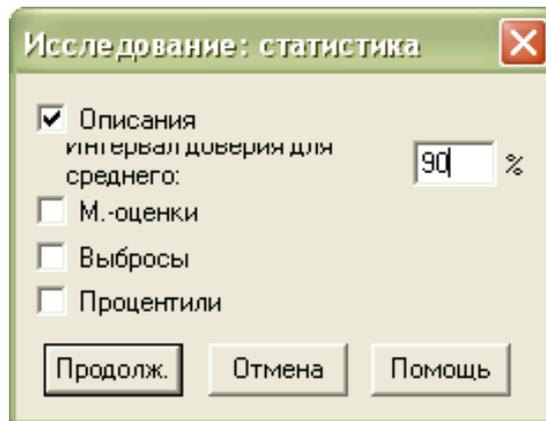
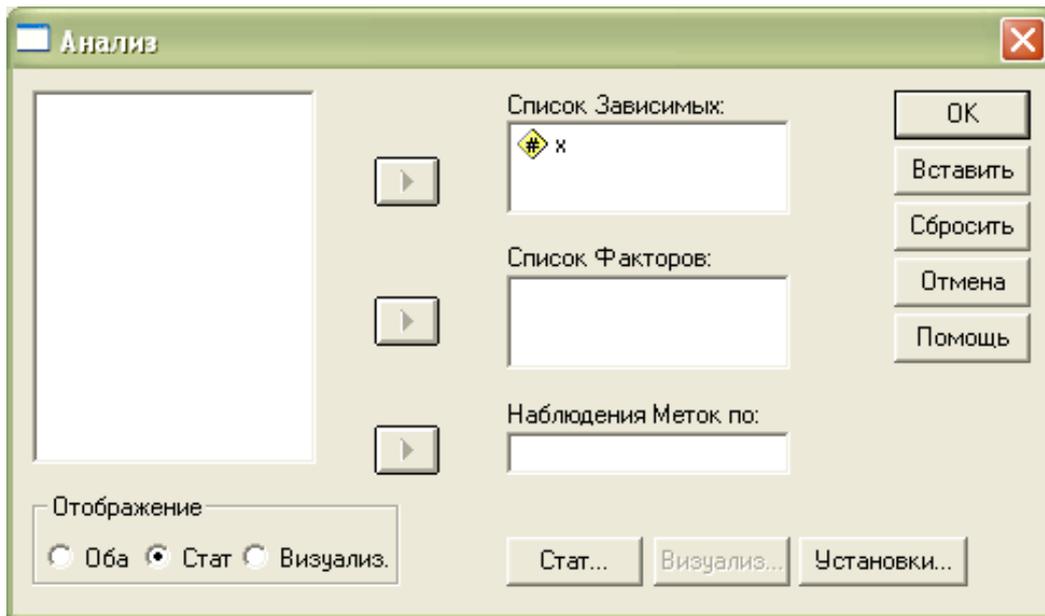


Рис. 2.2.2. Диалоговое окно «Анализ»

Перенесем переменную x в список зависимых переменных.

Поскольку важность для нас представляет только доверительный интервал для математического ожидания, то в меню *Display* (Отображение), выберем *Statistics* (Статистика).

Далее нажав, справа на *Statistics* (Статистика) зададим величину доверительной вероятности (0,9; 0,95; 0,99).

Нажав ОК, получим результаты.

Таким образом, после ввода трех значений доверительной вероятности получаем результаты в окне вывода.

Описательные

			Статистика	Стд. ошибка
x	Среднее		6,1860	,40107
	90% доверительный интервал для среднего	Нижняя граница	5,5115	
		Верхняя граница	6,8606	
	5% усеченное среднее		6,1809	
	Медиана		6,0000	
	Дисперсия		6,917	
	Стд. отклонение		2,63001	
	Минимум		1,00	
	Максимум		12,00	
	Размах		11,00	
	Межквартильный размах		4,00	
	Асимметрия		-,047	,361
	Эксцесс		-,379	,709

Описательные

			Статистика	Стд. ошибка
x	Среднее		6,1860	,40107
	95% доверительный интервал для среднего	Нижняя граница	5,3766	
		Верхняя граница	6,9954	
	5% усеченное среднее		6,1809	
	Медиана		6,0000	
	Дисперсия		6,917	
	Стд. отклонение		2,63001	
	Минимум		1,00	
	Максимум		12,00	
	Размах		11,00	
	Межквартильный размах		4,00	
	Асимметрия		-,047	,361
	Эксцесс		-,379	,709

Описательные

			Статистика	Стд. ошибка
x	Среднее		6,1860	,40107
	99% доверительный интервал для среднего	Нижняя граница	5,1039	
		Верхняя граница	7,2682	
	5% усеченное среднее		6,1809	
	Медиана		6,0000	
	Дисперсия		6,917	
	Стд. отклонение		2,63001	
	Минимум		1,00	
	Максимум		12,00	
	Размах		11,00	
	Межквартильный размах		4,00	
	Асимметрия		-,047	,361
	Эксцесс		-,379	,709

Рис. 2.2.3. Результаты анализа

Большая часть выведенных характеристик была рассмотрена в предшествующей работе.

Появились новые характеристики:

- 5% усеченное среднее – среднее значение, вычисленное без учета 5% наименьших и 5% наибольших значений;
- межквартильный размах – расстояние между первым и третьим квартилями;
- стд. ошибка – стандартная ошибка характеристик (среднее, асимметрия и эксцесс).

В интервале шириной, равной удвоенной стандартной ошибке, отложенному вокруг данной характеристики, располагается значение этой характеристики для генеральной совокупности с вероятностью примерно 67%.

2.3. Проверка гипотез по статистическим данным

Задание 2.3.

По имеющимся статистическим данным и результатам расчетов, выполненных в работах 2.1 и 2.2, проверить гипотезы:

1) о непротиворечивости данных выбранному теоретическому закону распределения;

2) о равенстве выборочного математического ожидания статистического распределения с гипотетическим (гипотетическое математическое ожидание принять «на глазок», исходя из физических свойств распределения, используя правило трех сигм).

Выполнение.

В SPSS существует возможность проверить, соответствует ли реальное распределение переменной нормальному, равномерному, экспоненциальному распределению или распределению Пуассона, при помощи теста Колмогорова–Смирнова.

Чтобы продемонстрировать работу данного теста, проверим на предмет наличия нормального распределения исходные значения переменной X (количество опозданий работников цеха за отчетные месяцы) из предыдущей работы.

Выберем в меню *Analyze* (Анализ) *Nonparametric Tests* (Непараметрические тесты) *1-Sample KS* (К–С одной выборки).

Появится диалоговое окно *One Sample Kolmogorov–Smirnov Test* (Тест Колмогорова–Смирнова для одной выборки) (рис. 2.3.2).

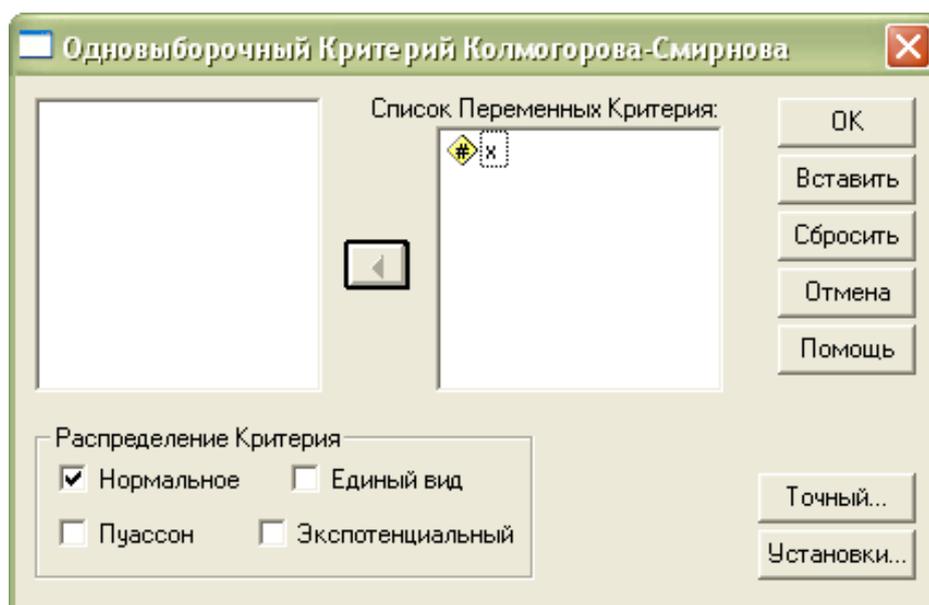


Рис. 2.3.2. Диалоговое окно «Тест Колмогорова–Смирнова для одной выборки»

Перенесем переменную x в поле тестируемых переменных.

Если щелкнуть по кнопке *Options* (Установки), то можно дополнительно организовать вывод характеристик дескриптивной статистики и квантилей.

Щелкнем на ОК.

В окне просмотра появятся следующие результаты.

Одновыборочный критерий Колмогорова-Смирнова

		x
N		43
Нормальные параметры ^{а,б}	Среднее	6,1860
	Стд. отклонение	2,63001
Разности экстремумов	Модуль	,146
	Положительные	,099
	Отрицательные	-,146
Статистика Z Колмогорова-Смирнова		,959
Асимпт. знч. (двухсторонняя)		,317

а. Сравнение с нормальным распределением.

б. Оценивается по данным.

Рис. 2.3.3. Результаты теста Колмогорова–Смирнова

Полученные результаты включают:

- среднее значение и стандартное отклонение;
- промежуточные результаты, полученные при выполнении теста Колмогорова–Смирнова;
- вероятность ошибки p .

Отклонение от нормального распределения считается существенным при значении $p < 0,05$; в этом случае для соответствующих переменных следует применять непараметрические тесты.

В рассматриваемом примере (значение $p = 0,317$), то есть вероятность ошибки является не значимой; поэтому значения переменной достаточно хорошо подчиняются нормальному распределению.

Аналогично можно провести анализ и по остальным видам распределения.

Выяснить, отличается ли среднее значение, полученное на основе данной выборки, от предварительно заданного контрольного значения, позволяет T -тест одной выборки.

Мы проверим, отличается ли средний показатель количества опозданий работников цеха за отчетные месяцы, полученный при нашем исследовании, от значения 7,32, которое могло быть определено в каком-либо другом исследовании (по правилу трех сигм для нормального закона с точностью 0,997 весь диапазон разброса случайной величины заключен в пределах $\pm 3\sigma$ от центра величины).

Выберем в меню команды *Analyze* (Анализ) *Compare Means* (Сравнение средних) *One-Sample T Test* (Одновыборочный T критерий).

Откроется диалоговое окно *One-Sample T Test* (Одновыборочный T критерий) (рис. 2.3.4).

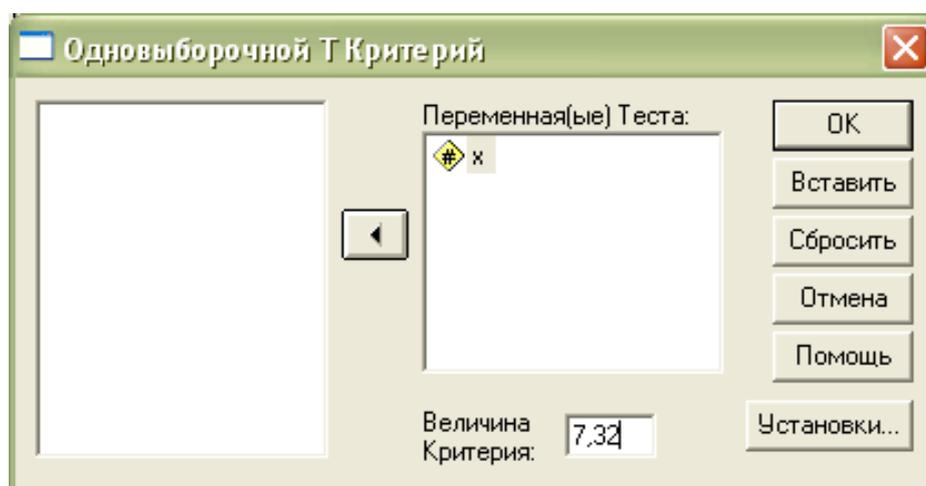


Рис. 2.3.4. Диалоговое окно «Одновыборочный T –критерий»

Перенесем переменную X в поле *Test Variable(s)* (Переменная(ые) Теста) и введем в поле *Test Value* (Величина Критерия) значение 7,32.

Кнопкой *Options* (Установки) можно задать вместо 95% любой другой доверительный интервал.

Значение доверительного интервала может принимать значения в промежутке от 1 до 99%. Запустим вычисления, щелкнув ОК.

Одновыборочные статистики

	N	Среднее	Стд. отклонение	Стд. ошибка среднего
x	43	6,1860	2,63001	,40107

Одновыборочный критерий

Тестовое значение = 7.32						
	t	ст.св.	Знч. (2-сторон)	Разность Средних	95% доверительный интервал разности средних	
					Нижняя граница	Верхняя граница
x	-2,827	42	,007	-1,13395	-1,9434	-,3246

Рис. 2.3.5. Результаты T -теста

Результаты, показанные в окне просмотра, свидетельствуют о том, что в данном исследовании среднее исходное количество опозданий работников цеха за отчетные месяцы, что составляет 6,186, значимо ($p = 0,007$) отличается от контрольного значения 7,32.

2.4. Корреляционный и регрессионный анализы

Задание 2.4.

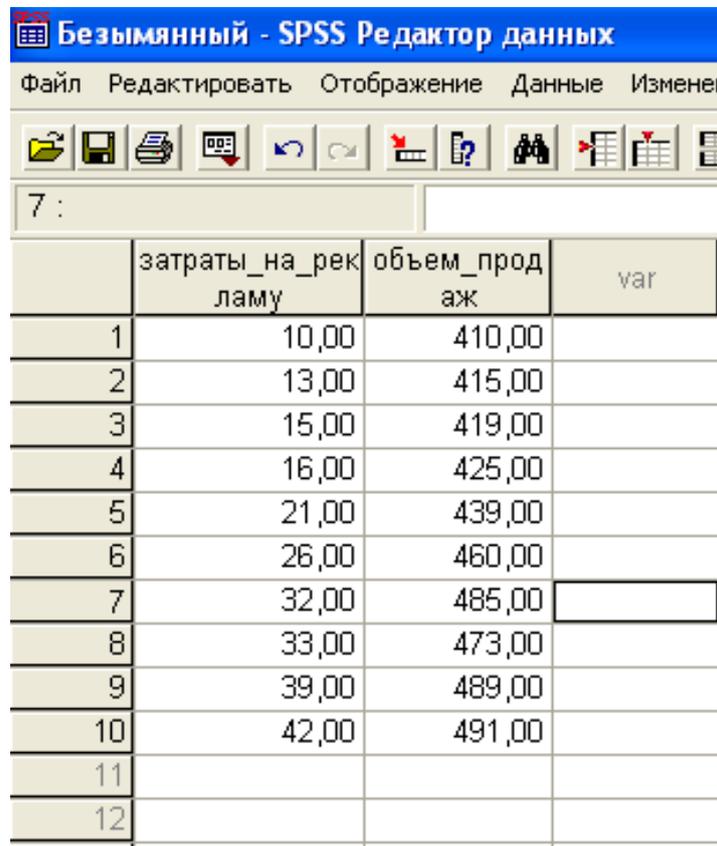
Произвести анализ зависимости объема продаж товара от затрат на рекламу данного товара. Данные приведены в табл. 2.4.1.

Таблица 2.4.1

Затраты на рекламу, тыс. руб.	10	13	15	16	21	26	32	33	39	42
Объемы продаж, тыс. руб.	10	15	10	25	39	60	85	73	89	91

Выполнение.

В начале работы создадим файл с данными как показано на рис. 2.4.1.



The screenshot shows the SPSS Data Editor window titled "Безымянный - SPSS Редактор данных". The menu bar includes "Файл", "Редактировать", "Отображение", "Данные", and "Изменить". The toolbar contains icons for file operations, editing, and data management. Below the toolbar, a small table shows the current data structure with columns: "затраты_на_рекламу", "объем_продаж", and "var". The main data table has 12 rows, with the first 10 rows containing numerical data.

	затраты_на_рекламу	объем_продаж	var
1	10,00	410,00	
2	13,00	415,00	
3	15,00	419,00	
4	16,00	425,00	
5	21,00	439,00	
6	26,00	460,00	
7	32,00	485,00	
8	33,00	473,00	
9	39,00	489,00	
10	42,00	491,00	
11			
12			

Рис. 2.4.1. Фрагмент файла данных

В пакете SPSS для корреляционного анализа есть раздел «Корреляция в меню. Анализ».

Для более наглядного представления имеющихся данных построим график зависимости «Затраты на рекламу – объем продаж» в виде диаграммы рассеяния.

- Выберем в меню *Graphs* (Визуализация) *Scatter-Dot* (Разброс/точка), откроется диалоговое окно *Scatter-Dot* (Разброс/точка) (рис. 2.4.2).

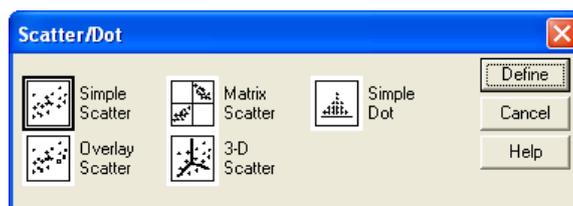


Рис. 2.4.2. Диалоговое окно *Scatter-Dot* (Разброс/точка)

- В диалоговом окне *Scatter-Dot* (Разброс/точка) щёлкнем на области *Simple Scatter* (Простой разброс).
- Щелчком по выключателю *Define* (Определить) откроем соответствующее диалоговое окно (рис. 2.4.3).
- Отообразим объем продаж в зависимости от затрат на рекламу, поэтому переменную *объемы_продаж* из списка исходных переменных перенесем в поле оси *Y*, а переменную *затраты_на_рекламу* – в поле оси *X*. И начнем построение диаграммы щелчком на ОК.

Результатом выполнения вышеуказанных команд будет следующий график (рис. 2.4.4).

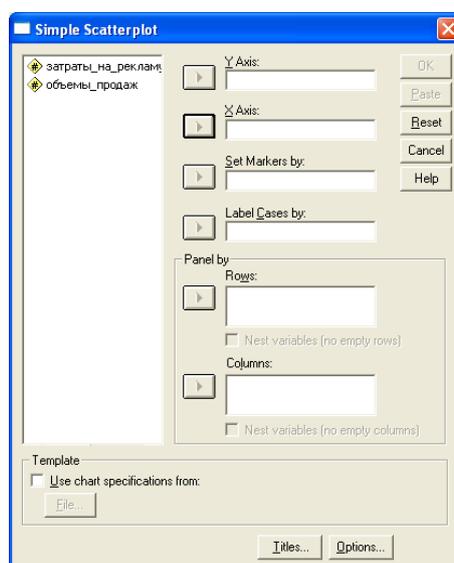


Рис. 2.4.3. Вид окна *Simple Scatterplot* (Простой график рассеяния).

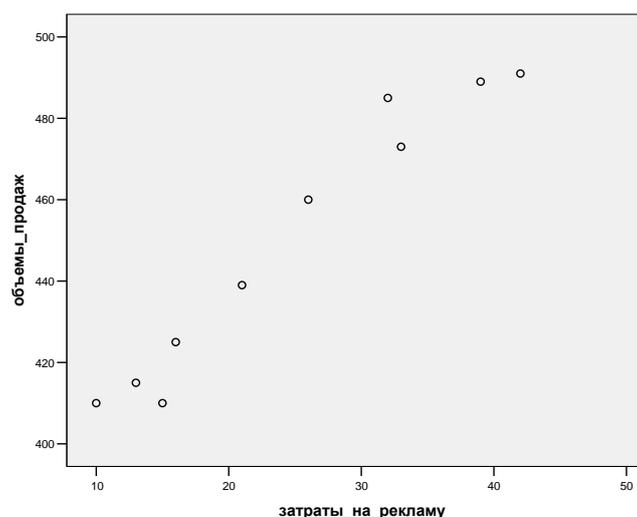


Рис. 2.4.4. Простой график рассеивания

Теперь определим основные корреляционные показатели. Для этого в меню Анализ выбираем Корреляция – Двумерный. Далее появится диалоговое окно (рис. 2.4.5).

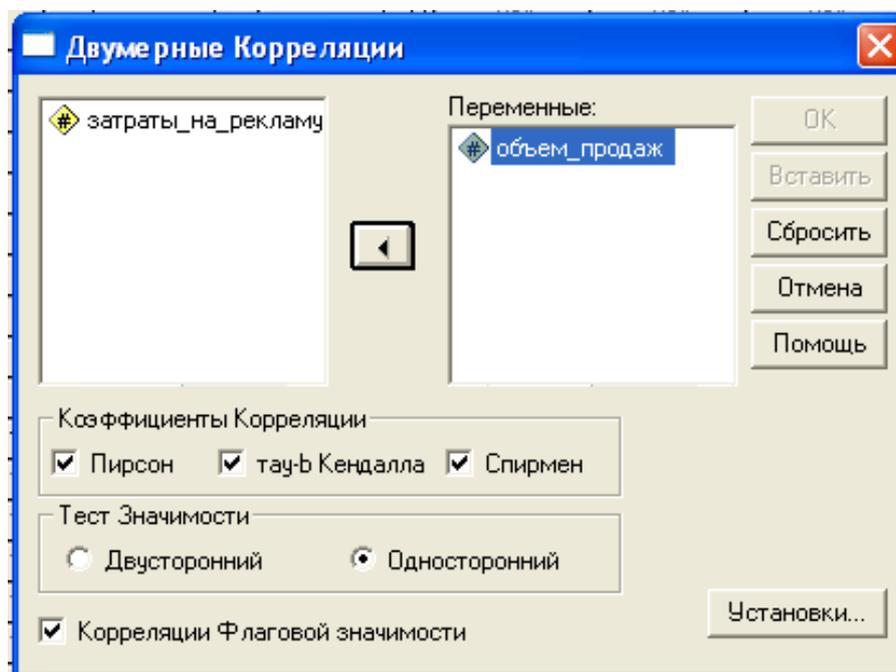


Рис. 2.4.5. Диалоговое окно «Двумерная корреляция»

Далее получаем следующий вывод (рис.2.4.6).

Регрессионный анализ служит для определения вида связи между переменными и дает возможность для прогнозирования значения одной (зависимой) переменной отталкиваясь от значения другой (независимой) переменной.

В пакете SPSS для этой цели имеется раздел *Regression* (Регрессия) (меню *Analyze* (Анализ)), который предоставляет пользователю широкий набор процедур регрессионного анализа.

Каждая процедура имеет модель регрессии, которая соотносит зависимую переменную с независимой переменной (или множеством независимых переменных).

Простая линейная регрессия лучше всего подходит для того, чтобы продемонстрировать основополагающие принципы регрессионного анализа.

По виду получившейся диаграммы рассеяния можно предположить о наличии линейной зависимости между исследуемыми показателями.

Описательные статистики

	Среднее	Стд. отклонение	N
затра- ты на рекламу	24,7000	11,37297	10
объем продаж	450,6000	32,63332	10

Корреляции(a)

		затра- ты на рекламу	объем продаж
затра- ты на рекламу	Корреляция Пирсона	1	,983(**)
	Знч.(1– сторон)		,000
объем_продаж	Корреляция Пирсона	,983(**)	1
	Знч.(1– сторон)	,000	

** Корреляция значима на уровне 0.01 (1–сторон.).
а Искл. целиком N=10

Корреляции(a)

		затра- ты на рекламу	объем_ продаж
тау–b Кен- далла	затра- ты_на_реклам у	Коэффициент корреляции	1,000
		Знч. (1–сторон)	,956(**)
			,000
	объем_продаж	Коэффициент корреляции	,956(**)
		Знч. (1–сторон)	1,000
			,000
ро Спир- мена	затра- ты_на_реклам у	Коэффициент корреляции	1,000
		Знч. (1–сторон)	,988(**)
			,000
	объем_продаж	Коэффициент корреляции	,988(**)
		Знч. (1–сторон)	1,000
			,000

** Корреляция значима на уровне 0.01 (1–сторонняя).
а Искл. целиком N = 10

Рис. 2.4.6. Вывод корреляций

Перейдем к построению регрессионной зависимости между показателями.

- Выберем в меню *Analyze* (Анализ) *Regression* (Регрессия) *Linear* (Линейная). Появится диалоговое окно *Linear Regression* (Линейная регрессия) (рис. 2.4.7).

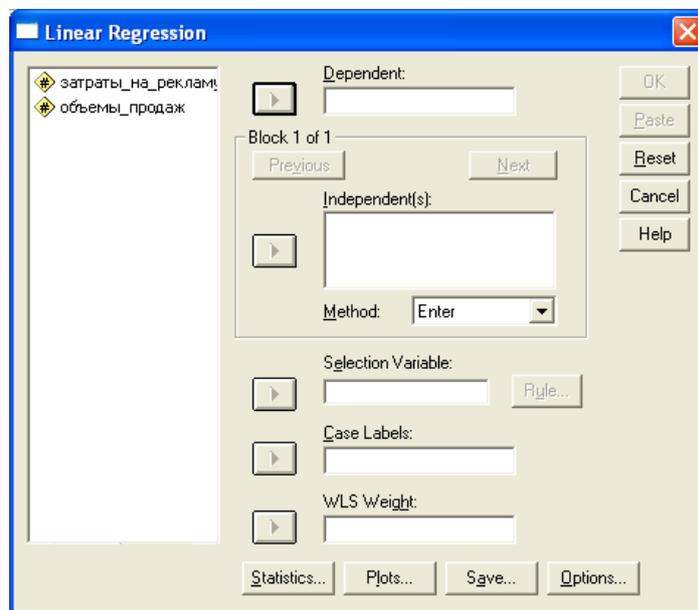


Рис. 2.4.7. Вид диалогового окна *Linear Regression* (Линейная регрессия)

- Перенесем переменную *объемы_продаж* в поле для зависимых переменных и присвоим переменной *затраты_на_рекламу* статус независимой переменной.
- Ничего больше не меняя, начните расчёт нажатием ОК.

Вывод основных результатов выглядит следующим образом (рис. 2.4.8).

Во второй таблице дается заключение о соответствии модели исходным данным, а именно приводится коэффициент детерминации, который характеризует качество получившейся модели.

В третьей таблице приведены величины, которые отражают два источника дисперсии: дисперсию, которая описывается уравнением регрессии (сумма квадратов, обусловленная регрессией) и дисперсию, которая не учитывается при записи уравнения (остаточная сумма квадратов). Также приведено значение *F*-критерия Фишера.

В последней таблице выводятся коэффициент регрессии *b* и смещение по оси ординат *a* под именем «константа».

То есть, уравнение регрессии выглядит следующим образом:

$$y = 2,820 \times x + 380,945,$$

где *y* показатель «Объемы продаж», *x* показатель «Затраты на рекламу».

Включенные/исключенные переменные(b)

Модель	Включенные переменные	Исключенные переменные	Метод
1	затраты_на_рекламу(a)	.	Принудительное включение

а Включены все запрошенные переменные

б Зависимая переменная: объем_продаж

Сводка для модели

Модель	R	R квадрат	Скорректированный R квадрат	Стд. ошибка оценки
1	,983(a)	,966	,962	6,39137

а Предикторы: (константа) затраты_на_рекламу

Дисперсионный анализ (b)

	Сумма квадратов	ст.св.	Средний квадрат	F	Знч.
Регрессия	9257,603	1	9257,603	226,627	,000(a)
Остаток	326,797	8	40,850		
Итого	9584,400	9			

а Предикторы: (константа) затраты_на_рекламу

б Зависимая переменная: объем_продаж

Коэффициенты(a)

Модель	Нестандартизованные коэффициенты	Стд. ошибка	Станд. коэфф.	t	Знч.
			Бета		
(Константа)	380,945	5,049		75,448	,000
затраты_на_рекламу	2,820	,187	,983	15,054	,000

а Зависимая переменная: объем_продаж

Рис. 2.4.8. Результат выполнения процедуры *Analyze – Regression – Linear*

2.5. Дисперсионный анализ

Задание 2.5.

Исследовать влияние работы трёх филиалов предприятия на получение им месячной прибыли. Данные приведены в таблице 2.5.1

Таблица 2.5.1

Филиал предприятия	Прибыль
1	35,8; 43,5; 30,7; 34,0; 33,2; 31,7; 31,6; 34,0; 33,2
2	31,4; 44,5; 43,0; 48,7; 41,6; 43,8; 42,2; 45,0; 41,9
3	32,8; 40,3; 40,5; 45,3; 41,4; 41,3; 40,8; 39,2

Выполнение.

В рамках однофакторного дисперсионного анализа предусмотрена процедура *One-Way ANOVA*, которая заключается в анализе влияния одного качественного фактора на количественную переменную.

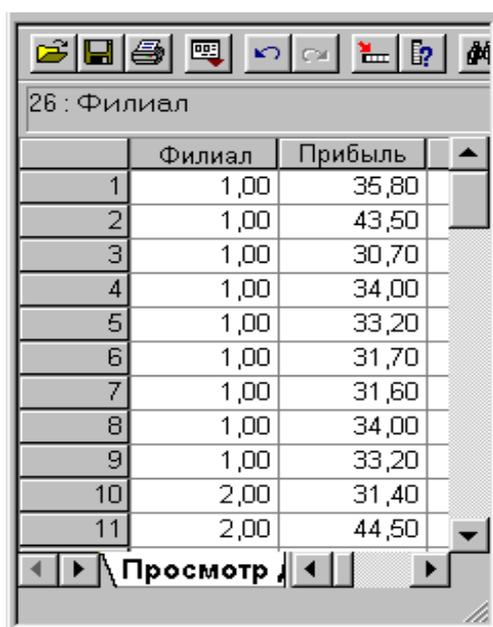
Требуется указать список количественных переменных и фактор, от которого они зависят.

Имеем задачу однофакторного дисперсионного анализа с наблюдениями на трех уровнях.

Качественным показателем, влияющим на количественную переменную (прибыль по филиалам), является номер филиала предприятия.

Выполним следующие действия.

- Создадим и загрузим файл данных (рис. 2.5.1).



26 : Филиал

	Филиал	Прибыль
1	1,00	35,80
2	1,00	43,50
3	1,00	30,70
4	1,00	34,00
5	1,00	33,20
6	1,00	31,70
7	1,00	31,60
8	1,00	34,00
9	1,00	33,20
10	2,00	31,40
11	2,00	44,50

Просмотр

Рис. 2.5.1. Фрагмент файла данных

- Выберем в меню команды *Analyze* (Анализ) *Compare Means* (Сравнение средних) *One–Way ANOVA* (Однофакторный дисперсионный анализ). Появится диалоговое окно односторонний ANOVA (рис. 2.5.2).
- Перенесем переменную *Прибыль* в список зависимых переменных, а переменную *Филиал* – в поле «Фактор».
- Зададим вывод описательной статистики, для этого щелкнем на кнопке «Параметры» и в открывшемся окне (рис. 2.5.3) установим флажок «Описательный».

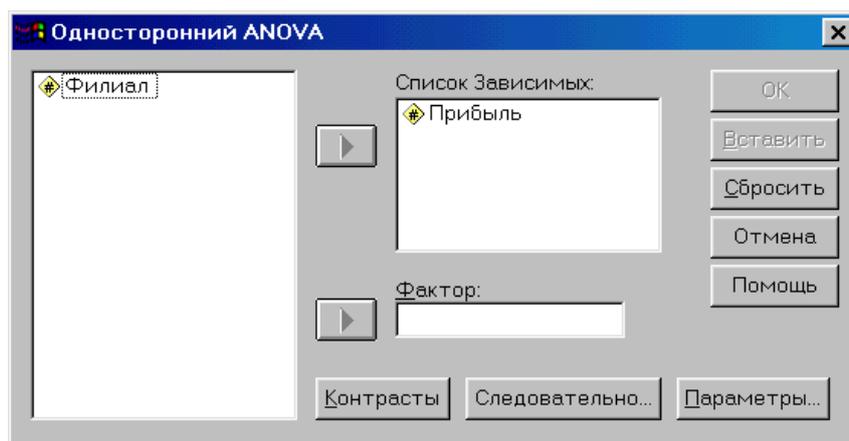


Рис. 2.5.2. Диалоговое окно «Однофакторный дисперсионный анализ»

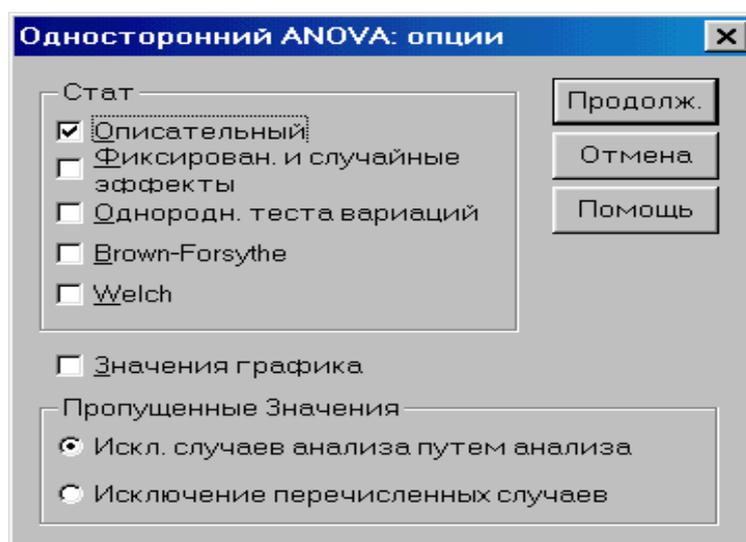


Рис. 2.5.3. Диалоговое окно «Однофакторный дисперсионный анализ: Опции»

Запустим тест, щелкнув на ОК.
Получим следующие результаты.

	N	Mean	Std. deviation	Std. Error	95% Confidence interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
1,00	9	34,1889	3,81426	1,27142	31,2570	37,1208	30,70	43,50
2,00	9	42,4556	4,67817	1,55939	38,8596	46,0515	31,40	48,70
3,00	8	40,2000	3,48220	1,23114	37,2888	43,1112	32,80	45,30
Total	26	38,9000	5,30434	1,04027	36,7575	41,0425	30,70	48,70

ANOVA

Прибыль

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	327,049	2	163,524	9,993	,001
Within Groups	376,351	23	16,363		
Total	703,400	25			

Рис. 2.5.4. Результат выполнения процедуры

Таким образом, имеем следующие характеристики:

- $n_1 = 9; n_2 = 9; n_3 = 8$.

- Средние значения переменной на i -м уровне $\bar{y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij}$.

$$\bar{y}_1 = 34,1889; \quad \bar{y}_2 = 42,4556; \quad \bar{y}_3 = 40,2.$$

- Среднее значение переменной по всем значениям

$$\bar{y} = \frac{1}{n} \sum_{i=1}^k \sum_{j=1}^{n_i} y_{ij} = 38,9.$$

- Сумму квадратов отклонений всех наблюдений от общего

$$\text{среднего } S = \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y})^2 = 703,4.$$

- Сумму квадратов отклонений средних групповых значений

$$S_F = \sum_{i=1}^k (\bar{y}_i - \bar{y})^2 n_i = 327,049.$$

- Остаточную сумму квадратов отклонений

$$S_{OCT} = \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2 = 376,351.$$

Отметим справедливость соотношения $S = S_F + S_{OCT}$.

Нулевая гипотеза принимается при $F_H < F_{кр} = F(1 - \alpha, k - 1, n - k)$

Проверим гипотезу для уровня 0,05: $F_{кр} = 3,443$; $F_H = 9,993$.

Имеем, что $F_H > F_{кр}$, нулевую гипотезу отвергаем и делаем вывод, что работа в каждом из филиалов влияет на месячную прибыль предприятия.

Наглядным представлением результатов являются графики средних значений и их доверительных интервалов (простая диаграмма величины ошибки) (рис.2.5.7).

Построим подобный график, для этого выполним следующие действия:

- Выберем команду «Визуализация/Колонка ошибок» и в открывшемся окне (рис. 2.5.5) выберем вариант «Простой» и нажмем на кнопку «Определ.».
- В появившемся диалоговом окне (рис. 2.5.6) заполним следующие поля. Переменная в рассматриваемом случае это переменная *Прибыль*. Ось категорий – поле факторной переменной (переменной, содержащей категории), для данного примера это переменная *Филиал*. Изобразить панели – что следует отразить на графике: доверительный интервал для математического ожидания, стандартную ошибку математического ожидания или среднеквадратичное отклонение. Нас интересует доверительный интервал для математического ожидания. Уровень – пределы доверительного интервала (по умолчанию стоит значение 95%).

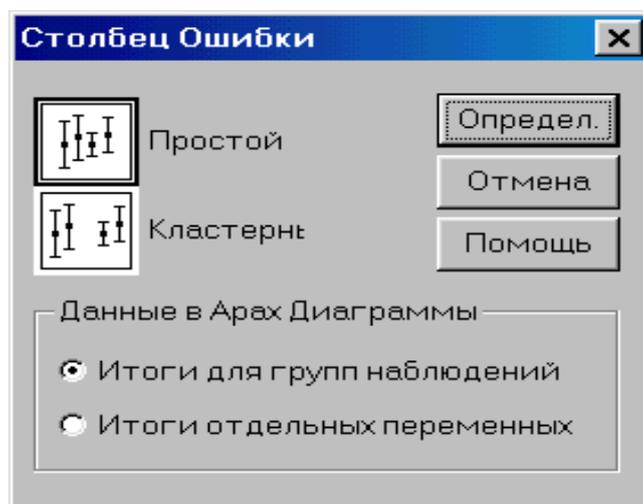


Рис. 2.5.5. Диалоговое окно «Колонка Ошибок»

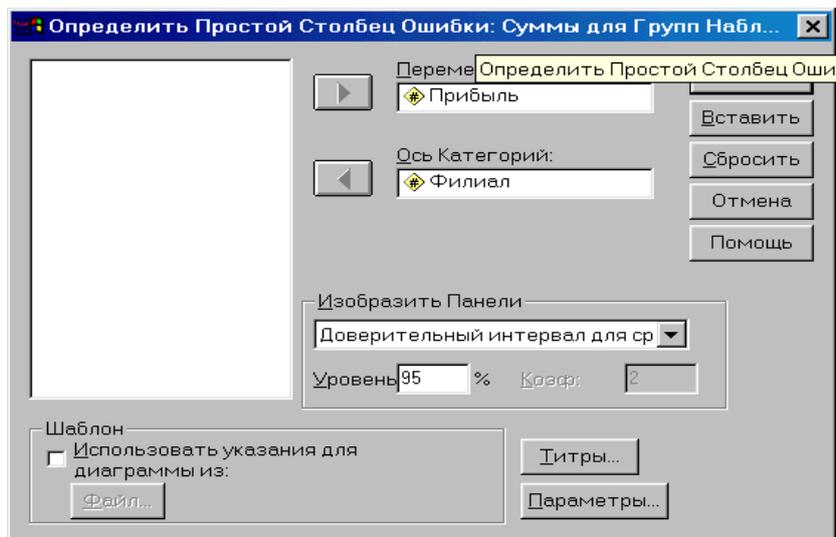


Рис. 2.5.6. Диалоговое окно

Запустим выполнение, щелкнув на ОК.
Получим следующий график (рис. 2.5.7).

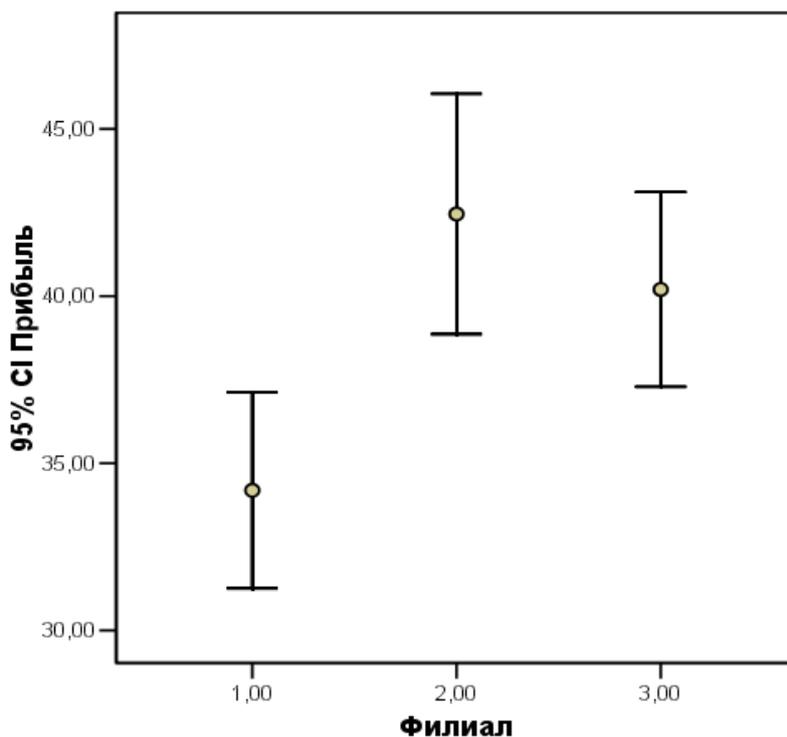


Рис. 2.5.7. Результат выполнения процедуры «Визуализация/Колонка ошибок»

2.6. Ряды динамики

Задание 2.6.

Необходимо выявить сезонную компоненту (расход воды жителями какого-то города).

Выполнение.

Для этого необходимо ввести данные, как показано на рис. 2.6.1.

Затем выбираем:

«Анализ – Временные ряды – Сезонное разложение» (рис.2.6.2).

Далее выбираем модель мультипликативную (умножение), перетаскиваем нашу исследуемую величину (рис.2.6.3).

Как видно, здесь уже автоматически подсчитана периодичность.

Нажимаем ОК.

Получаем вывод (рис.2.6.4).

Если вернуться в редактор данных, то увидим, что к исходным данным добавились еще переменные: *ERR_1*, *SAS_1*, *SAF_1*, *STC_1* (рис. 2.6.5).

Здесь:

первая переменная – это отклонение объема потребления от средней за год,

вторая – сезонная характеристика ряда,

третья – сезонный фактор, четвертая – тренд.

Теперь сделаем все тоже, но только с аддитивной моделью.

Для этого в диалоговом окне (рис.2.6.6) отметим положительная модель.

В результате получим следующий вывод рис.2.6.7.

Если вернуться в редактор данных, то снова можно убедиться в появлении новых переменных: *ERR_2*, *SAS_2*, *SAF_2*, *STC_2*, имеющих аналогичный смысл что и *ERR_1*, *SAS_1*, *SAF_1*, *STC_1*.

лаборраб_6.sav - SPSS Редактор данных

Файл Редактировать Отображение Данные Изменение А

0 : объем_потреблени

	объем пот	year	month	date
1	80,50	1968	1	JAN 1968
2	84,60	1968	2	FEB 1968
3	126,60	1968	3	MAR 1968
4	162,00	1968	4	APR 1968
5	140,90	1968	5	MAY 1968
6	137,90	1968	6	JUN 1968
7	139,80	1968	7	JUL 1968
8	136,60	1968	8	AUG 1968
9	134,30	1968	9	SEP 1968
10	140,80	1968	10	OCT 1968
11	127,10	1968	11	NOV 1968
12	96,40	1968	12	DEC 1968
13	101,50	1969	1	JAN 1969
14	90,10	1969	2	FEB 1969
15	131,90	1969	3	MAR 1969
16	159,00	1969	4	APR 1969
17	155,50	1969	5	MAY 1969
18	147,30	1969	6	JUN 1969
19	125,20	1969	7	JUL 1969
20	124,90	1969	8	AUG 1969
21	129,30	1969	9	SEP 1969
22	123,40	1969	10	OCT 1969
23	94,60	1969	11	NOV 1969
24	84,10	1969	12	DEC 1969
25	66,40	1970	1	JAN 1970
26	74,30	1970	2	FEB 1970
27	114,70	1970	3	MAR 1970
28	128,40	1970	4	APR 1970
29	125,00	1970	5	MAY 1970
30	135,20	1970	6	JUN 1970
31	140,80	1970	7	JUL 1970
32	128,70	1970	8	AUG 1970

Обзор данных / Обзор переменных /

SPSS Процессор

Рис. 2.6.1. Редактор данных (выделенный столбец – y_t)

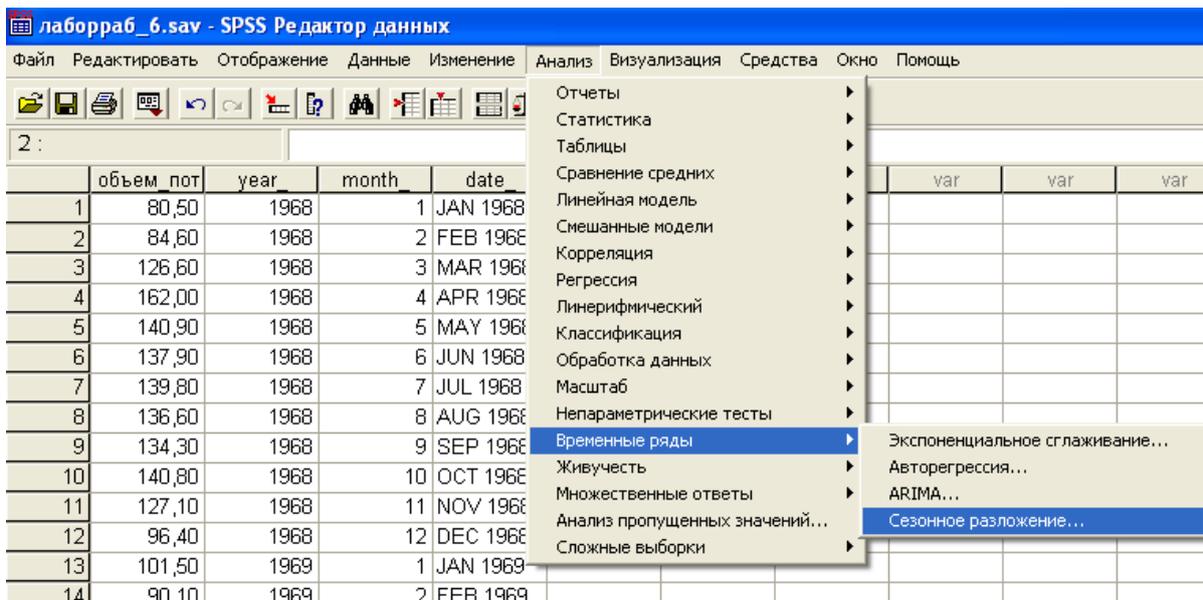


Рис. 2.6.2. Выбор необходимого пункта меню

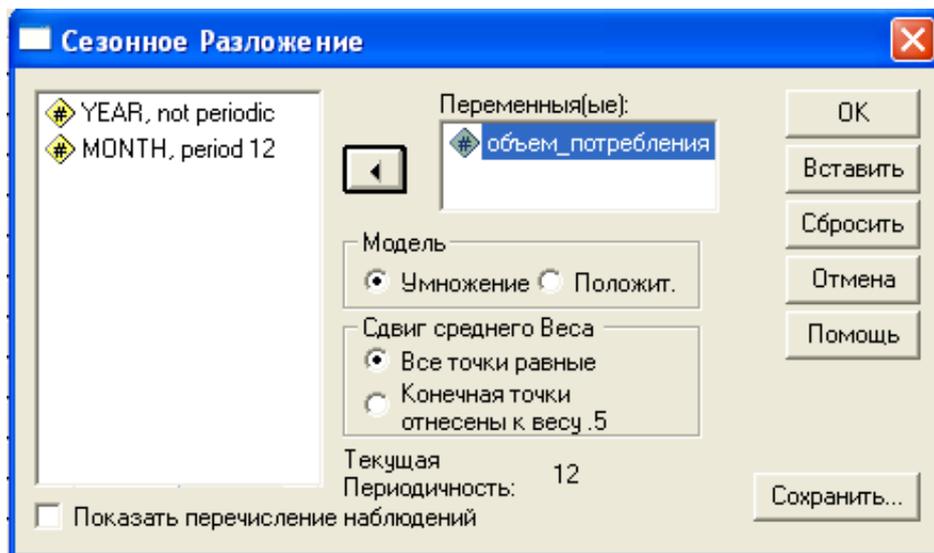


Рис. 2.6.3 Диалоговое окно «Сезонное Разложение»

Seasonal Decomposition Model Description

Model Name	MOD_4
Model Type	Multiplicative
Series Name	1 объем_потребления
Length of Seasonal Period	12
Computing Method of Moving Averages	Span equal to the periodicity and all points weighted equally

Applying the model specifications from MOD_4

Seasonal Factors

Series Name: объем_потребления

PPeriod	Seasonal Factor (%)
1	65,6
2	69,4
3	101,1
4	114,6
5	120,4
6	119,4
7	112,5
8	113,5
9	106,3
10	110,2
11	92,2
12	74,9

Рис.2.6.4. Результат выполнения «Сезонного Разложения»
(мультипликативная модель)

The screenshot shows the SPSS Trends editor interface. The main window title is "Trends chapter 14.sav - SPSS Редактор данных". The menu bar includes "Файл", "Редактировать", "Отображение", "Данные", "Изменение", "Анализ", "Визуализация", "Средства", "Окно", "Помощь". The toolbar contains various icons for file operations and analysis. The data table below shows the results of the seasonal decomposition for the series "объем_потребления".

	объем_потребления	year_	month_	date_	ERR_1	SAS_1	SAF_1	STC_1
1	80,50	1968	1	JAN 1968	1,01799	122,68586	,65615	120,51765
2	84,60	1968	2	FEB 1968	,98907	121,95442	,69370	123,30241
3	126,60	1968	3	MAR 1968	,98707	125,26696	1,01064	126,90738
4	162,00	1968	4	APR 1968	1,11022	141,38466	1,14581	127,34883
5	140,90	1968	5	MAY 1968	,94505	117,00155	1,20426	123,80497
6	137,90	1968	6	JUN 1968	,95305	115,49405	1,19400	121,18303
7	139,80	1968	7	JUL 1968	1,02779	124,21574	1,12546	120,85722
8	136,60	1968	8	AUG 1968	,97972	120,34586	1,13506	122,83717
9	134,30	1968	9	SEP 1968	,99991	126,38639	1,06261	126,39777
10	140,80	1968	10	OCT 1968	,99064	127,79866	1,10173	129,00584
11	127,10	1968	11	NOV 1968	1,02766	137,91604	,92158	134,20382
12	96,40	1968	12	DEC 1968	,94251	128,70584	,74899	136,55699
13	101,50	1969	1	JAN 1969	1,11406	154,69087	,65615	138,85303
14	90,10	1969	2	FEB 1969	,95228	129,88289	,69370	136,39167
15	131,90	1969	3	MAR 1969	,96862	130,51116	1,01064	134,73869
16	159,00	1969	4	APR 1969	1,05053	138,76642	1,14581	132,09129
17	155,50	1969	5	MAY 1969	1,00757	129,12520	1,20426	128,15515
18	147,30	1969	6	JUN 1969	1,00969	123,36674	1,19400	122,18241
19	125,20	1969	7	JUL 1969	,95229	111,24328	1,12546	116,81618
20	124,90	1969	8	AUG 1969	,96025	110,03805	1,13506	114,59276
21	129,30	1969	9	SEP 1969	1,07048	121,68101	1,06261	113,66927
22	123,40	1969	10	OCT 1969	1,00104	112,00536	1,10173	111,88896
23	94,60	1969	11	NOV 1969	,94328	102,65034	,92158	108,82302
24	84,10	1969	12	DEC 1969	1,04867	112,28383	,74899	107,07307
25	66,40	1970	1	JAN 1970	,95019	101,19679	,65615	106,50152

Рис. 2.6.5. В редакторе данных появились новые переменные

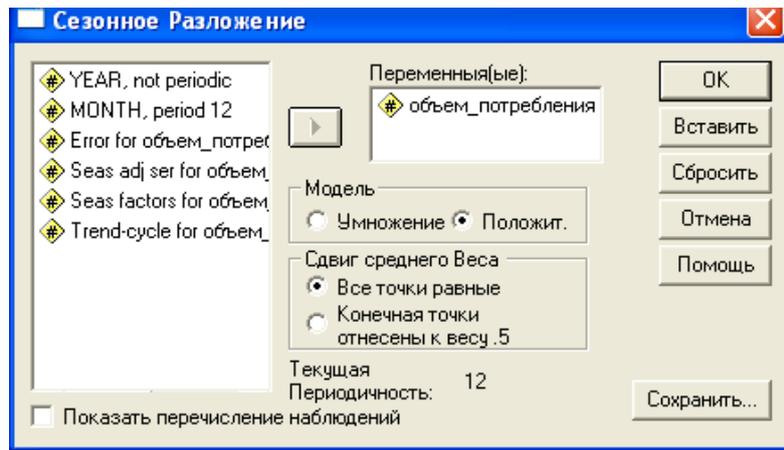


Рис. 2.6.6. Диалоговое окно «Сезонное Разложение»

Seasonal Decomposition Model Description

Model Name		MOD_5
Model Type		Additive
Series Name	1	объем_потребления
Length of Seasonal Period		12
Computing Method of Moving Averages		Span equal to the periodicity and all points weighted equally

Applying the model specifications from MOD_5

Seasonal Factors Series Name: объем_потребления

PPeriod	Seasonal Factor
1	-47,45827
2	-42,74865
3	1,60455
4	19,35391
5	27,00391
6	27,57491
7	17,48019
8	19,13917
9	8,60519
10	15,04109
11	-10,75506
12	-34,84096

Рис. 2.6.7. Результат выполнения «Сезонного Разложения»
(аддитивная модель)

2.7. Индексный метод

Задание 2.7.

Имеются данные о производстве продукции и численности работников фирмы по двум филиалам.

Таблица 2.7.1

Филиалы фирмы	Базовый период		Отчетный период	
	Продукция, млн. руб.	Рабочие, чел.	Продукция, млн. руб.	Рабочие, чел.
1	150	100	140	130
2	110	90	135	80

Определить индексы производительности труда постоянного состава, переменного состава и структурных сдвигов.

Выполнение.

Путем вычислений в SPSS можно образовать новые переменные и добавить их в файл данных.

Таким образом, работу со сложными вычислениями можно переложить на компьютер, который сделает ее быстро и, главное, без ошибок.

Для этого поступим для решения рассматриваемой задачи следующим образом

Прежде всего, определим переменные, как представлено на рис. 2.7.1 и введем значения переменных (см. рис. 2.7.2).

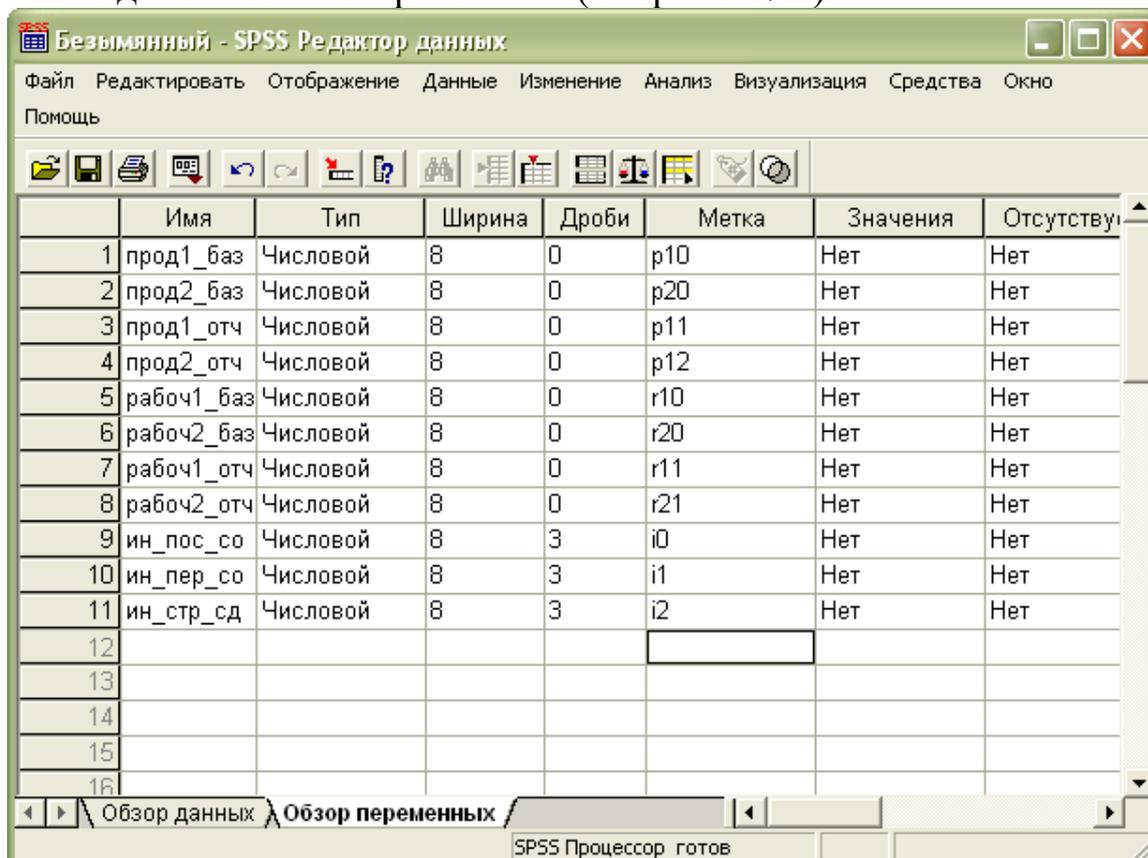


Рис. 2.7.1. Определение переменных

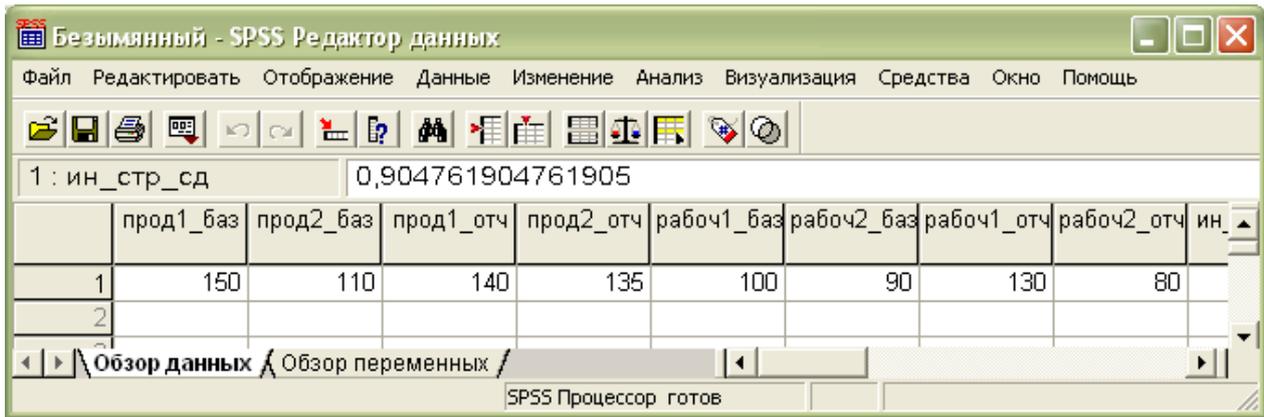


Рис. 2.7.2. Ввод значений переменных

Далее в меню выберем *Transform* (Изменение) *Compute* (Просчет).

Откроется диалоговое окно *Compute Variable* (Вычислить переменную) уже со списком определенных переменных (см. рис. 2.7.3).

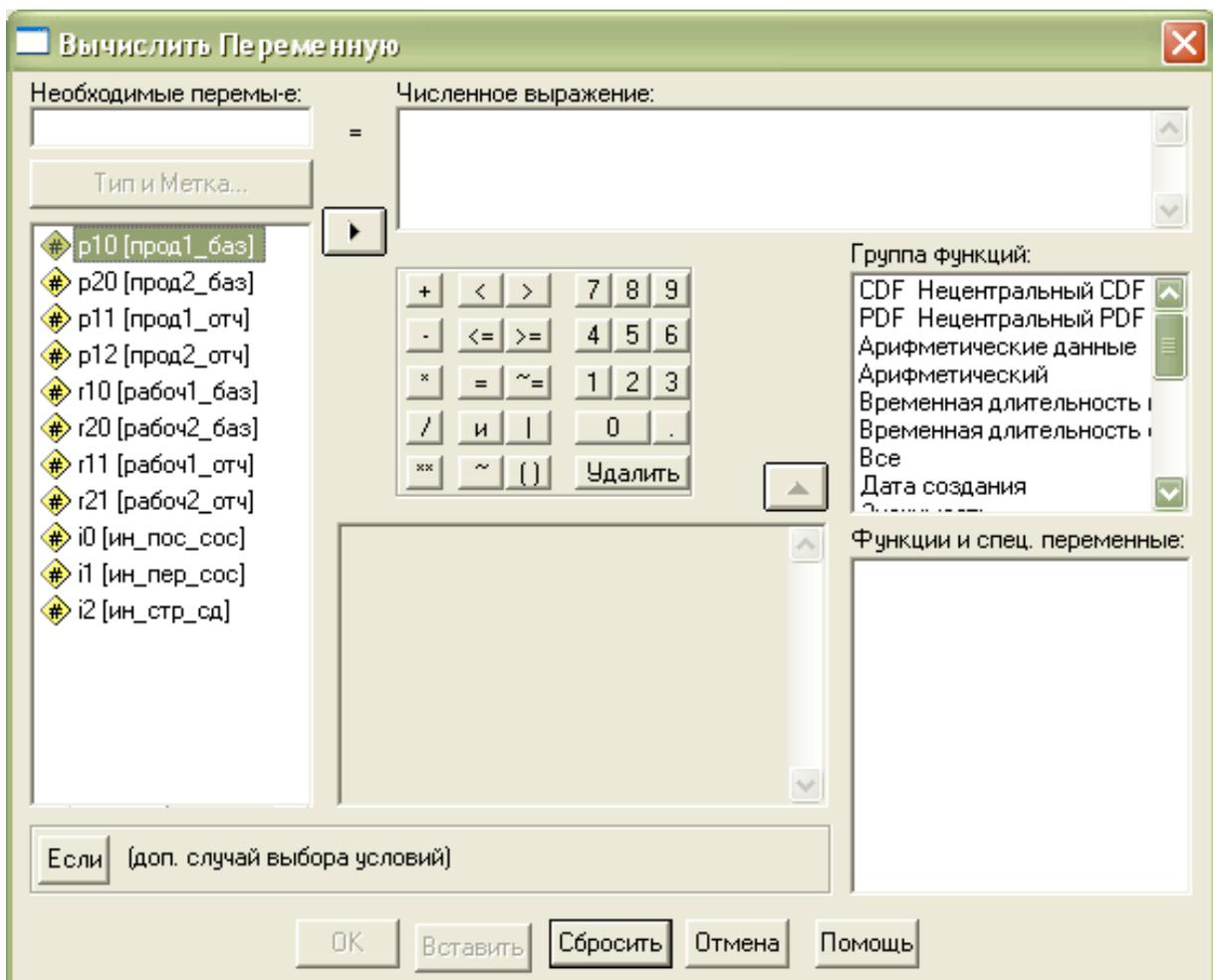


Рис. 2.7.3. Диалоговое окно «Вычислить переменную»

В поле *Target Variable* (Необходимая переменная) указывается имя переменной, которой присваивается вычисленное значение.

В качестве выходной переменной может служить уже существующая или новая переменная.

В поле *Numeric Expression* (Численное выражение) вводится выражение, применяемое для определения значения выходной переменной.

В этом выражении могут использоваться имена существующих переменных, константы, арифметические операторы и функции.

Далее в поле *Numeric Expression* (Численное выражение) создаем выражения для последних трех переменных (индексы производительности труда постоянного состава, переменного состава и структурных сдвигов).

При этом используем функцию *SUM* из списка справа.

Данные операции представлены на рис. 2.7.4 – 2.7.6.

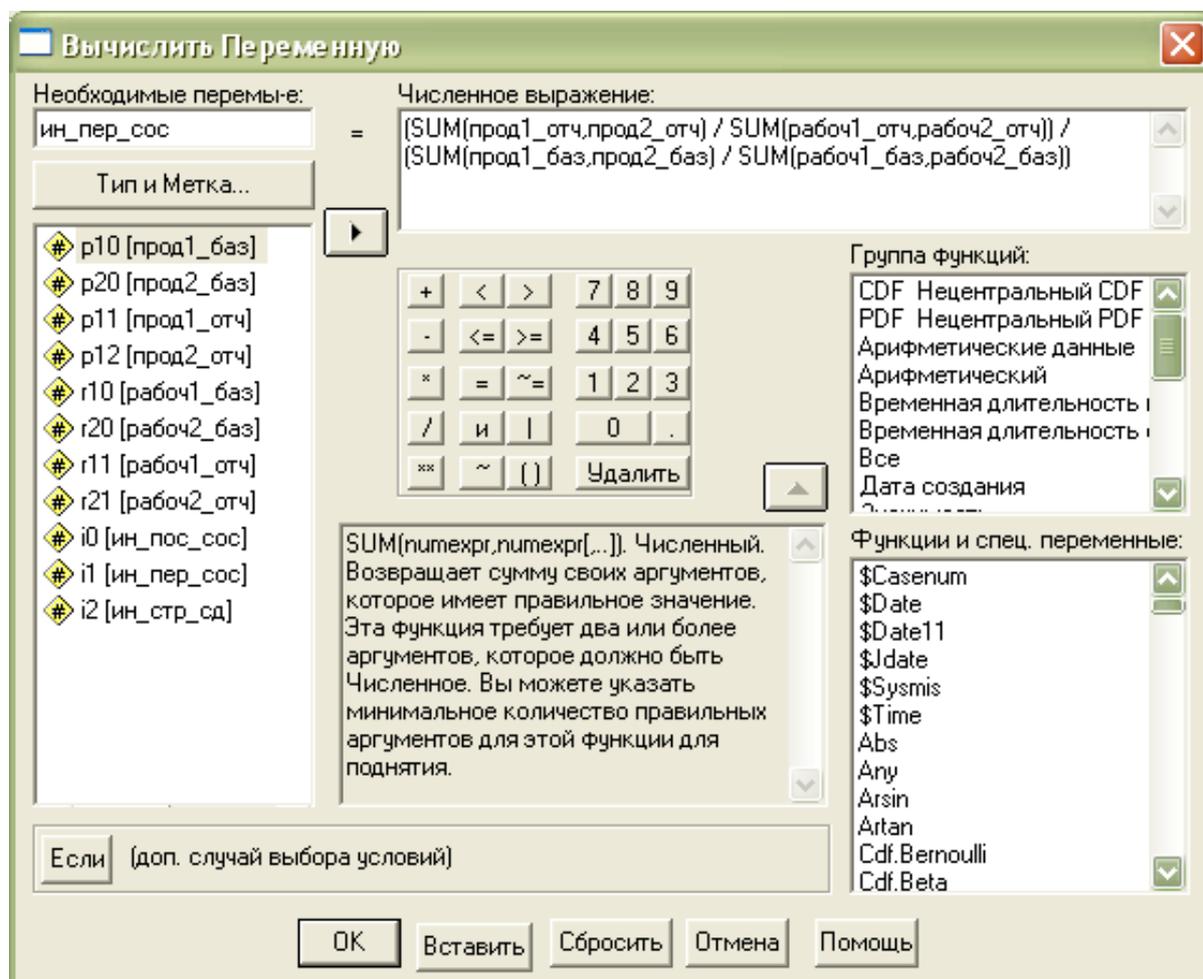


Рис. 2.7.4. Определение индекса переменного состава

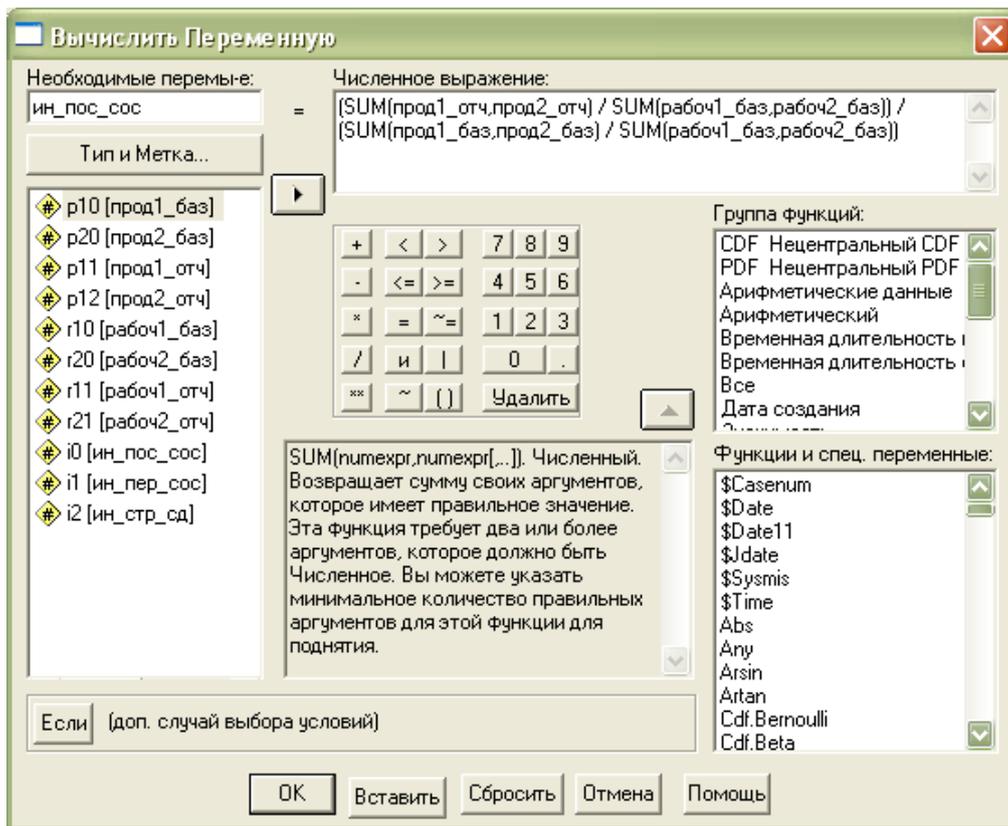


Рис. 2.7.5. Определение индекса постоянного состава

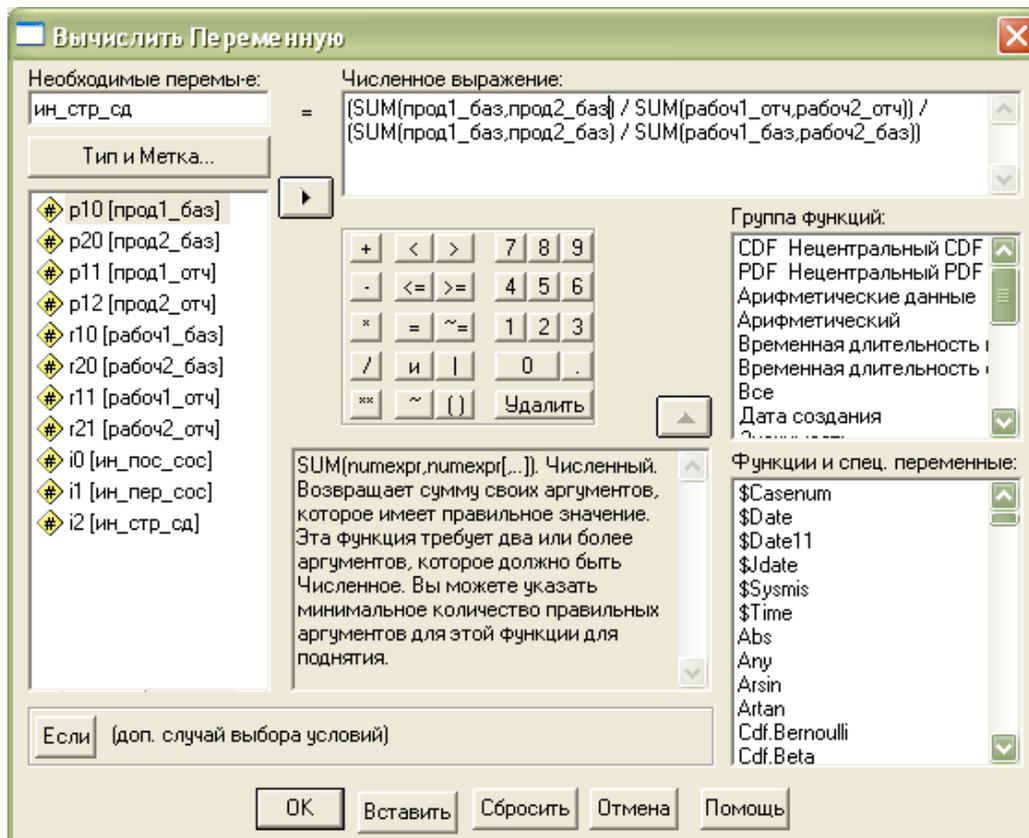


Рис. 2.7.6. Определение индекса структурных сдвигов

Таким образом, вычисляются значения неизвестных индексов (см. рис. 2.7.8).

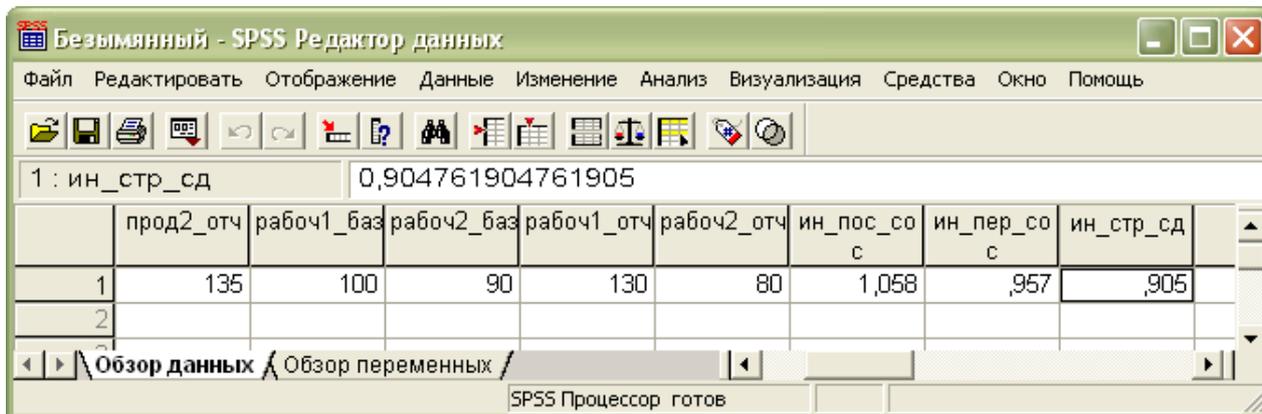


Рис. 2.7.8. Результаты вычисления переменных

2.8. Кластерный анализ

Задание 2.8.

Провести кластерный анализ 20 сортов картофеля по двум показателям: процентное содержание крахмала и процентное содержание белка урожайность в нем.

Данные для анализа приведены в таблице 2.8.1.

Таблица 2.8.1

№ п/п	1	2	3	4	5	6	7	8	9	10
Крахмал (%)	28,6	29,4	25,3	19,4	19,6	18,8	26,9	25,7	24,6	24,7
Белок (%)	1,6	1,5	3,2	4,5	2,8	4,4	3,2	3,3	3	3,4
№ п/п	11	12	13	14	15	16	17	18	19	20
Крахмал (%)	20,1	26,7	25,9	27,5	18,5	20,7	28,3	26,9	17,8	18
Белок (%)	2,8	1,9	3	1,6	4,1	2,7	2	1,3	4,1	4,2

Выполнение.

Загрузим файл данных. Проведем простейший кластерный анализ имеющихся данных. Для этого построим диаграмму рассеяния.

- Выберем в меню «Визуализация – Интерактив – Диаграмма рассеяния». Появится диалоговое окно «Диаграмма рассеяния» (рис. 2.8.1). Выбираем «Простой» – «Определ.».

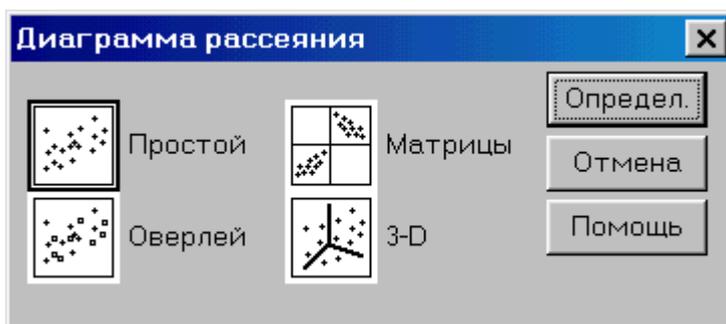


Рис. 2.8.1 Вид диалогового окна «Диаграмма рассеяния»

Появится диалоговое окно «Простой график рассеяния» (рис. 2.8.2).

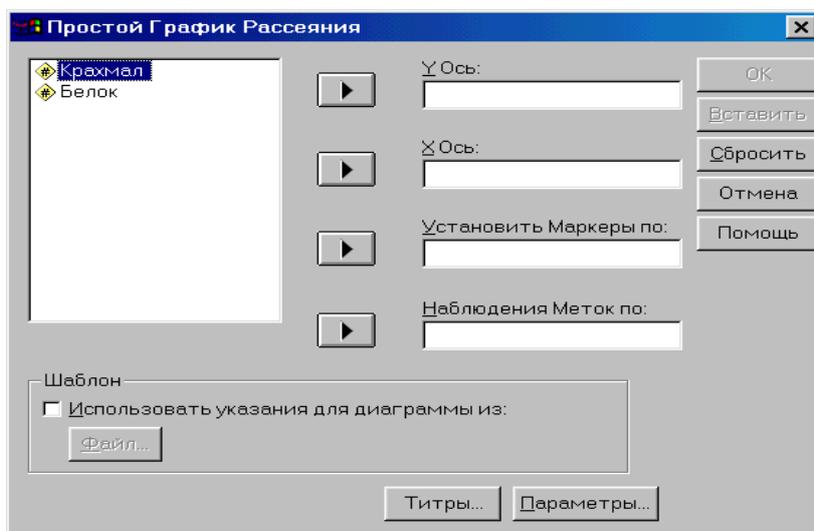


Рис. 2.8.2 Вид диалогового окна «Простой график рассеяния»

- Переменную *Белок* поместим в поле оси *X*, а переменную *Крахмал* в поле оси *Y*.
- Ничего больше не меняя, начнем расчёт нажатием ОК.

Получим диаграмму рассеяния, представленную на рис. 2.8.3.

При помощи диаграммы рассеяния для двух переменных: *Крахмал* и *Белок*, был проведен самый простой кластерный анализ.

Был выбран такой вид графического представления, с помощью которого можно отчётливо распознать группирование в кластеры.

В рассматриваемом случае по виду диаграммы можно увидеть, что наблюдения сгруппировались в четыре кластера.

Далее будем проводить иерархический кластерный анализ.

В иерархических методах каждое наблюдение образует сначала свой отдельный кластер.

На первом шаге два соседних кластера объединяются в один; этот процесс может продолжаться до тех пор, пока не останутся только два кластера.

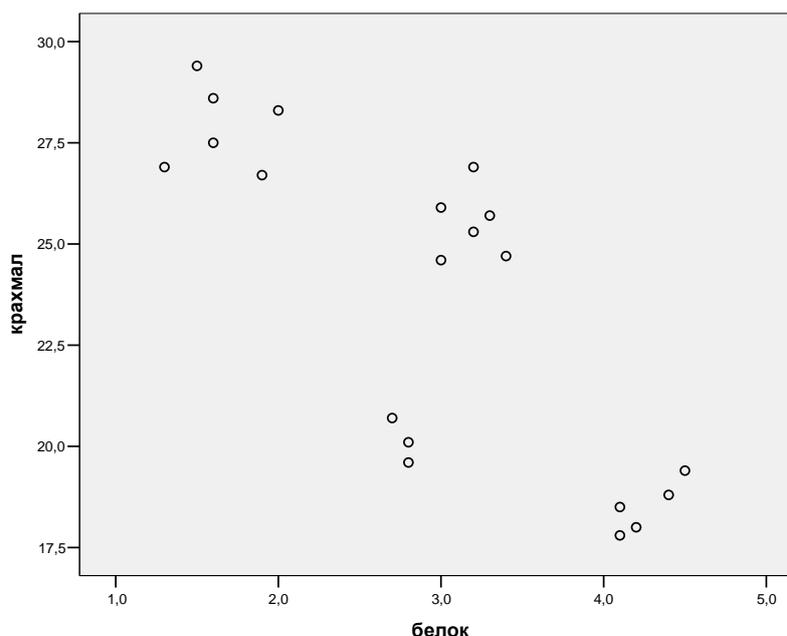


Рис. 2.8.3. Диаграмма рассеяния переменных *Крахмал* и *Белок*

Соберём имеющиеся 20 сортов картофеля в кластеры, используя имеющиеся параметры *Крахмал* и *Белок*.

Выберем в меню Анализ – Классификация – Иерархический кластерный анализ. Появится диалоговое окно «Иерархический Анализ Кластера» (см. рис. 2.8.4).

Переменные *Крахмал* и *Белок* поместите в поле «Переменные».

- Щелчком по выключателю «Статистика» откроем диалоговое окно «Иерархический кластерный анализ: Статистика» (рис. 2.8.5). Оставим флажок напротив опции «Режим накопления». Активируем опцию «Диапазон решений» и введем числа 3 и 5 в качестве границ области. (Хотя на основании графического представления на диаграмме рассеяния (рис. 2.8.3) и ожидается результат в виде четырёх кластеров, но не можем быть полностью уверены в достижении этого результата).
- Вернувшись в главное диалоговое окно, щёлкните по выключателю «Графики». В появившемся диалоговом окне (рис. 2.8.6) активируем опцию вывода «Древовидной диаграммы» и посредством опции «Нет» отмените вывод ориентации графика.
- С помощью кнопки «Метод» возможно выбрать метод образования кластеров, а также метод расчета дистанционной меры и меры подобия соответственно. Здесь (рис. 2.8.7) в поле «Преобразование значений» установите «Множества–z (стандартизацию) значений».

- Вернемся назад в главное диалоговое окно и начните расчёт нажатием ОК.

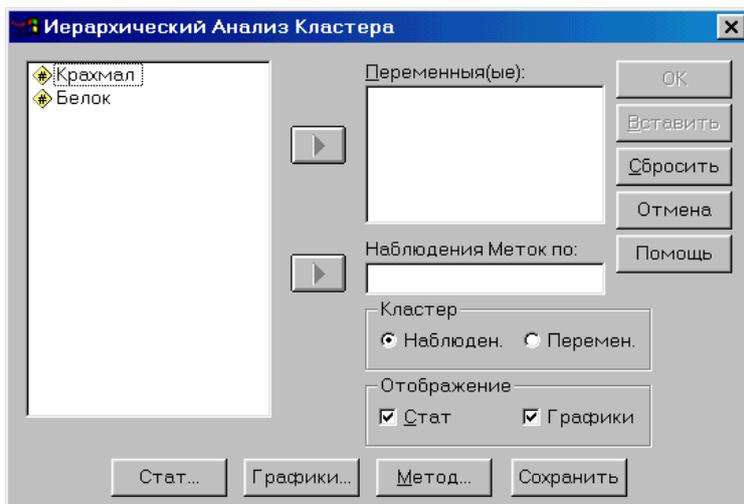


Рис. 2.8.4. Вид диалогового окна «Иерархический кластерный анализ»

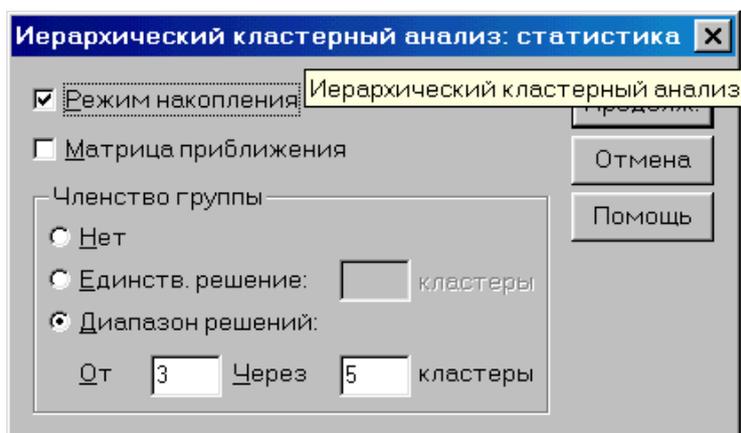


Рис. 2.8.5. Вид диалогового окна «Иерархический кластерный анализ: Статистики»

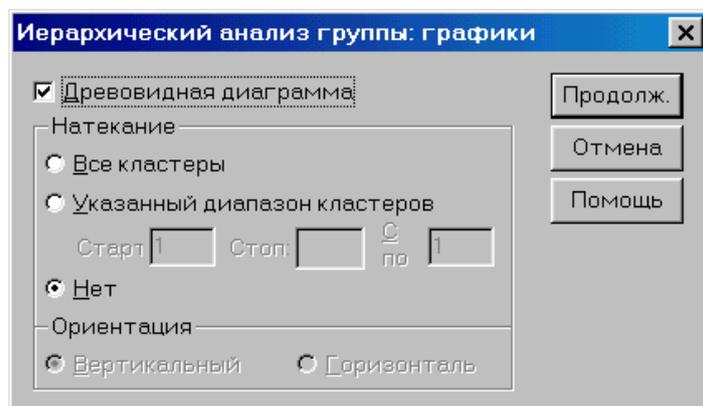


Рис. 2.8.6. Вид диалогового окна «Иерархический кластерный анализ: Диаграммы»

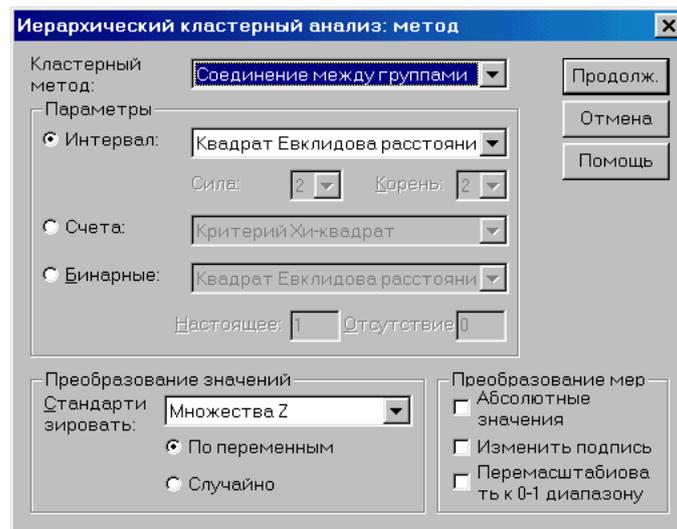


Рис. 2.8.7. Вид диалогового окна «Иерархический кластерный анализ: Метод»

Вывод основных результатов расчета выглядит следующим образом (рис. 2.8.8).

В таблице *Agglomeration Schedule* (Порядок агломерации) можно выяснить очерёдность построения кластеров, а также их оптимальное количество.

По двум колонкам, расположенным под общей шапкой *Cluster Combined* (Объединение в кластеры), можно увидеть, что на первом шаге были объединены наблюдения 19 и 20.

Эти два сорта картофеля максимально похожи друг на друга и отделены друг от друга очень малое расстояние.

Эти два наблюдения образуют кластер с номером 1, в то время как кластер 20 в обзорной таблице больше не появляется.

На следующем шаге происходит объединение наблюдений 5 и 11, затем 3 и 8 и т.д.

Под этим коэффициентом подразумевается расстояние между двумя кластерами, определенное на основании выбранной дистанционной меры с учётом предусмотренного преобразования значений.

В нашем случае это квадрат евклидова расстояния, определенный с использованием стандартизованных значений (в примере использовалось Множество–z).

На этом этапе, где эта мера расстояния между двумя кластерами увеличивается скачкообразно, процесс объединения в новые кластеры необходимо остановить, так как в противном случае были бы объединены уже кластеры, находящиеся на относительно большом расстоянии друг от друга.

Agglomeration Schedule							Cluster Membership			
Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage	Case	5 Clusters	4 Clusters	3 Clusters
	Cluster 1	Cluster 2		Cluster 1	Cluster 2		1:Case 1	2:Case 2	3:Case 3	4:Case 4
1	19	20	,012	0	0	4	1	1	1	1
2	5	11	,015	0	0	7	1	1	1	1
3	3	8	,019	0	0	8	2	2	2	2
4	15	19	,027	0	1	13	3	3	3	3
5	4	6	,032	0	0	13	4	4	4	2
6	1	2	,049	0	0	15	6	3	3	3
7	5	16	,058	2	0	17	7	2	2	2
8	3	10	,066	3	0	11	8	2	2	2
9	7	13	,100	0	0	14	9	2	2	2
10	14	18	,109	0	0	15	10:Case 10	2	2	2
11	3	9	,128	8	0	14	11:Case 11	4	4	2
12	12	17	,166	0	0	16	12:Case 12	5	1	1
13	4	15	,173	5	4	19	13:Case 13	2	2	2
14	3	7	,181	11	9	17	14:Case 14	1	1	1
15	1	14	,247	6	10	16	15:Case 15	3	3	3
16	1	12	,330	15	12	18	16:Case 16	4	4	2
17	3	5	2,013	14	7	18	17:Case 17	5	1	1
18	1	3	3,524	16	17	19	18:Case 18	1	1	1
19	1	4	7,015	18	13	0	19:Case 19	3	3	3
							20:Case 20	3	3	3

***** HIERARCHICAL CLUSTER ANALYSIS *****

Dendrogram using Average Linkage (Between Groups)

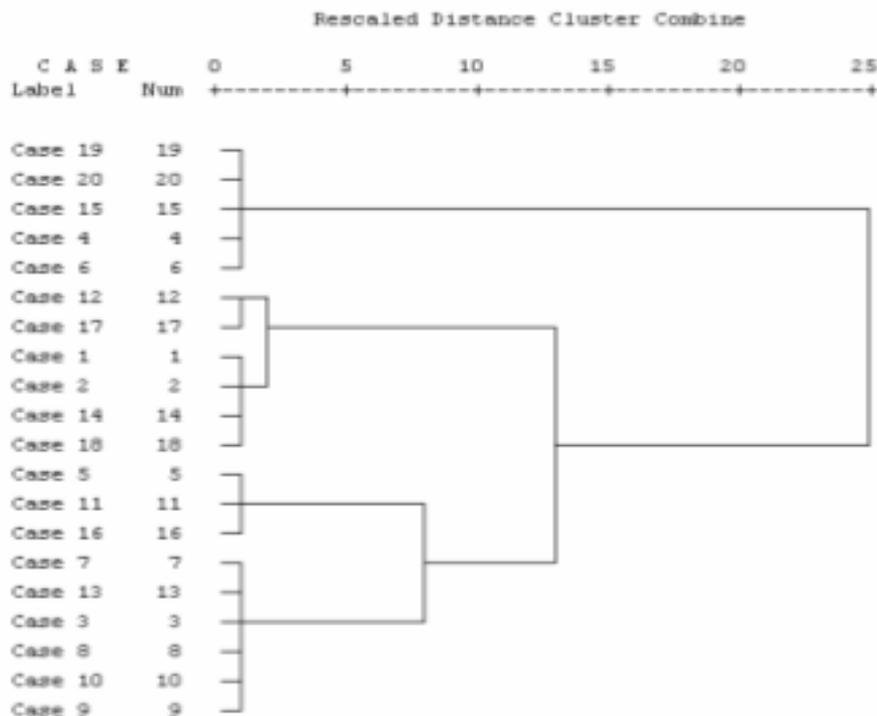


Рис. 2.8.8. Результат выполнения процедуры

Для определения, какое количество кластеров следовало бы считать оптимальным, решающее значение имеет показатель, выводимый под заголовком *Coefficients* (Коэффициент).

В приведенном примере – это скачок с 0,33 до 2,013. Это означает, что после образования четырех кластеров мы больше не должны производить никаких последующих объединений, а результат с четырьмя кластерами является оптимальным.

Оптимальным считается число кластеров равное разности количества наблюдений (здесь: 20) и количества шагов, после которого коэффициент увеличивается скачкообразно (здесь: 16).

В таблице *Cluster Membership* (Принадлежность к кластеру) приведена информация о принадлежности каждого наблюдения кластеру, для результатов расчёта содержащих 5, 4 и 3 кластера.

В заключение приводится затребованная нами древовидная диаграмма, которая визуализирует процесс слияния, приведенный в обзорной таблице порядка агломерации.

Она идентифицирует объединённые кластеры и значения коэффициентов на каждом шаге.

При этом отображаются не исходные значения коэффициентов, а значения, приведенные к шкале от 0 до 25.

2.9. Двухфакторный дисперсионный анализ

Задание 2.9.

Провести факторный анализ сдачи зачетной сессии 20 студентов по четырем предметам. Данные приведены в таблице 2.9.1.

Таблица 2.9.1

№ п/п	1	2	3	4	5	6	7	8	9	10
социология	72	76	64	95	87	78	64	74	82	69
информатика	74	75	71	61	66	93	67	68	71	61
культурология	61	61	65	76	70	62	51	59	70	51
№ п/п	11	12	13	14	15	16	17	18	19	20
социология	72	79	62	70	62	79	92	75	90	94
информатика	65	78	72	52	77	89	93	55	94	78
культурология	58	64	50	67	50	63	74	60	72	75

Выполнение.

- Создадим файл данных, присваивая каждой переменной результаты успеваемости студентов по конкретному предмету.
- Выберем в меню *Analyze* (Анализ) *Data Reduction* (Обработка данных) *Factor* (Факторный анализ). Появится диалоговое окно *Factor Analysis* (Факторный анализ) (рис. 2.9.1).

- Переменные «социология», «информатика», «культурология», «математика» поместим в поле исследуемых переменных.
- После чего, нажав кнопку *Descriptives* (Описание) в появившемся окне установим флажок напротив опции *Coefficients* (Коэффициенты), далее *Continue* (рис. 2.9.2).
- Далее нажмем кнопку *Rotation* (Вращение), которая позволяет выбрать метод вращения. Активируем метод Варимакса, установив флажок напротив опции *Varimax* и оставим активированным вывод повернутой матрицы факторов (рис. 2.9.3).
- Ничего больше не меняя, начнем расчёт нажатием ОК.

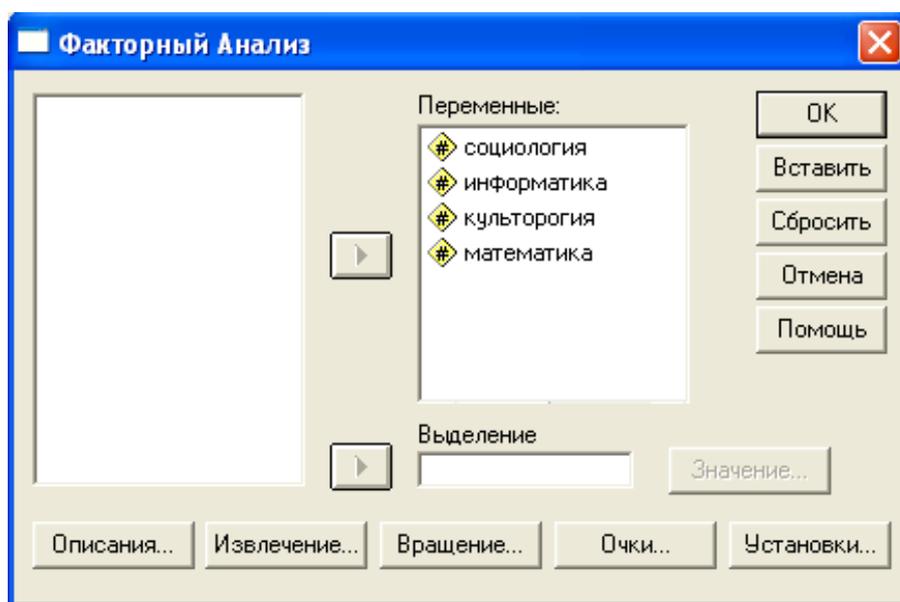


Рис. 2.9.1. Вид диалогового окна «Факторный анализ»

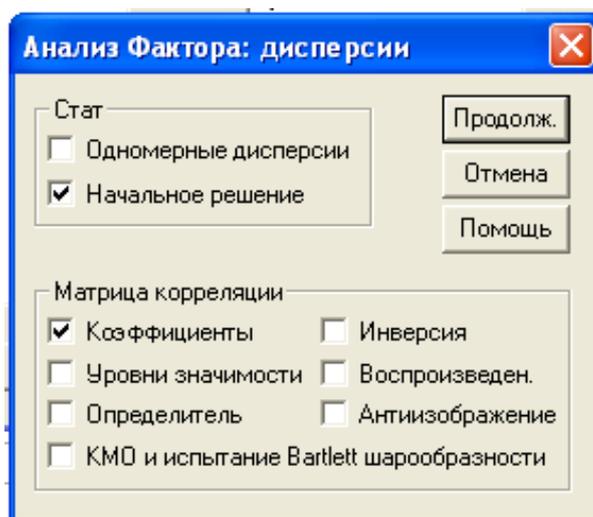


Рис. 2.9.2. Вид диалогового окна «Факторный анализ: Описание»

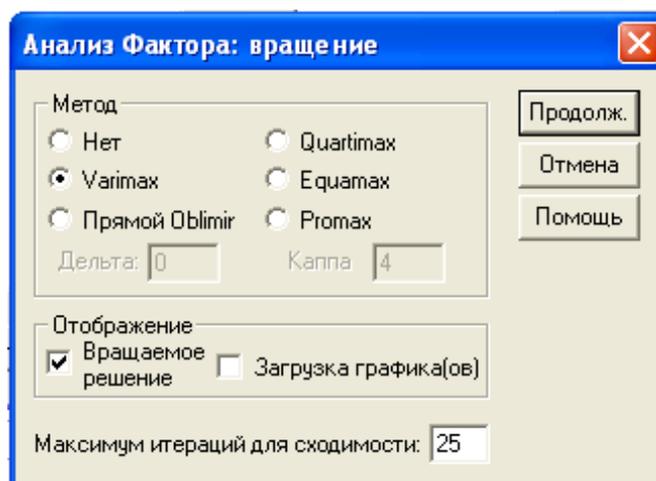


Рис. 2.9.3. Вид диалогового окна «Факторный анализ: Вращение»

Вывод основных результатов, необходимых для дальнейших расчетов, выглядит следующим образом (рис. 2.9.4):

Матрица корреляций:

Correlation Matrix

	социология	информатика	культуро логия	математика
Correlation				
социология	1,000	,333	,881	,433
информатика	,333	1,000	,225	,979
культурология	,881	,225	1,000	,313
математика	,433	,979	,313	1,000

Повёрнутая матрица компонентов:

Rotated Component Matrix^a

	Component	
	1	2
социология	,232	,943
информатика	,989	,119
культурология	,103	,968
математика	,971	,221

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 3 iterations.

Объяснённая суммарная дисперсия

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2,589	64,722	64,722	2,589	64,722	64,722	1,986	49,646	49,646
2	1,285	32,136	96,858	1,285	32,136	96,858	1,888	47,212	96,858
3	,111	2,780	99,638						
4	,014	,362	100,000						

Extraction Method: Principal Component Analysis.

Рис. 2.9.4. Результат выполнения описанной процедуры

Таким образом, видно, что матрица коэффициентов корреляции R и матрица факторных нагрузок A , при влияющем одном факторе будут следующими:

$$R = \begin{pmatrix} 1 & 0,333 & 0,881 & 0,433 \\ 0,333 & 1 & 0,225 & 0,979 \\ 0,881 & 0,225 & 1 & 0,313 \\ 0,433 & 0,979 & 0,313 & 1 \end{pmatrix} \quad A = \begin{pmatrix} 0,232 \\ 0,989 \\ 0,103 \\ 0,971 \end{pmatrix}$$

При этом, если найти произведение $A \times A^T$, то получим:

$$A \cdot A^T = \begin{pmatrix} 0,232 \\ 0,989 \\ 0,103 \\ 0,971 \end{pmatrix} \begin{pmatrix} 0,232 & 0,989 & 0,103 & 0,971 \end{pmatrix} = \begin{pmatrix} 0,054 & 0,229 & 0,024 & 0,225 \\ 0,229 & 0,978 & 0,102 & 0,960 \\ 0,024 & 0,102 & 0,011 & 0,100 \\ 0,225 & 0,960 & 0,100 & 0,943 \end{pmatrix}$$

Видно, что выбранный один фактор не учитывает всю корреляцию (например, почти отсутствует корреляция между первым и третьим показателями), поэтому следует ввести в рассмотрение второй фактор, тогда матрица факторных нагрузок будет следующая:

$$A = \begin{pmatrix} 0,232 & 0,943 \\ 0,989 & 0,119 \\ 0,103 & 0,968 \\ 0,971 & 0,221 \end{pmatrix}$$

$$A \times A^T = \begin{pmatrix} 0,232 & 0,943 \\ 0,989 & 0,119 \\ 0,103 & 0,968 \\ 0,971 & 0,221 \end{pmatrix} \begin{pmatrix} 0,232 & 0,989 & 0,103 & 0,971 \\ 0,943 & 0,119 & 0,968 & 0,221 \end{pmatrix} = \begin{pmatrix} 0,943 & 0,342 & 0,937 & 0,434 \\ 0,342 & 0,992 & 0,217 & 0,987 \\ 0,937 & 0,217 & 0,948 & 0,314 \\ 0,434 & 0,987 & 0,314 & 0,992 \end{pmatrix}$$

Видно, что получившаяся матрица, с определенной погрешностью (связанной с остаточной корреляцией), отражает исходную матрицу коэффициентов корреляции.

Попытаемся объяснить отобранные факторы.

Имеющиеся 4 показателя можно описать двумя факторами.

Фактор 1 наиболее полно характеризует первый и третий показатели, фактор 2 – второй и четвертый.

2.10. Многомерный анализ

Задание 2.10.

Исследовать взаимосвязь между тремя показателями:

1 – производительность труда (X);

2 – возраст (Y);

3 – стаж работников (V).

Таблица 2.10.1

№	Производительность труда, ед./день	Возраст, лет	Стаж, лет
1	3,00	20,00	2,00
2	4,12	25,00	3,00
3	8,96	38,00	15,00
4	5,33	27,00	5,00
5	7,29	45,00	20,00
6	4,98	33,00	8,00
7	5,14	29,00	6,00
8	7,11	42,00	9,00
9	3,58	41,00	4,00
10	6,37	30,00	7,00

Выполнение.

Вычисление парных коэффициентов корреляции было рассмотрено ранее (см. раб. 2.4).

Для нашей задачи имеем $r_{xy} = 0,595$, $r_{xv} = 0,827$, $r_{yv} = 0,728$.

Таким образом, можно рассчитать коэффициенты множественной корреляции.

Что касается частных (парциальных) коэффициентов корреляции, то для их вычисления в SPSS имеются специальные средства.

Создав файл с данными задачи, приступим к исследованию, выбрав в меню *Analyze* (Анализ) *Correlate* (Корреляция) *Partial* (Частный).

Откроется диалоговое окно (см. рис. 2.10.1).

Перенесите переменные *произв_труда* и *возраст* в поле «Переменные», а переменную *стаж* в поле «Регулирование для», оставим предварительную установку для двухстороннего теста значимости.

Нажав ОК, получим результаты (рис. 2.10.2).

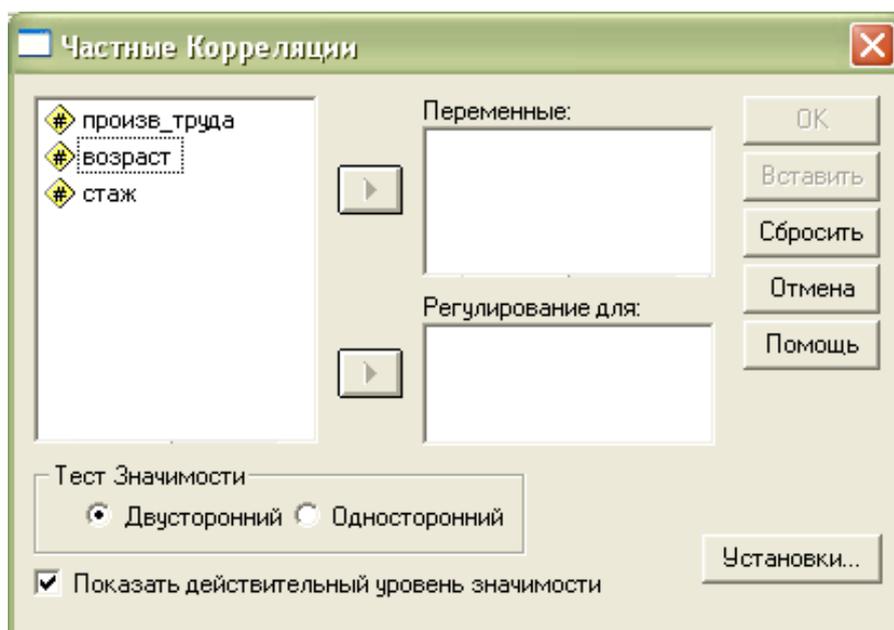


Рис. 2.10.1. Диалоговое окно «Частные корреляции»

Корреляции

Контрольные			произв труда	возраст
стаж	произв_труда	Корреляция	1,000	-,018
		Значимость (2-сторон.)	.	,963
		ст.св.	0	7
возраст	возраст	Корреляция	-,018	1,000
		Значимость (2-сторон.)	,963	.
		ст.св.	7	0

Рис. 2.10.2. Результаты вычисления частного коэффициента корреляции $r_{xy/v}$

Меняя переменные еще 2 раза, получаем результаты других двух частных коэффициентов корреляции.

Корреляции

Контрольные			произв труда	стаж
возраст	произв_труда	Корреляция	1,000	,714
		Значимость (2-сторон.)	.	,031
		ст.св.	0	7
стаж	стаж	Корреляция	,714	1,000
		Значимость (2-сторон.)	,031	.
		ст.св.	7	0

Рис. 2.10.3. Результаты вычисления частного коэффициента корреляции $r_{xv/y}$

Корреляции

Контрольные			стаж	возраст
произв_труда	стаж	Корреляция	1,000	,522
		Значимость (2-сторон.)	.	,149
		ст.св.	0	7
	возраст	Корреляция	,522	1,000
		Значимость (2-сторон.)	,149	.
		ст.св.	7	0

Рис. 2.10.4. Результаты вычисления частного коэффициента корреляции $r_{y\backslash x}$

В данном случае всё ещё выводится старый вариант таблицы результатов, соответствующий прежним версиям SPSS.

Результаты включают: частный корреляционный коэффициент, число степеней свободы (число наблюдений минус 3) и уровень значимости.

2.11. Множественная линейная регрессия

Задание 2.11.1.

Построить линейную модель по следующим данным, собранным о сотрудниках фирмы.

Таблица 2.11.1

№	Образование, лет	Текущая з/п, \$	Начальная з/п, \$	Проработанное время, мес.	Предшествующий опыт, мес.
1	15	57,000	27,000	98	144
2	16	40,200	18,750	98	36
3	16	60,625	22,500	91	44
4	14	39,900	15,750	91	59
5	12	26,250	10,950	89	0
6	12	22,200	15,000	88	324
7	16	42,300	26,250	87	126
8	8	30,750	15,000	87	451
9	12	26,700	12,900	87	18
10	12	20,850	12,000	87	163
11	15	35,250	15,000	87	54
12	15	26,700	15,000	87	56
13	12	26,550	13,050	87	11
14	16	66,750	52,500	83	258

15	18	66,875	31,980	79	30
16	12	30,000	15,750	79	308
17	16	83,750	21,750	79	12
18	12	24,450	10,950	75	32
19	8	31,950	15,750	74	408
20	15	33,900	15,750	67	96

Выполнение.

При построении линейной модели в качестве исследуемой переменной (зависимой) рассмотрим *размер текущей з/п*.

Выберем в меню *Analyze* (Анализ) *Regression* (Регрессия) *Linear* (Линейная). Появится диалоговое окно (см. рис. 2.11.1).

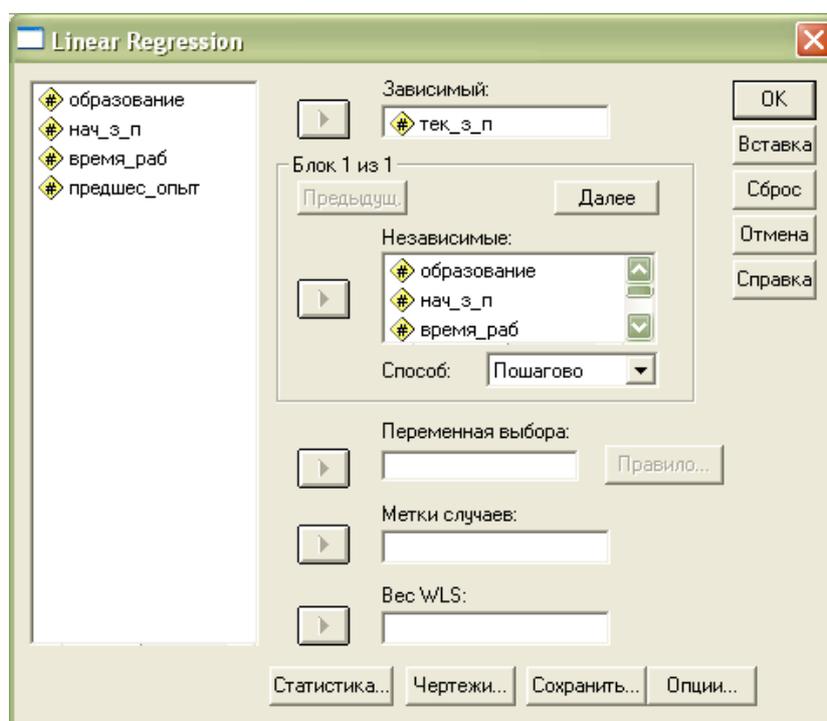


Рис. 2.11.1. Диалоговое окно «Линейная регрессия»

Поместим переменную *тек_з_п* в поле для зависимых переменных, а остальные переменные в поле для независимых переменных.

Для множественного анализа с несколькими независимыми переменными не рекомендуется оставлять метод включения всех переменных, установленный по умолчанию.

Этот метод соответствует одновременной обработке всех независимых переменных, выбранных для анализа, и поэтому он может рекомендоваться для использования только в случае простого анализа с одной независимой переменной.

Для множественного анализа следует выбрать один из пошаговых методов. При прямом методе независимые переменные, которые имеют наибольшие коэффициенты частной корреляции с зависимой переменной, пошагово увязываются в регрессионном уравнении.

При обратном методе начинают с результата, содержащего все независимые переменные, и затем исключают независимые переменные с наименьшими частными корреляционными коэффициентами, пока соответствующий регрессионный коэффициент не оказывается незначимым (в данном случае уровень значимости равен 0,1).

Наиболее распространенным является пошаговый метод, который устроен так же, как и прямой метод, однако после каждого шага переменные, используемые в данный момент, исследуются по обратному методу.

При пошаговом методе могут задаваться блоки независимых переменных; в этом случае заданные блоки на одном шаге обрабатываются совместно.

Выберем пошаговый метод, но воздержимся от блочной формы ввода данных, не задавая больше ни каких дополнительных расчётов, и начнем вычисление нажатием ОК.

Сводка для модели

Модель	R	R квадрат	Скорректированный R квадрат	Стд. ошибка оценки
1	,880 ^a	,775	,774	\$8,115.356
2	,891 ^b	,793	,793	\$7,776.652
3	,897 ^c	,804	,803	\$7,586.187
4	,900 ^d	,810	,809	\$7,465.139

- a. Предикторы: (константа) Beginning Salary
- b. Предикторы: (константа) Beginning Salary, Previous Experience (months)
- c. Предикторы: (константа) Beginning Salary, Previous Experience (months), Months since Hire
- d. Предикторы: (константа) Beginning Salary, Previous Experience (months), Months since Hire, Educational Level (years)

Рис. 2.11.2. Результаты анализа «Сводка для модели»

Из таблицы «Сводка для модели» следует, что вовлечение переменных в расчет производилось за четыре шага, то есть переменные *образование, начальная з/п, проработанное время, предшествующий опыт работников* поочередно внедрялись в уравнение регрессии.

Для каждого шага происходит вывод коэффициентов множественной регрессии, меры определённости, смещенной меры определённости и стандартной ошибки.

К указанным результатам пошагово присоединяются результаты расчёта дисперсии, которые здесь не приводятся.

Также, пошаговым образом, производится вывод соответствующих коэффициентов регрессии и значимость их отличия от нуля (рис. 2.11.3).

Коэффициенты^а

Модель		Нестандартизованные коэффициенты		Стандартизованные коэффициенты	t	Знч.
		B	Стд. ошибка	Бета		
1	(Константа)	1928,206	888,680		2,170	,031
	нач_з_п	1,909	,047	,880	40,276	,000
2	(Константа)	3850,718	900,633		4,276	,000
	нач_з_п	1,923	,045	,886	42,283	,000
	предшес_опыт	-22,445	3,422	-,137	-6,558	,000
3	(Константа)	-10266,6	2959,838		-3,469	,001
	нач_з_п	1,927	,044	,888	43,435	,000
	предшес_опыт	-22,509	3,339	-,138	-6,742	,000
	время_раб	173,203	34,677	,102	4,995	,000
4	(Константа)	-16149,7	3255,470		-4,961	,000
	нач_з_п	1,768	,059	,815	30,111	,000
	предшес_опыт	-17,303	3,528	-,106	-4,904	,000
	время_раб	161,486	34,246	,095	4,715	,000
	образование	669,914	165,596	,113	4,045	,000

а. Зависимая переменная: тек_з_п

Рис. 2.11.3. Коэффициенты линейной модели

Уравнение регрессии для прогнозирования значения *тек_з_п* выглядит следующим образом:

$$\text{тек_з_п} = 669,914 \times \text{образование} + 161,486 \times \text{время_раб} - 17,303 \times \text{предшес_опыт} + 1,768 \times \text{нач_з_п}$$

При помощи соответствующих опций можно организовать вывод большого числа дополнительных статистических характеристик и графиков.

Можно также создать много дополнительных переменных и добавить их в исходный файл данных.

Пример 2.11.2. Имеются 12 наблюдений за тремя показателями:

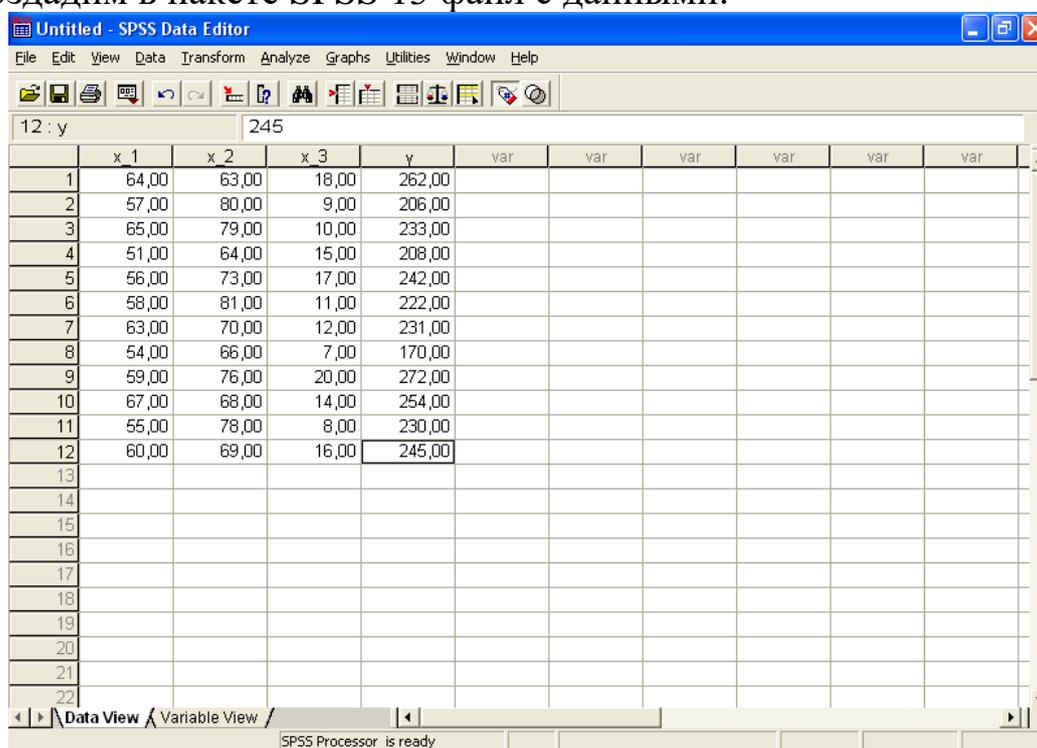
X_1	64	57	65	51	56	58	63	54	59	67	55	60
X_2	63	80	79	64	73	81	70	66	76	68	78	69
X_3	18	9	10	15	17	11	12	7	20	14	8	16
Y	262	206	233	208	242	222	231	170	272	254	230	245

Используя пакет для обработки статистических данных получить уравнение множественной линейной регрессии и проанализировать качество полученной модели.

Решение. Построим уравнение регрессии в виде

$$Y = a_0 + a_1X_1 + a_2X_2 + a_3X_3.$$

Создадим в пакете SPSS 13 файл с данными:



The screenshot shows the SPSS Data Editor window titled 'Untitled - SPSS Data Editor'. The window contains a data grid with 12 rows and 10 columns. The first four columns are labeled 'x_1', 'x_2', 'x_3', and 'y'. The data values are as follows:

	x_1	x_2	x_3	y	var	var	var	var	var	var
1	64,00	63,00	18,00	262,00						
2	57,00	80,00	9,00	206,00						
3	65,00	79,00	10,00	233,00						
4	51,00	64,00	15,00	208,00						
5	56,00	73,00	17,00	242,00						
6	58,00	81,00	11,00	222,00						
7	63,00	70,00	12,00	231,00						
8	54,00	66,00	7,00	170,00						
9	59,00	76,00	20,00	272,00						
10	67,00	68,00	14,00	254,00						
11	55,00	78,00	8,00	230,00						
12	60,00	69,00	16,00	245,00						
13										
14										
15										
16										
17										
18										
19										
20										
21										
22										

Выберем в меню пункты (опции) *Analyze...* (Анализ) *Regression...*(Регрессия) *Linear...* (Линейная). Появится диалоговое окно *Linear Regression* (Линейная регрессия) (рис.5.4).

Перенесем переменную Y в поле для зависимых переменных и присвоим переменным x_1 , x_2 , x_3 статус независимых переменных.

После нажатия кнопки «Statistics...» в появившемся окне установим флажки напротив опций «*R squared change*» и «*Descriptives*», далее нажимаем клавишу «*Continue*» (рис.5.5).

Расчет параметров модели начинается нажатием клавиши ОК.

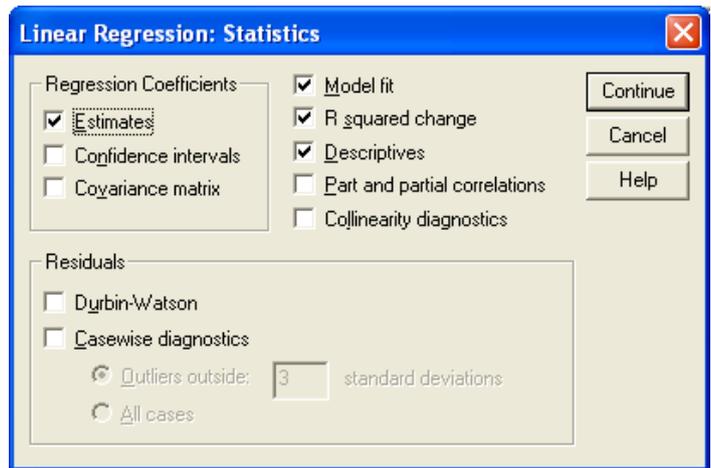
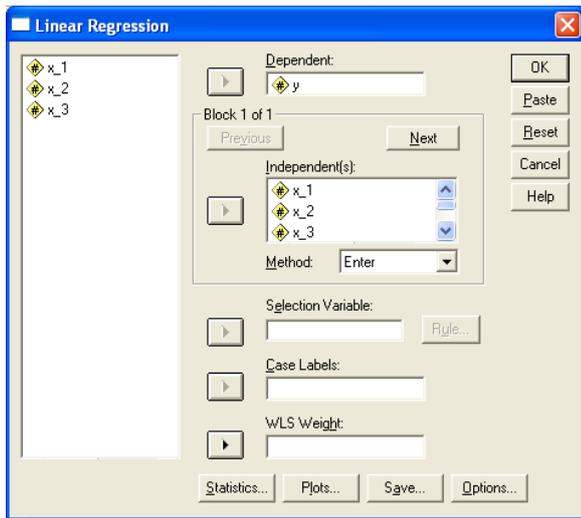


Рис.2.11.4. Вид диалогового окна «Линейная регрессия»

Рис.2.11.5. Вид диалогового окна «Статистика линейной регрессии»

Вывод основных результатов выглядит следующим образом:

Descriptive Statistics

	Mean	Std.Deviation	N
y	231.2500	27.71322	12
x_1	59.0833	4.87029	12
x_2	72.2500	6.44029	12
x_3	13.0833	4.20948	12

Variables Entered/Removed^b

Model	Variables Entered	Variables Removed	Method
1	x_3, ^a x_1, x_2		Enter

^aAll requested entered

^bDependent Variable: y

Correlations

		y	x_1	x_2	x_3
Pearson Correlation	y	1.000	.595	.019	.770
	x_1	.595	1.000	-.012	.204
	x_2	.019	-.012	1.000	-.353
	x_3	.770	.204	-.353	1.000

Sig. (1-tailed)	y	.	.021	.476	.002
	x_1	0.21	.	.485	.263
	x_2	.476	.485	.	.130
	x_3	.002	.263	.130	.
N	y	12	12	12	12
	x_1	12	12	12	12
	x_2	12	12	12	12
	x_3	12	12	12	12

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.934 ^a	.873	.825	11.58090	.873	18.331	3	8	.001

a. Predictors: (Constant), x_3, x_1, x_2

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	7375.313	3	2458.438	18.331	.001 ^a
	Residual	1072.937	8	134.117		
	Total	8448.250	11			

a. Predictors: (Constant), x_3, x_1, x_2

b. Depended Variable: y

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-78.166	60.771		-1.286	.234
	x_1	2.496	.734	.439	3.402	.009
	x_2	1.303	.581	.303	2.243	.055
	x_3	5.183	.907	.787	5.712	.000

Основные числовые характеристики исследуемых показателей будут следующие: $\bar{y} = 231.25$; $\bar{x}_1 = 59.1$; $\bar{x}_2 = 72.3$; $\bar{x}_3 = 13.1$;

$$\sigma_y = 27.71; \sigma_{x_1} = 4.87; \sigma_{x_2} = 6.44; \sigma_{x_3} = 4.21.$$

Параметры модели регрессии представлены в таблице «Coefficients».

Таким образом уравнение множественной регрессии имеет вид:

$$\bar{y} = a_0 + a_1x_1 + a_2x_2 + a_3x_3 = -78.166 + 2.496x_1 + 1.303x_2 + 5.183x_3.$$

Для сравнения влияния каждой независимой переменной вычислим стандартизированные коэффициенты регрессии

$$\beta_1 = 0.439; \beta_2 = 0.303; \beta_3 = 0.787.$$

Увеличение показателя X_1 только на одно значение σ_{x_1} увеличивает в среднем показатель Y на $0.439\sigma_y$.

Аналогично, увеличение X_2 на значение σ_{x_2} увеличивает Y на $0.303\sigma_y$. Увеличение X_3 на σ_{x_3} увеличивает Y на $0.787\sigma_y$.

Коэффициенты эластичности будут равны

$$\mathcal{E}_1 = a_1 \frac{\bar{x}_1}{\bar{y}} = 2.496 \frac{59.1}{231.25} = 0.638;$$

$$\mathcal{E}_2 = a_2 \frac{\bar{x}_2}{\bar{y}} = 1.303 \frac{72.3}{231.25} = 0.407;$$

$$\mathcal{E}_3 = a_3 \frac{\bar{x}_3}{\bar{y}} = 5.183 \frac{13.1}{231.25} = 0.294.$$

Увеличение переменных X_1 , X_2 и X_3 на 1% от своих средних значений приводит в среднем к увеличению Y соответственно на 0.638 %, 0.407 % и 0.294 %.

Проверим значимость модели множественной регрессии по F -критерию. Выдвигаем гипотезу H_0 : уравнение незначимо.

$$F_H = \frac{R_{y(x_1x_2x_3)}^2 (n-m-1)}{(1-R_{y(x_1x_2x_3)}^2)m},$$

где $n=12$ – количество наблюдений, $m=3$ – количество показателей.

Коэффициент детерминации модели

$$R_{y(x_1x_2x_3)}^2 = 0.873.$$

Наблюдаемое значение критерия

$$F_{\hat{I}} = \frac{0.873 \cdot (12-3-1)}{(1-0.873) \cdot 3} = 18.331.$$

Отметим, что вычисления в пакете SPSS 13 дают то же значение

$$F_H = 18.331.$$

По таблице F -распределения найдем критическую точку

$$F_{кр} = F(1-\alpha; m; n-m-1) = F(1-0.05; 3; 8) = 4.07.$$

Так как $F_H > F_{кр}$, то есть все основания отвергнуть гипотезу H_0 , то есть модель регрессии значима при уровне значимости $\alpha=0.05$.

Проверим на значимость коэффициенты a_1 , a_2 и a_3 с помощью T -критерия Стьюдента.

Имеем

$$H_0 : a_j = 0; \quad H_1 : a_j > 0, \quad j = 1, 2, 3.$$

$$T_{H(a_j)} = \frac{a_j}{\sigma_{a_j}},$$

где $\sigma_{a_j} = \sigma_\varepsilon \sqrt{b_{jj}}$, b_{jj} – диагональные элемента матрицы $(X^T X)^{-1}$.

Вычислим дисперсию остатков модели по соотношению

$$\begin{aligned} \sigma_\varepsilon^2 &= \frac{1}{n-m} (Y^T Y - A^T X^T Y) = \\ &= \frac{1}{12-3} \left((262, 206, \dots, 245) \begin{pmatrix} 262 \\ 206 \\ \vdots \\ 245 \end{pmatrix} - (-78.166, 2.496, 1.303, 5.183) \begin{pmatrix} 2775 \\ 164840 \\ 200532 \\ 37294 \end{pmatrix} \right) = \\ &= 1072.94. \end{aligned}$$

Средняя ошибка (точность) модели

$$\sigma_\varepsilon = \sqrt{1072.94} = 32.76.$$

Тогда

$$\begin{aligned} \sigma_{a_1} &= \sigma_\varepsilon \sqrt{b_{11}} = 32.76 \sqrt{\frac{467119}{116332896}} = 2.076; \quad T_{H(a_1)} = \frac{a_1}{\sigma_{a_1}} = \frac{2.496}{2.076} = 1.202; \\ \sigma_{a_2} &= \sigma_\varepsilon \sqrt{b_{22}} = 32.76 \sqrt{\frac{292492}{116332896}} = 1.643; \quad T_{H(a_2)} = \frac{a_2}{\sigma_{a_2}} = \frac{1.303}{1.643} = 0.793; \\ \sigma_{a_3} &= \sigma_\varepsilon \sqrt{b_{33}} = 32.76 \sqrt{\frac{714151}{116332896}} = 2.567; \quad T_{H(a_3)} = \frac{a_3}{\sigma_{a_3}} = \frac{5.183}{2.567} = 2.019. \end{aligned}$$

По таблице квантилей T -распределения Стьюдента найдем критическую точку.

$$T_{кр} = T(1 - \alpha; n - 2) = T(1 - 0.05; 12 - 2) = T(0.95; 10) = 1.812.$$

Так как $T_{H(a_1)} < T_{кр}$ и $T_{H(a_2)} < T_{кр}$, то нет оснований отвергать гипотезу H_0 , то есть коэффициенты a_1 и a_2 незначимы.

Так как $T_{H(a_3)} > T_{кр}$, то гипотезу H_0 отвергаем, то есть коэффициент a_3 значим.

2.12. Дискриминантный анализ

Задание 2.12.

Построить линейную модель по следующим данным, собранным о сотрудниках фирмы (табл. 2.12.1) (0 – премия не выдана сотруднику, 1 – премия выдана сотруднику).

Таблица 2.12.1

Исходные данные для исследования

№	Образование, лет	Текущая з/п, \$	Начальная з/п, \$	Проработанное время, мес.	Предшествующий опыт, мес.	Премия
1	15	57,000	27,000	98	144	1
2	16	40,200	18,750	98	36	1
3	16	60,625	22,500	91	44	1
4	14	39,900	15,750	91	59	0
5	12	26,250	10,950	89	0	0
6	12	22,200	15,000	88	324	0
7	16	42,300	26,250	87	126	1
8	8	30,750	15,000	87	451	0
9	12	26,700	12,900	87	18	1
10	12	20,850	12,000	87	163	0
11	15	35,250	15,000	87	54	1
12	15	26,700	15,000	87	56	1
13	12	26,550	13,050	87	11	1
14	16	66,750	52,500	83	258	0
15	18	66,875	31,980	79	30	0
16	12	30,000	15,750	79	308	0
17	16	83,750	21,750	79	12	1
18	12	24,450	10,950	75	32	0
19	8	31,950	15,750	74	408	1
20	15	33,900	15,750	67	96	0

Выполнение.

Для решения поставленной задачи необходимо после того, как сформированы необходимые переменные, выбрать в меню *Analyze* (Анализ) *Classify* (Классификация) *Discriminant* (Дискриминант).

Откроется диалоговое окно *Discriminant Analysis* (Дискриминантный Анализ) (см. рис. 2.12.1).

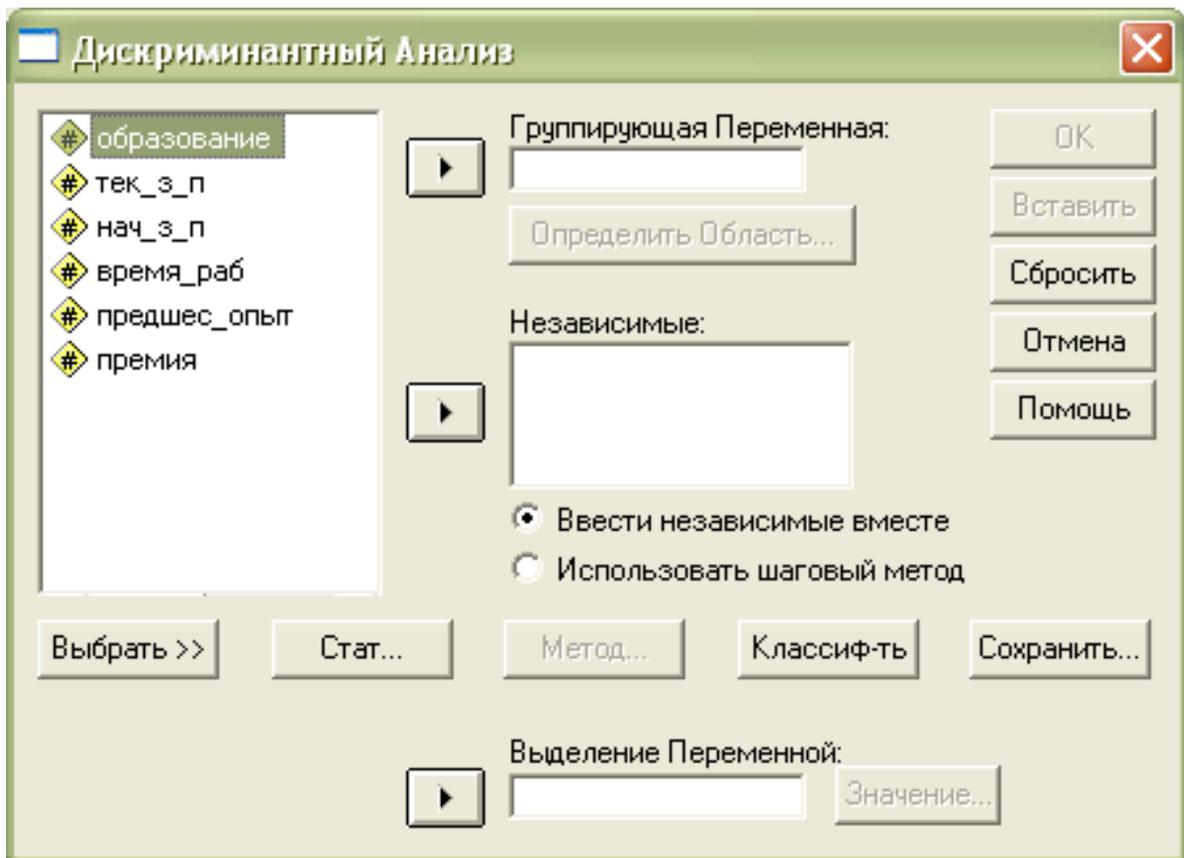


Рис. 2.12.1. Диалоговое окно *Discriminant Analysis* (Дискриминантный Анализ)

Далее необходимо сделать следующее:

- Поместить переменную *премия* в поле, предназначенное для группирующих переменных.
- После щелчка по выключателю *Define Range* (Определить Область) введите минимальное и максимальное значения этой переменной: 0 и 1.
- Остальным переменным присвоить статус независимых переменных. Для начала оставить установленный по умолчанию метод: *Enter independents together* (Ввести независимые вместе), при котором в анализе одновременно будут участвовать все независимые переменные.
- После щелчка по выключателю *Statistics* (Статистика) активировать опции: *Means* (Средние значения), *Univariate ANOVAs* (Одномерные тесты ANOVA), *Unstandardized Function Coefficients* (Нестандартизированные коэффициенты функции) и *Within-group Correlation Matrices* (Корреляционная матрица внутри группы).

- Через выключатель *Classify* (Классифицировать) сделать дополнительно запрос на вывод диаграмм по отдельным группам (*Separate-groups Plots*), результатов для отдельных наблюдений (*Casewise results*) и сводной таблицы (*Summary table*).

Довольно полезный график для объединенных групп, который был реализован в ранних версиях SPSS, и сейчас можно активировать в диалоговом окне, однако вместо графика в окне отображения результатов будет появляться предупреждение о том, что такая гистограмма в анализах более не доступна.

- При помощи выключателя *Save* (Сохранить) активировать сохранение значения дискриминантной функции в дополнительной переменной (*Discriminant Scores*).
- Начать расчёт нажатием ОК.

После вводного обзора действительных и пропущенных значений приводятся средние значения, стандартные отклонения, количество наблюдений для каждой группы в отдельности и суммарные показатели для обеих групп.

Затем проводится тест, насколько значимо различаются между собой переменные в обеих группах; наряду с тестовой величиной, в качестве которой служит Лямбда Уилкса («*Wilks-Lambda*»), применяется также и простой дисперсионный анализ.

Далее следует корреляционная матрица между всеми переменными, в которой приводятся коэффициенты, усредненные для обеих групп.

Следующими шагами являются расчёт и анализ коэффициентов дискриминантной функции.

Значения этой функции должны как можно отчётливой разделять обе группы. Мерой удачности этого разделения служит корреляционный коэффициент между рассчитанными значениями дискриминантной функции и показателем принадлежности к группе.

Следующая таблица дает представление о том, как сильно отдельные переменные, применяемые в дискриминантной функции, коррелируют со стандартизированными значениями этой дискриминантной функции.

При этом корреляционные коэффициенты были рассчитаны в обеих группах по отдельности и затем усреднены.

И в заключение, приводятся сами коэффициенты дискриминантной функции (рис. 2.12.2).

Коэффициенты канонической дискриминантной функции

	Функция
	1
образование	-,078
тек_з_п	,000
нач_з_п	,000
время_раб	,095
предшес_опыт	-,003
(Константа)	-7,505

Ненормированные коэффициенты

Рис. 2.12.2. Результаты вычисления коэффициентов

Интерес представляет также таблица, в которой построчно для каждого наблюдения приводится информация о значении дискриминантной функции и определяется принадлежность к одной из двух групп.

Таким образом, получили дискриминантную функцию:

$$d = -0,78 \times \text{образование} + \\ + 0,095 \times \text{время_раб} - \\ - 0,003 \times \text{предшес_опыт} - \\ - 7,505.$$

3. ИНДИВИДУАЛЬНЫЕ ЗАДАНИЯ ДЛЯ ИССЛЕДОВАНИЙ

ЗАДАНИЕ 1.

По имеющимся данным построить закон распределения заданной случайной величины (см. варианты задания).

Необходимо:

1. Построить вариационный ряд исследуемой случайной величины.
2. Произвести группировку данных вариационного ряда на 6 – 10 интервалов (разрядов, групп).
3. Вычислить и представить графически эмпирические функции распределения исследуемой случайной величины.
4. Выровнять (аппроксимировать) имеющиеся данные подходящим теоретическим законом распределения заданной случайной величины.
5. Проанализировать полученные результаты.

ВАРИАНТЫ ЗАДАНИЯ

1. Количество сделок (договоров) фирмы по месяцам (ед.): 18, 22, 23, 24, 25, 24, 27, 25, 23, 22, 23, 27, 26, 24, 22, 20, 21, 22, 23, 22, 21, 20, 19, 16, 17, 20, 22, 22, 23, 27, 30, 28, 27, 25, 27, 25, 22, 20, 20, 21, 23, 27, 29, 33, 30, 29, 26, 24, 23, 21, 18, 20, 22, 23, 24, 26.

2. Прибыль фирмы (млн.р./мес.): 8.3; 8.7; 10.3; 10.7; 9.9; 8.6; 8.9; 9.1; 9.9; 9.8; 10.2; 8.9; 9.4; 9.3; 9.2; 9.6; 9.4; 9.3; 9.2; 10.1; 9.1; 10.2; 10.4; 9.0; 8.8; 12.1; 11.3; 10.9; 10.3; 9.8; 10.3; 10.4; 10.9; 10.8; 10.7; 9.3; 11.2; 11.4; 8.2; 12.7; 10.3; 10.8; 10.3; 10.4; 10.9; 10.8.

3. Рентабельность собственных средств предприятия (%): 22.3; 24.5; 21.8; 19.7; 22.3; 24.1; 22.8; 21.7; 21.5; 20.4; 20.7; 21.5; 21.2; 20.3; 20.8; 21.4; 21.6; 21.9; 22.0; 22.2; 22.8; 19.7; 21.4; 22.9; 22.1; 22.3; 21.6; 21.3; 21.7.

4. Отклонение от плана нормы выработки рабочего по неделям (%): 0.2; 0.3; 0.1; 0; 0; -0.1; -0.3; -0.1; 0; 0; 0.3; 0.2; 0.1; -0.1; -0.2; -0.3; -0.1; 0; 0.2; 0.1; 0; 0.1; 0; 0; 0; 0.4; 0.3; 0.2; -0.1; 0.1; -0.1; 0.2; 0.3; 0.8; 0.9; 0.5; 0; 0.3; -0.5; -0.6; 0.1; 0; 0.6; 1.5; 0.7; 0.2; 0.1.

5. Выпуск изделий по месяцам (шт.): 32, 30, 28, 27, 24, 25, 27, 32, 29, 28, 27, 22, 20, 22, 25, 24, 26, 25, 20, 18, 19, 26, 33, 35, 30, 28, 27, 27, 23, 19, 20, 27, 28, 30.

6. Валовые издержки предприятия (тыс. р./мес.): 30, 40, 52, 44, 37, 38, 42, 41, 44, 33, 38, 41, 39, 37, 30, 37, 42, 53, 50, 49, 51, 46, 45, 40, 42, 51, 52, 47, 46, 54, 38, 61, 57, 50, 44, 47.

7. Месячная прибыль фирмы (тыс.р.): 3.4; 2.8; 7.1; 6.2; 4.4; 4.7; 4.8; 4.9; 5.0; 5.1; 3.8; 3.9; 4.7; 4.6; 4.5; 5.7; 5.8; 6.3; 6.4; 5.9; 7.7; 6.2; 4.4; 5.6; 5.0; 6.0; 7.0; 5.8; 6.1; 5.8; 6.4; 6.3; 6.0; 6.9; 6.1; 6.2; 4.8; 4.7.

8. Величина основных фондов предприятия (млн. р.): 60, 70, 82, 91, 99, 103, 100, 85, 74, 63, 84, 76, 87, 88, 89, 90, 93, 96, 91, 99, 81, 77, 67, 71, 68, 70, 76, 81, 96, 96, 93, 84, 96, 99, 102, 105, 100, 98, 84, 77.

9. Урожайность пшеницы (ц/га): 21.3; 21.4; 20.6; 20.8; 22.4; 23.8; 19.3; 17.4; 18.9; 21.4; 22.8; 22.6; 23.1; 22.9; 19.4; 19.5; 20.5; 20.6; 20.7; 21.0; 22.0; 21.8; 20.9; 20.0; 21.4; 23.0.

10. Уровень инфляции (% в месяц): 3.1; 3.7; 6.4; 3.8; 4.2; 3.0; 3.2; 3.7; 3.0; 3.1; 3.6; 4.1; 4.8; 5.5; 3.0; 3.1; 3.2; 3.0; 3.7; 3.5; 3.1; 3.2; 3.0; 3.1; 3.0; 3.4; 3.5; 3.6; 4.1; 3.7; 4.2; 3.8; 4.6; 4.7; 5.1; 5.8; 3.1; 3.0; 5.4; 6.0; 6.2; 3.0; 3.2; 3.4; 3.5; 3.3; 3.3; 3.0; 4.0.

11. Месячная рентабельность предприятия (%): 19.7; 20.4; 22.8; 20.7; 21.2; 20.3; 20.4; 20.3; 20.7; 22.1; 22.7; 19.1; 21.2; 21.7; 21.2; 21.4; 21.6; 20.4; 20.5; 21.7; 21.0; 20.6; 20.8; 20.9; 22.0; 19.7; 20.1; 23.0; 20.4; 22.3; 22.7; 21.4; 21.5.

12. Валовая продукция цеха (тыс. р./мес.): 160, 172, 168, 173, 174, 169, 171, 172, 170, 171, 168, 164, 158, 172, 171, 170, 166, 164, 162, 161, 161, 168, 170, 170, 163, 162, 161, 167, 164, 167, 158, 159, 171, 170, 164, 160, 163, 165, 166, 166, 156, 170, 164.

13. Количество опозданий работников цеха за отчетные месяцы: 8, 7, 6, 7, 6, 2, 1, 4, 6, 8, 10, 12, 9, 8, 7, 6, 7, 7, 6, 9, 4, 10, 11, 7, 6, 9, 4, 6, 2, 3, 7, 6, 5, 4, 4, 3, 9, 1, 7, 6, 9, 4, 3.

14. Количество бракованных деталей за смену: 5, 1, 2, 7, 4, 3, 2, 1, 0, 2, 2, 3, 5, 4, 4, 1, 7, 6, 4, 3, 3, 4, 5, 5, 3, 3, 4, 6, 1, 7, 8, 2, 1, 3, 4, 5, 8, 7, 7, 6, 4, 3, 0, 4, 3, 2, 1, 6, 4, 5, 4, 3, 2, 8, 7, 7, 6, 6, 5.

15. Товарооборот на душу населения (тыс. р.): 0.9; 0.8; 0.6; 0.4; 0.5; 0.6; 0.7; 0.8; 0.4; 0.9; 0.8; 0.7; 0.9; 1.2; 1.1; 1.0; 0.7; 1.0; 1.4; 1.3; 1.2; 1.4; 0.8; 0.7; 0.9; 0.3.

16. Месячный процент премии на предприятии (%): 15, 16, 18, 20, 20, 25, 27, 28, 22, 18, 16, 15, 15, 16, 18, 21, 23, 25, 25, 22, 18, 16, 20, 19, 18, 16, 21, 23, 26, 28, 30, 32.

17. Запасы продукции на складе по месяцам (тыс. р.): 10.8; 10.4; 17.3; 20.1; 15.3; 23.7; 14.2; 12.1; 11.3; 12.0; 18.4; 13.4; 18.3; 14.7; 16.0; 17.9; 16.2; 16.4; 15.8; 19.3; 18.6; 8.8; 10.1; 9.4; 9.7; 19.4; 20.2; 25.4; 24.3; 20.9; 23.1; 9.1; 9.4; 20.8; 21.2; 22.1; 11.4.

18. Коэффициент ритмичности работы цеха по сменам: 0.62; 0.71; 0.64; 0.79; 0.54; 0.59; 0.64; 0.70; 0.82; 0.67; 0.69; 0.71; 0.81; 0.74; 0.70; 0.72; 0.60; 0.84; 0.76; 0.78; 0.69; 0.66; 0.60; 0.80; 0.58; 0.56.

19. Себестоимость добычи одной тонны руды по месяцам (руб.): 10.5; 10.2; 10.7; 10.4; 10.3; 10.7; 10.6; 10.2; 10.3; 10.5; 10.4; 10.7; 10.4; 10.5; 10.6; 10.4; 10.1; 10.3; 10.2; 10.1; 10.5; 10.5; 10.4; 10.3; 10.7; 10.6; 11.0; 10.8; 10.7; 10.0; 10.5; 11.1; 10.3; 10.4; 10.6; 10.5; 10.5; 10.6; 10.7; 10.5; 10.4; 10.2.

20. Средняя выработка на одного рабочего в фирме по декадам (тыс. р.): 83, 70, 64, 69, 72, 80, 84, 78, 74, 70, 53, 58, 60, 66, 67, 60, 64, 33, 47, 65, 84, 77, 56, 54, 55, 51, 48, 66, 67, 60, 40.

21. Месячная выручка от реализации продукции (тыс. р.): 570, 592, 604, 581, 680, 730, 654, 642, 686, 672, 591, 703, 574, 688, 608, 630, 650, 644, 666, 682, 599, 704, 710, 520, 656, 642, 610, 670, 594, 600, 700, 680, 676, 647, 662, 641, 634, 690.

22.Количество заказов на компьютеры (шт./день): 3, 2, 1, 0, 1, 2, 4, 6, 5, 4, 3, 2, 1, 1, 2, 5, 6, 7, 6, 5, 4, 3, 2, 2, 3, 1, 2, 1, 0, 3, 4, 6, 4, 2, 1, 0, 1, 3, 4, 7, 8, 5, 1, 4, 3, 2, 0, 3, 6, 5, 3, 1.

23.Месячное потребление электроэнергии (квт.ч): 171, 183, 199, 185, 183, 177, 179, 181, 188, 179, 177, 179, 180, 182, 190, 179, 174, 172, 166, 167, 170, 171, 172, 175, 178, 181, 188, 180, 174, 175, 175, 166, 168, 171, 176, 175, 174, 171, 173, 186, 180, 178, 197, 193, 185, 176.

24.Квартальный процент премии на предприятии (%): 15, 16, 18, 20, 20, 25, 27, 28, 22, 18, 16, 15, 15, 16, 18, 21, 30, 23, 25, 25, 22, 18, 16, 20, 19, 18, 16, 21, 23, 26, 28, 30, 18, 21, 30, 23, 32.

25.Динамика отклонений от номинала длины заготовок, производимых станком–автоматом (мм): -2, -3, 0, -1, 2, 4, 2, -3, -4, -2, 0, 3, 2, 4, 3, 2, -2, -3, -2, 0, 1, 2, 3, 0, -1, 1, 2, 3, 4, 3, 4, 6, 0, -2, -5, -6, -3, -2, 0, 2, 3, 4, 3, 2, 1.

26.Себестоимость производства детали по месяцам (руб.): 11.4; 11.3; 10.7; 10.6; 10.4; 10.7; 10.4; 10.5; 10.6; 10.4; 10.1; 10.3; 10.2; 10.1; 11.5; 10.5; 11.4; 10.3; 10.7; 10.6; 11.0; 10.8; 10.7; 10.0; 10.5; 11.1; 10.3; 10.4; 10.6; 11.5; 10.5; 10.6; 10.7; 11.5; 10.4; 11.2.

27.Средняя выработка на одного рабочего в фирме по месяцам (тыс. долл.): 66, 60, 64, 69, 72, 80, 84, 78, 74, 70, 53, 58, 60, 66, 67, 60, 64, 33, 47, 65, 64, 77, 56, 54, 65, 55, 51, 48, 66, 67, 60, 55, 52, 57, 43.

28.Месячные объемы продаж продукции (тыс. р.): 770, 692, 614, 781, 680, 730, 654, 642, 686, 672, 791, 703, 574, 688, 608, 630, 650, 644, 666, 682, 599, 704, 710, 720, 656, 642, 610, 670, 594, 600, 700, 680, 676, 647, 662, 641, 634, 690.

29.Коэффициент ритмичности работы предприятия по сменам: 0.67; 0.69; 0.71; 0.62; 0.71; 0.64; 0.79; 0.54; 0.59; 0.64; 0.70; 0.82; 0.67; 0.69; 0.71; 0.81; 0.74; 0.70; 0.72; 0.60; 0.84; 0.76; 0.78; 0.69; 0.66; 0.60; 0.80; 0.58; 0.56.

30.Средняя выработка на одного рабочего предприятия по сменам (тыс. р.): 55, 61, 68, 83, 70, 64, 69, 72, 80, 84, 78, 74, 70, 53, 58, 60, 66, 67, 60, 64, 33, 47, 65, 84, 77, 56, 54, 55, 51, 48, 66, 67, 60, 66.

ЗАДАНИЕ 2.

По имеющимся данным и результатам расчетов, выполненных в работе 1, построить доверительные интервалы для математического ожидания, дисперсии и среднего квадратического отклонения заданной случайной величины.

При этом следует:

1. Распределение случайной величины взять из задания 1.
2. Доверительные интервалы параметров строить с учетом полученного закона распределения случайной величины как точными так и приближенными методами (там, где это возможно) для различных значений доверительной вероятности (0,9; 0,95; 0,99).
3. Проанализировать полученные результаты.

ЗАДАНИЕ 3.

По схеме случайной выборки из N объектов исследования генеральной совокупности было отобрано n . Получено соответствующее распределение анализируемого показателя (см. варианты задания 1).

Необходимо:

- 1) определить границы, в которых с вероятностью P_0 заключено среднее значение показателя во всей генеральной совокупности;
- 2) вычислить вероятность того, что доля объектов исследования в выборке со значением показателя, не меньшим чем среднее значение выборки плюс среднее квадратическое отклонение, отличается от доли таких приборов во всей генеральной совокупности (партии) не больше, чем на ε (по абсолютной величине);
- 3) рассчитать необходимую численность выборки для определения среднего значения показателя с предельной ошибкой, равной $\varepsilon/2$;
- 4) проанализировать полученные результаты.

ЗАДАНИЕ 4.

Имеются наблюдения за двумя показателями (см. варианты заданий).

Исследовать взаимосвязь этих показателей с использованием ЭВМ.

Требуется:

- 1) вычислить основные характеристики показателей (среднее значение, дисперсию, среднее квадратическое отклонение);
- 2) оценить тесноту взаимосвязи показателей, вычислив их коэффициент корреляции;
- 3) вычислить коэффициенты регрессионного уравнения взаимосвязи исследуемых показателей;
- 4) оценить качество выполненного корреляционного и регрессионного анализа (проверить значимость коэффициента корреляции, адекватность и точность уравнения регрессии);
- 5) проанализировать полученные результаты.

Исходные данные к заданию 4:

№ вар	Исследуемые показатели	Значения показателей							
		X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
1	X_1, X_3	10,3	21,4	34,9	30,8	41,4	61,5	50,8	47,9
2	X_2, X_4	11,2	23,5	36,2	31,9	40,2	62,2	51,9	42,5
3	X_3, X_5	10,9	22,6	37,0	36,5	40,3	62,9	55,5	44,8
4	X_4, X_6	11,4	21,9	36,1	34,8	39,4	63,1	54,8	46,2
5	X_5, X_7	12,0	19,6	38,3	35,3	38,6	63,7	55,3	48,3
6	X_6, X_8	11,5	20,0	39,4	34,9	37,9	63,5	57,9	42,2
7	X_1, X_8	12,2	19,7	40,2	36,2	36,5	63,9	58,2	51,6
8	X_2, X_7	12,9	18,4	40,3	37,0	34,8	64,2	59,0	55,4
9	X_3, X_6	13,1	20,5	41,9	36,1	35,3	64,6	60,1	52,3
10	X_4, X_5	13,7	18,7	42,9	33,3	34,9	65,0	61,4	54,6
11	X_1, X_2	14,5	19,3	42,5	39,4	33,2	65,3	63,5	54,7
12	X_2, X_3	13,9	17,4	44,8	36,2	32,0	65,3	62,6	55,6
13	X_3, X_4	14,2	16,4	46,2	40,3	31,9	66,1	66,9	59,1
14	X_4, X_8	14,6	18,9	48,3	41,9	30,5	67,9	69,6	51,0
15	X_5, X_6	15,0	16,2	49,9	47,9	29,5	66,4	72,9	54,3
16	X_6, X_7	15,3	17,1	47,9	42,5	29,2	68,4	78,5	61,3
17	X_7, X_8	16,1	15,8	49,5	44,8	28,9	69,5	79,8	62,9
18	X_8, X_2	15,9	16,6	49,8	46,2	27,1	62,6	80,2	64,1
19	X_7, X_1	16,4	15,5	51,2	48,3	27,7	61,9	81,3	65,3
20	X_6, X_3	17,2	14,9	51,3	42,2	26,5	69,6	82,2	68,6
21	X_1, X_7	16,8	14,3	52,2	51,6	25,9	70,0	88,6	67,1
22	X_2, X_6	17,5	15,8	51,6	55,4	24,2	69,7	88,4	61,2
23	X_3, X_5	17,8	13,6	55,4	52,3	24,6	69,8	92,3	63,4
24	X_4, X_8	18,2	14,2	52,3	54,6	23,0	71,3	94,6	69,8
25	X_3, X_8	17,6	13,8	54,6	54,7	22,3	72,1	94,7	71,3
26	X_5, X_8	18,3	13,1	54,7	55,6	21,3	72,5	95,6	77,1
27	X_6, X_1	18,4	12,6	61,3	59,1	21,1	74,3	96,1	78,5
28	X_5, X_1	19,5	12,4	62,9	51,0	19,9	76,6	97,0	74,3
29	X_4, X_1	18,9	11,5	64,1	54,3	18,4	76,5	98,3	76,6
30	X_7, X_2	19,4	13,5	65,3	61,2	17,2	77,2	99,2	79,5

ЗАДАНИЕ 5.

По имеющимся статистическим данным оценить и исследовать взаимосвязь между двумя показателями. Требуется:

1. Вычислить основные числовые характеристики каждой исследуемой случайной величины (переменной): математические ожидания, дисперсии, средние квадратические отклонения.

2. По имеющейся корреляционной таблице определить значение коэффициента корреляции двух показателей.

3. Получить уравнения регрессии и построить их графики.

4. Вычислить ожидаемое (прогнозное) значение одной величины при значении другой, равном математическому ожиданию соответствующей величины плюс два ее средних квадратических отклонения.

5. Проанализировать полученные результаты.

ВАРИАНТЫ ЗАДАНИЯ

1. Зависимость между затратами на маркетинг и месячной прибылью фирмы

Затраты на маркетинг, тыс.р.	Прибыль фирмы, млн. р.				
	2,0–2,2	2,2–2,4	2,4–2,6	2,6–2,8	2,8–3,0
10 – 20	4	–	–	–	–
20 – 30	3	8	7	–	–
30 – 40	–	9	8	3	3
40 – 50	–	3	4	12	21
Более 50	–	–	6	17	19

2. Взаимосвязь основных производственных фондов и величиной валовой продукции предприятий

Объем валовой продукции, тыс.руб.	Среднегодовая стоимость основных фондов, тыс.руб.				
	До 80	80 – 100	100–120	120–140	более 140
До 50	6	–	–	–	–
50 – 70	2	3	–	–	–
70 – 90	5	6	2	–	–
Более 90	–	–	4	3	11

3. Зависимость между затратами на рекламу и доходом фирмы

Затраты на рекламу, тыс.р.	Доход, тыс. р.				
	100–150	150–200	200–250	250–300	300–350
5 – 10	9	–	–	3	–
10 – 15	12	4	2	6	2
15 – 20	16	8	12	1	1
20 – 25	–	3	6	13	7

4. Распределение сотрудников по стажу и выполненному объему работ

Стаж работы, лет	Объем работы за год, тыс. р.					
	18–20	20–22	22–24	24–26	26–28	28–30
0 – 5	17	–	–	–	–	–
5 – 10	6	3	1	4	–	3
10 – 15	4	5	7	8	1	–
15 – 20	–	8	6	2	8	1
20 – 25	–	–	3	–	9	20

5. Распределение производительности труда и объема товарооборота

Товарооборот на душу населения, тыс. р.	Производительность труда, тыс. р/чел.				
	15–20	20–25	25–30	30–35	35–40
0,7 – 0,9	18	1	–	–	–
0,9 – 1,1	4	12	23	1	–
1,1 – 1,3	3	7	16	13	6
1,3 – 1,5	–	–	2	17	14

6. Зависимость между рентабельностью собственных средств предприятия и процентной ставкой за кредит

Рентабельность, %	Процентная ставка, %					
	10–12	12–14	14–16	16–18	18–20	20–22
18 – 20	–	–	2	1	3	4
20 – 22	2	3	9	2	2	1
22 – 24	7	6	4	1	–	–
24 – 26	3	1	–	–	–	–

7. Простой оборудования и себестоимость продукции за месяц

Простой, ч.	Себестоимость, тыс. р.				
	5–5,2	5,2–5,4	5,4–5,6	5,6–5,8	5,8–6,0
50 – 60	5	1	–	1	–
60 – 70	2	8	2	2	1
70 – 80	1	3	9	6	7
80 – 90	–	–	4	5	9

8. Распределение жилищного фонда на душу населения и производительности труда в городах области (региона)

Жилищный фонд (m^2)	Производительность труда, тыс. р/чел				
	14–18	18–22	22–26	26–30	30–34
Менее 9	2	1	–	2	–
9 – 12	9	7	3	–	1
12 – 15	4	14	7	1	2
15 – 18	–	3	18	13	3
Более 18	–	–	2	2	6

9. Зависимость между затратами на НИР и себестоимостью продукции

Затраты на НИР, тыс. р.	Себестоимость, тыс. р.				
	2,0–2,2	2,2–2,4	2,4–2,6	2,6–2,8	2,8–3,0
10 – 20	–	6	1	4	9
20 – 30	3	12	17	8	2
30 – 40	8	19	13	13	3
Более 40	14	16	4	–	–

10. Распределение населения по возрасту и количеству безработных

Возраст, лет	Количество безработных, %				
	2,0–2,2	2,2–2,4	2,4–2,6	2,6–2,8	2,8–3,0
20 – 30	12	2	–	–	1
30 – 40	3	16	6	2	3
40 – 50	4	5	12	3	15
Более 50	–	1	9	18	28

11. Взаимосвязь объема производства и затрат на управление фирмами

Объем производства, млн. р/год	Затраты на управление, тыс. р/год				
	50–60	60–70	70–80	80–90	90–100
6	8	2	–	–	–
6 – 7	2	4	6	3	4
7 – 8	1	7	10	8	6
Более 8	–	2	4	7	12

12. Зависимость между экономической рентабельностью предприятия и затратами на восстановление основных производственных фондов

Рентабельность, %	Затраты на восстановление ОПФ, тыс. р.				
	2,0–2,2	2,2–2,4	2,4–2,6	2,6–2,8	2,8–3,0
12 – 14	2	1	4	1	8
14 – 16	3	2	12	6	3
16 – 18	4	9	7	3	1
18 – 20	10	4	5	2	–

13. Взаимосвязь скорости движения автомобиля и расходом горючего

Скорость движения, км/ч	Расход горючего на 100 км пути, л				
	7–8	8–9	9–10	10–11	более 11
До 50	3	–	–	3	5
50 – 70	2	3	7	4	–
70 – 90	5	13	2	–	–
90 – 110	7	8	18	2	–
Более 110	–	–	4	5	17

14. Распределение предприятий по величине собственных и заемных средств

Заемные средства, тыс. р.	Собственные средства, тыс. р.				
	250–300	300–350	350–400	400–450	450–500
До 50	–	1	2	10	17
50 – 100	4	8	13	39	8
100 – 150	12	9	6	5	3
150 – 200	13	4	2	–	–

15. Распределение рабочих по стажу работы и месячной заработной плате

Стаж работы, лет	Заработная плата, долл./мес.				
	130–160	160–190	190–210	210–240	240–270
До 10	14	7	4	–	–
10 – 15	1	16	12	5	–
15 – 20	–	–	3	10	19
Более 20	–	–	–	18	39

16. Зависимость между количеством внесенных удобрений и урожайностью

Количество удобрений, ц/га	Урожайность, ц/га				
	20–22	22–24	24–26	26–28	28–30
До 1	18	–	–	1	7
1 – 2	8	9	2	3	10
2 – 3	2	3	17	12	24
3 – 4	1	–	9	7	11

17. Распределение фирм по величине фондов и себестоимости продукции

Основные фонды, млн. р.	Себестоимость единицы продукции, р.				
	30–34	34–38	38–42	42–46	46–50
До 60	–	–	–	–	4
60 – 70	–	4	1	4	16
70 – 80	13	12	6	19	24
80 – 100	35	13	18	15	–
Более 100	51	15	12	7	–

18. Зависимость между месячной прибылью фирмы и количеством нарушений технологического процесса производства

Количество нарушений технологического процесса	Месячная прибыль, тыс. р.				
	200–300	300–400	400–500	500–600	600–700
До 4	–	–	2	3	13
4 – 8	1	4	2	12	9
8 – 12	8	9	7	–	1
Более 12	10	–	–	–	–

19. Зависимость между месячными затратами на рекламу и доходом фирмы

Затраты на рекламу, тыс. р.	Доход, тыс. р.				
	100–150	150–200	200–250	250–300	300–350
До 10	9	–	–	1	–
10 – 15	12	4	2	6	2
15 – 20	16	8	12	8	9
20 – 25	–	3	6	13	7

20. Распределение сотрудников фирмы по стажу работы и выполненному за год объему работ

Стаж, лет	Объем работы, тыс. р.					
	18–20	20–22	22–24	24–26	26–28	28–30
0 – 5	3	–	–	–	6	–
5 – 10	15	3	1	4	2	3
10 – 15	4	15	27	8	1	12
15 – 20	–	8	16	2	22	16
20 – 25	–	–	3	–	–	30

21. Распределение магазинов по торговой площади и ежедневной выручкой

Торговая площадь, (кв. метры)	Ежедневная выручка, тыс. р.					
	2–4	4–6	6–8	8–10	10–12	более 12
10 – 30	2	2	–	–	–	–
30 – 50	6	3	3	–	–	–
50 – 70	4	8	6	1	2	1
70 – 90	–	4	7	2	4	8
Более 90	–	–	–	2	3	10

22. Зависимость между рентабельностью собственных средств предприятия и средним износом оборудования

Износ оборудования, %	Рентабельность, %					
	10–12	12–14	14–16	16–18	18–20	20–22
До 20	–	–	2	1	23	8
20–40	2	13	20	7	2	1
40–60	14	16	3	4	–	–
Более 60	13	1	–	–	–	–

23. Простой оборудования и себестоимость продукции предприятия за месяц

Простой, ч.	Себестоимость, тыс. р.				
	5–5,2	5,2–5,4	5,4–5,6	5,6–5,8	5,8–6,0
До 20	5	1	–	–	–
20 – 40	2	8	2	2	9
40 – 60	1	13	19	6	7
60 – 80	–	–	4	5	23
Более 80	–	–	1	15	9

24. Зависимость между рентабельностью собственных средств предприятия и величиной заемных средств

Рентабельность, %	Заемные средства, тыс. р.				
	200–300	300–400	400–500	500–600	600–700
18 – 20	–	–	1	1	13
20 – 22	1	4	2	12	9
22 – 24	3	9	7	4	1
24 – 26	10	2	–	–	–

25. Распределение производительности труда и объема товарооборота на душу населения

Товарооборот, тыс. р.	Производительность труда, тыс. р/чел.				
	15–20	20–25	25–30	30–35	35–40
0,7 – 0,9	2	1	–	–	–
0,9 – 1,1	4	5	3	1	–
1,1 – 1,3	3	7	6	5	6
1,3 – 1,5	–	–	2	7	3
1,5 – 1,7	–	–	–	6	5

ЗАДАНИЕ 6.

Исследовать влияние на количественную переменную качественного показателя (показателей).

ВАРИАНТЫ ЗАДАНИЯ

1. Результаты испытания трех сортов пшеницы

Сорт пшеницы	Урожайность по участкам, ц/га
1	18,3; 19,4; 20,8; 17,4; 20,3; 19,8; 20,3; 20,0; 19,3; 17,8
2	17,9; 22,4; 23,1; 20,5; 17,3; 19,7; 20,1; 19,4; 20,5
3	19,1; 23,6; 21,5; 22,7; 22,8; 23,4; 22,1; 20,8; 19,3; 23,1; 23,0

2. Влияние типа удобрения на урожайность ячменя

Тип удобрения	Урожайность, ц/га
Калийное	20,1; 20,4; 19,3; 26,7; 21,4; 22,5; 19,3; 19,8; 22,4; 19,8; 29,1; 19,7; 20,4;
Фосфатное	22,3; 23,4; 19,8; 22,3; 24,1; 22,3; 23,7; 22,6; 22,8
Азотное	20,5; 20,7; 22,4; 21,7; 21,4; 20,8; 23,5; 24,6; 23,1; 22,0

3. Данные испытаний озимой пшеницы

Сорт	Урожайность, ц/га
Новоукраинка 84	32,4; 31,8; 33,5; 30,7; 34,0; 33,2; 31,7; 31,6; 34,0; 33,2
Безостая 4	41,4; 44,5; 43,0; 42,7; 41,6; 43,8; 42,2; 45,0; 41,9
Скороспелка 3	42,8; 40,3; 40,5; 40,3; 41,4; 41,3; 40,8; 39,2; 40,5; 40,3

4. Зависимость уровня вибрации (в % от предельно допустимого) от марки пылесоса

Марка пылесоса	Уровень вибрации
“Метеор”	76, 80, 74, 77, 89, 73, 70, 91, 84, 73
“Полнос”	86, 79, 73, 91, 92, 85, 87, 93, 78, 97, 84
“Надежда”	64, 80, 91, 78, 66, 68, 74, 83, 76, 71

5. Зависимость разрывной нагрузки волокна от способа изготовления

Способ изготовления	Разрывная нагрузка, кг/мм ²				
1	180	160	180	140	165
2	196	156	200	160	150
3	280	270	296	208	220
4	150	190	230	170	180

6. Данные испытания износа шин мотоциклов после пробега в 8 тыс. км

Марка мотоцикла	Износ шин, мм
“ИЖ”	8,2; 9,0; 8,6; 7,8; 8,1; 7,4; 8,9; 9,9; 8,4
“Ява”	6,5; 8,5; 6,6; 8,0; 7,3; 7,1; 6,0; 5,8
“Урал”	7,6; 7,5; 7,8; 7,4; 7,8; 8,9; 7,1; 7,2; 7,9

7. Зависимость урожайности (ц/га) яблок от сорта и вида почв

Сорт	Почва			
	Чернозем	Торфяники	Супесчаники	Глиноземы
Пепин	220	270	230	210
Антоновка	270	280	225	215
Славянка	245	274	231	228
Уэлси	264	251	243	214
Богатырь	277	284	245	202

8. Прирост удоев молока (в литрах) в зависимости от вида добавок в рацион кормления для различных пород коров

Порода коров	Виды добавок				
	1	2	3	4	5
1	10	12	11	15	18
2	25	30	27	18	35
3	40	43	47	42	41
4	33	40	45	37	36

9. Зависимость скорости реакции (мин) от типа катализатора и среды окисления

Катализатор	Атмосферный воздух	Чистый кислород	Обогащенный кислородом воздух
Железо	61,5	60,1	60,3
Перманганат калия	54,2	53,1	54,2
Платина	50,0	49,0	49,1
Марганец	52,2	52,0	52,0

10. Урожайность картофеля (ц/га) в зависимости от вида удобрений и сроков посева

Сроки посева	Виды удобрений			
	Торф	Калийная соль	Суперфосфат	Сернокислый аммоний
Раньше срока	261	266	230	240
Своевременно	265	269	237	239
С опозданием	263	260	229	236

11. Данные о месячной прибыли филиалов предприятия

Филиал предприятия	Прибыль, тыс.руб.
1	35,8; 43,5; 30,7; 34,0; 33,2; 31,7; 31,6; 34,0; 33,2
2	31,4; 44,5; 43,0; 48,7; 41,6; 43,8; 42,2; 45,0; 41,9
3	32,8; 40,3; 40,5; 45,3; 41,4; 41,3; 40,8; 39,2

12. Зависимость урожайности (ц/га) яблок от сорта и вида почв

Сорт	Почва			
	Чернозем	Торфяники	Супесчаники	Глиноземы
Пепин	245	250	215	220
Антоновка	260	280	225	215
Славянка	245	274	231	228
Уэлси	264	256	243	214
Богатырь	270	284	245	200

13. Данные испытания износа шин мотоциклов после пробега в 7 тыс. км

Марка мотоцикла	Износ шин, мм
“ИЖ”	8,2; 9,0; 8,6; 7,8; 8,1; 7,4; 8,9; 9,9; 8,4; 8,0; 7,3; 7,1;
“Ява”	6,5; 8,5; 6,6; 8,0; 7,3; 7,1; 6,0; ; 6,0; 7,0; 7,3; 5,8
“Урал”	7,6; 7,5; 7,8; 7,4; 7,8; ; 8,0; 7,3; 7,1; 8,9; 7,1; 7,2; 7,9

14. Данные испытаний озимой пшеницы

Сорт	Урожайность, ц/га
Новоукраинская	31,8; 33,5; 30,7; 34,0; 33,2; 31,7; 31,6; 34,0; 33,2
Богатырская	43,4; 44,5; 43,0; 42,7; 41,6; 43,8; 42,2; 45,0; 41,9
Скороспелка	40,3; 41,4; 42,8; 40,3; 40,5; 40,3; 41,4; 41,3; 40,8; 39,2

15. Зависимость скорости реакции (мин) от типа катализатора и среды окисления

Катализатор	Атмосферный воздух	Чистый кислород	Обогащенный кислородом воздух
Железо	62,5	60,6	59,3
Перманганат калия	54,2	53,1	54,7
Платина	50,0	49,0	49,1
Марганец	51,2	54,0	52,0

16. Зависимость урожайности (ц/га) овса от сорта и вида почв

Сорт овса	Почва			
	Чернозем	Торфяники	Супесчаники	Глиноземы
С1	25,0	27,0	23,0	21,0
С2	27,0	28,0	22,5	21,3
С3	21,5	27,4	18,1	22,8
С4	20,4	25,1	20,3	21,4
С5	19,7	19,4	18,5	20,2

17. Зависимость уровня вибрации (в % от предельно допустимого) от марки холодильника

Марка холодильника	Уровень вибрации
“ОКА-6”	76, 80, 74, 77, 89, 73, 70, 91, 84, 73, 78, 66, 68, 79
“Стинол”	77, 86, 79, 73, 91, 92, 85, 87, 93, 78, 97, 84
“Полнос”	64, 80, 77, 66, 68, 91, 78, 66, 78, 74, 83, 76, 71

18. Данные о безотказном сроке работы двигателей различных типов

Тип двигателя	Срок службы, месяцев
Тип 1	32,4; 28,8; 33,5; 30,7; 34,0; 33,2; 31,7; 31,6; 34,0; 33,9
Тип 2	41,4; 46,5; 43,0; 42,7; 41,6; 43,8; 42,2; 45,0; 41,9
Тип 3	42,8; 40,3; 40,5; 40,3; 31,4; 41,3; 40,8; 39,2

19. Месячный прирост веса (кг) в зависимости от вида добавок в рацион кормления для различных пород бычков

Порода бычков	Виды добавок				
	Д1	Д2	Д3	Д4	Д5
1	18	17	19	15	18
2	24	30	27	18	25
3	20	27	18	25	21
4	23	22	24	21	20

20. Данные о месячной прибыли филиалов предприятия

Филиал предприятия	Прибыль, тыс.руб.
1	55,8; 43,5; 45,3; 41,4; 41,3; 34,0; 33,2; 31,7; 31,6; 34,0; 33,2
2	31,4; 44,5; 43,0; 48,7; 51,6; 43,8; 52,2; 45,0; 41,9
3	42,8; 40,3; 40,5; 45,3; 41,4; 41,3; 40,8; ; 45,3; 41,4; 41,3

21. Данные о безотказном сроке работы приборов различных типов

Тип прибора	Срок службы, месяцев
Тип 1	32,4; 28,8; 33,5; 30,7; 34,0; 33,2; 31,7; 31,6; 34,0; 33,9
Тип 2	26,4; 36,5; 43,0; 42,7; 41,6; 33,8; 42,2; 35,0; 41,9
Тип 3	22,8; 40,3; 40,5; 40,3; 31,4; 41,3; 40,8; 39,2

22. Прибыль магазина (млн. руб.) от продажи телевизоров

Марка телевизора	Квартал года			
	Зима	Весна	Лето	Осень
СОНИ	25,0	27,0	23,0	26,0
ПАНАСОНИК	27,0	28,0	22,5	21,3
ГОРИЗОНТ	21,5	20,4	18,1	22,8
ЭЛЕКТРОН	20,4	25,1	20,3	23,4
ВИТЯЗЬ	29,7	24,4	30,5	27,2

23. Данные о месячной прибыли цехов предприятия

№ цеха	Прибыль, тыс.руб.
1	55,8; 43,5; 41,4; 41,3; 34,0; 33,2; 31,7; 31,6; 34,0; 33,2
2	31,4; 44,5; 43,0; 48,7; 51,6; 43,8; 52,2; 45,0; 41,9
3	42,8; 40,3; 40,5; 45,3; 41,4; 41,3; 40,8; 45,3; 41,4; 41,3

24. Прибыль магазина (млн. руб.) от продажи пылесосов

Марка пылесоса	Квартал года			
	Зима	Весна	Лето	Осень
Тайфун	25,0	27,0	23,0	26,0
Ветерок	27,0	28,0	22,5	21,3
Вихрь	21,5	20,4	18,1	22,8
Элка	20,4	25,1	20,3	23,4
Москва	29,7	24,4	30,5	27,2

25. Данные о коэффициентах загрузки оборудования по цехам предприятия

№ цеха	Коэффициент загрузки оборудования по месяцам (%)
1	44,4; 43,5; 41,4; 41,3; 34,0; 33,2; 31,7; 31,6; 34,0; 33,2; 40,5
2	52,2; 44,5; 43,0; 48,7; 51,6; 43,8; 52,2; 45,0; 41,9; 43,5
3	42,1; 40,3; 40,5; 45,3; 41,4; 41,3; 40,8; 45,3; 41,4; 41,3

26. Данные о безотказном сроке работы приборов различных типов

Тип прибора	Срок службы, месяцев
Тип 1	33,5; 30,7; 28,8; 33,5; 30,7; 34,0; 33,2; 31,7; 31,6; 34,0; 33,9
Тип 2	26,4; 36,5; 43,0; 33,5; 39,7; 42,7; 41,6; 33,8; 42,2; 35,0; 41,9
Тип 3	22,8; 40,3; 40,5; 40,3; 31,4; 41,3; 40,8; 33,5; 30,7; 39,2

27. Прибыль магазина (млн. руб.) от продажи стиральных машин

Марка машины	Квартал года			
	Зима	Весна	Лето	Осень
Сименс	25,8	27,0	23,0	26,0
Сибирь	27,4	28,0	22,5	21,3
Ардо	21,5	20,4	18,1	22,8
Индезит	21,4	25,1	20,3	23,4
Аристон	29,7	24,4	30,5	27,2

ЗАДАНИЕ 7.

По имеющимся данным (табл.) получить уравнение множественной линейной регрессии $Y = a_0 + a_1X_1 + a_2X_2 + a_3X_3$ и проанализировать качество полученной модели.

Варианты заданий приведены ниже в таблице.

Варианты задания для многомерного регрессионного анализа

№ вар.	Значения показателей												
1	X_1	3,1	3,4	2,8	2,9	3	2,7	3,2	3,1	3,5	4	2,6	3,4
	X_2	15	12	13	11	18	20	18	20	14	13	10	12
	X_3	0,6	0,4	0,6	0,5	0,8	1	1,3	0,9	0,4	0,3	0,6	0,8
	Y	13,5	16,2	14,6	15,9	13	11,5	18,4	11,3	14,1	16,8	16	20
2	X_1	3	3,5	4,6	3,2	4,3	3,4	3,9	3,1	3,5	4,7	4,1	3,3
	X_2	0,1	0,3	0,8	0,2	0,1	0,3	0,5	0,3	0,1	0,9	0,6	0,7
	X_3	1,5	1,3	1,8	1,3	1,7	1,5	1,1	1,4	1,9	1,4	1,2	1,7
	Y	1,5	4,3	10,1	3,1	5,4	4,2	7,1	3,2	2,2	11,9	8,3	5,2
3	X_1	1,4	1,7	1,3	1,4	1,3	1,1	1,7	1,2	1,6	1,4	1,2	1,5
	X_2	16	9	16	8	10	15	8	16	11	15	8	12
	X_3	0,3	0,8	0,2	0,1	0,3	0,5	0,3	0,7	0,4	0,7	0,9	0,6
	Y	10,1	5,8	9,3	1,3	3,4	9,1	2,6	11,2	5,9	10,1	4,2	7,3
4	X_1	23	32	26	29	20	29	23	26	34	24	36	29
	X_2	8	6	5	9	5	8	6	5	9	6	5	7
	X_3	38	34	41	44	37	43	36	39	45	37	41	37
	Y	114	91	136	128	133	124	118	129	117	125	119	107
5	X_1	91	122	110	93	125	109	113	95	120	97	121	103
	X_2	438	422	423	424	410	439	413	430	432	419	437	420
	X_3	204	219	231	206	222	218	207	223	233	211	216	228
	Y	960	841	621	854	740	901	888	691	666	777	987	604
6	X_1	16	18	20	17	18	16	19	20	18	21	19	16
	X_2	0,3	0,4	0,6	0,5	0,8	0,6	0,1	0,9	0,4	0,3	0,6	0,8
	X_3	9,4	8,8	9,3	8,3	9,2	8,9	8,5	9,4	9,5	8,6	8,9	9
	Y	0,2	3,5	5,1	3,8	4	2	4,1	5,7	2,2	6,2	5,3	2,4
7	X_1	0,79	0,86	0,94	0,85	0,95	0,81	0,88	0,91	0,83	0,96	0,84	0,89
	X_2	3,03	3,22	3,33	3,05	3,07	3,19	3,34	3,09	3,41	3,12	3,15	3,37
	X_3	1,19	0,99	1,26	0,97	1,36	1,41	1,04	1,43	1,39	1,29	1,07	1,14
	Y	1,69	3,05	2,34	2,77	1,55	1,18	3,15	1,18	1,75	1,96	2,54	2,84
	X_1	3,3	3,1	2,8	2,9	3	2,7	3,6	3,1	3,5	4	2,6	3,4

8	X_2	8,8	9,5	8,5	8,9	8,6	8,1	9,5	9,1	8,3	8,7	8,9	8,3
	X_3	3,8	3,1	4	3,7	3,4	3,9	3,5	3,2	3,8	4	3,9	3,3
	Y	9,3	14	5,5	8,3	10	4,7	14	13,2	9	10,4	5,9	12,3
9	X_1	3,5	3,2	3,1	3,5	4	2,6	3,4	2,9	3,7	3,3	2,6	3,4
	X_2	14	13	10	12	13	16	17	20	14	13	19	12
	X_3	0,9	0,6	0,7	0,5	0,8	1	1,3	0,9	0,9	1	1,2	0,8
	Y	19,9	15	19,3	17,1	21	14	20,6	10,5	20,5	20,7	13	19,8
10	X_1	0,59	0,71	0,79	0,61	0,69	0,76	0,57	0,65	0,77	0,62	0,78	0,67
	X_2	0,03	0,15	0,13	0,21	0,06	0,16	0,07	0,13	0,18	0,08	0,23	0,09
	X_3	1,28	1,22	1,29	1,26	1,01	1,27	1,32	1,03	1,33	1,09	1,17	1,19
	Y	26,1	28,8	30,1	28,7	23,8	29,7	27,4	24,8	30,7	24,3	29,6	26,1
11	X_1	1,2	1,6	1,4	1,9	1,5	1,8	1,2	1,3	1,7	1,4	1,2	1,5
	X_2	15	8	16	11	15	8	12	13	13	14	8	12
	X_3	0,7	0,4	0,7	0,9	0,6	0,7	0,9	0,1	0,4	0,7	0,9	0,6
	Y	10,4	2,9	11,4	8,5	10	4,4	8	6,3	8,1	9,3	4	7,4
12	X_1	3,5	3,7	3,2	3,1	3,9	3,7	3,8	3,1	3,4	3,9	3,1	3,6
	X_2	0,3	0,1	0,5	0,6	0,7	0,1	0,4	0,7	0,2	0,4	0,3	0,7
	X_3	1,4	1,1	1,5	1,1	1,7	1,2	1,2	1,3	1,9	1,8	1,6	1,3
	Y	5,7	4,3	6,8	5,9	8,5	4,7	6,2	6,8	7	7,9	6	7,3
13	X_1	20	32	26	35	20	30	23	27	37	24	39	29
	X_2	15	14	18	16	11	15	17	12	18	13	19	14
	X_3	35	39	35	45	42	38	34	41	44	37	43	36
	Y	77	78	48	80	125	69	56	89	63	85	49	65
14	X_1	20	28	14	29	16	33	23	30	34	25	32	27
	X_2	41	57	49	42	59	55	44	53	58	47	54	48
	X_3	14	11	19	15	12	16	19	12	17	13	15	18
	Y	87	109	103	117	81	150	127	115	166	100	137	143
15	X_1	12,1	10,4	11,6	12,6	10,6	11,4	11,9	10,9	12,8	12	11	11,3
	X_2	7,4	8,2	9	7,6	8,8	8,3	7,7	9,1	7,8	8,4	8,2	8,8
	X_3	3,1	4,6	4,3	5,3	3,8	5,1	3,9	4,6	4	3,7	4,2	4
	Y	87	82	90	92	84	88	87	87,4	92,4	90	85	87,9
16	X_1	9,7	8,9	9,6	8,1	9,9	9,7	9,2	8,3	9,4	8,4	9,4	8,6
	X_2	13,3	13,6	13,8	14,4	14	14,6	13,9	14,7	14,1	13,4	14,3	13,4
	X_3	22,5	21	22	20	21,5	20,5	23	21,5	20,5	22	20	23,5
	Y	41,6	42,1	40,4	42,3	38,1	36,6	46	46,7	37,6	46,4	37	48
	X_1	3,4	3,2	3,6	3,9	3,7	3,1	3,9	3,7	3,8	3,1	3,9	3,4

17	X_2	0,3	0,5	0,6	0,8	0,1	0,4	0,7	0,8	0,4	0,2	0,7	0,3
	X_3	1,4	1,1	1,4	1,9	1,4	1,2	1,6	1,3	1,9	1,7	1,9	1,4
	Y	5,7	5,4	7	10	5,3	5,3	8,5	7,8	8	6	9,3	5,5
18	X_1	16	13	10	20	19	14	11	17	15	12	18	13
	X_2	7	11	6	3	5	9	3	8	6	5	4	8
	X_3	3	4	4	2	9	2	1	3	1	7	5	6
	Y	7	8	14	3	20	3	10	5	3	22	15	17
19	X_1	18	9	10	15	17	11	12	7	20	14	8	16
	X_2	63	80	79	64	73	81	70	66	76	68	78	69
	X_3	23	28	15	29	16	33	24	30	34	27	32	29
	Y	150	173	98	183	119	209	147	167	240	172	191	194
20	X_1	3,6	4,6	4,3	5,3	3,8	4,9	3,9	4,6	4,1	3,8	4,5	4
	X_2	7,6	8,2	9	7,9	9,1	8,3	7,7	9,8	7,8	8,9	9	8,8
	X_3	17,5	19,6	19,4	17,6	19,9	18	18,3	18,3	19,4	17,4	18,9	19
	Y	60	68	68,8	68,4	67	68,4	61,9	71,1	64,3	64,1	70	66,1
21	X_1	7,4	8,2	9	7,6	8,8	8,3	7,7	9,1	7,8	8,4	7,9	9,1
	X_2	13,9	14,7	14,1	13,4	14,3	13,4	13,9	14,9	14	13,9	14,6	14,7
	X_3	11	12	17	20	14	8	16	17	15	14	16	19
	Y	18,1	18,9	29	45,4	21,3	4,9	32,5	29,3	29	22,4	33,1	35,2
22	X_1	10,5	9,3	9,7	9,2	10,1	9,4	10,3	9,1	9,4	10,3	9,2	10,2
	X_2	2	2,8	2,6	2,9	2,1	3,1	2,6	2,4	2,7	3,1	2,5	2,7
	X_3	5,1	6,5	5,9	6,3	5,8	6,5	5,4	6,6	6	5,6	6,2	5,7
	Y	29,1	39,6	34,4	35,5	36,6	36,1	28,5	43,1	34,7	27,7	38,3	31
23	X_1	21,3	22,9	22,4	23,1	21,3	22,4	21,4	23,4	22,6	21,5	22,8	20,8
	X_2	32,4	32,8	30,9	33	32,1	31,4	30,1	31,8	30,4	33,2	31,9	30,5
	X_3	10,3	11,9	11,1	11,8	10,1	11,4	11,3	10,8	11,4	10,4	11,7	10,6
	Y	50,3	52,3	45,4	52,9	48,1	50	49,1	41,2	46,6	52,5	49,9	48,1
24	X_1	4,5	4,1	4,6	4,1	4,2	4,6	4,2	4,3	4,8	4,4	4,5	4,5
	X_2	0,3	0,5	0,3	0,7	0,4	0,7	0,9	0,6	0,7	0,9	0,1	0,3
	X_3	1,4	1,1	1,4	1,9	1,4	1,2	1,6	1,3	1,9	1,7	1,9	1,4
	Y	7	6,3	7,1	9,6	6,8	8,2	10	7,6	10,3	10,1	7,6	6,6
25	X_1	64	57	65	51	56	58	63	54	59	67	55	60
	X_2	63	80	79	64	73	81	70	66	76	68	78	69
	X_3	18	9	10	15	17	11	12	7	20	14	8	16
	Y	262	206	233	208	242	222	231	170	272	254	190	245
	X_1	3,6	3,5	4	2,6	3,4	2,9	3,7	3,3	4,1	3,9	4,6	3,2

26	X_2	10,6	11,4	11,9	10,9	12,8	13	11,4	10,9	13,1	11,8	11,3	12,5
	X_3	18,3	19,6	19,4	20,4	19,9	20,6	18,3	18,3	19,4	20,1	18,9	20,3
	Y	69	72	77,1	67,1	76	75,1	72	68,9	82,2	78	76,8	74,8
27	X_1	23	32	26	29	20	29	23	26	34	24	36	29
	X_2	14	12	15	13	11	16	11	16	14	12	13	15
	X_3	93	73	92	84	95	75	96	79	90	81	99	83
	Y	304	220	296	263	333	212	334	239	272	266	300	244
28	X_1	10,5	11,3	11,7	11,2	10,9	11,4	10,3	11,1	11,4	10,3	11,2	10,2
	X_2	3,6	3,2	4,9	4,6	3,8	3,2	4,7	3,9	3,4	4	3,5	4,3
	X_3	7,1	6,6	6	6,8	6,9	7,6	7	6,4	7,2	6,7	7,3	7,1
	Y	39,6	36,3	17,2	27,5	34,4	45,3	29,9	28,6	39,9	31,7	41,3	33,7
29	X_1	20,1	19,7	20,6	18,3	18,3	19,4	20,1	18,9	20,3	21,5	19,9	20,8
	X_2	31,9	31,8	30,9	31,2	32,1	31,4	30,1	31,8	30,4	33	31,9	30,5
	X_3	14,9	15	13,8	14,4	14	14,6	13,9	14,7	14,1	13,4	14,3	13,4
	Y	83,6	85,9	70,3	86,5	86,9	83	71,2	87,1	72,3	69,9	80,8	67
30	X_1	4,5	4,1	4,6	4,1	4,2	4,6	4,2	4,3	4,8	4,4	4,5	4,6
	X_2	0,3	0,5	0,3	0,7	0,4	0,7	0,9	0,6	0,7	0,9	0,1	0,4
	X_3	5,6	5	5,6	5,1	5,8	5,2	5,5	5,9	5,3	6	5,4	5,5
	Y	12,3	11	12,4	12,2	13	13,1	14,3	14,4	13,3	16,2	11	12,9

Задание 8.

На конкурсе проектов n участников получили некоторые оценки (по сто-бальной системе) за два основных качественных (ранговых) показателя (варианты задания). Связаны ли между собой эти показатели? Дать физическую интерпретацию показателям. Проанализировать полученные результаты.

Варианты задания

Вариант 1

№ проекта	1	2	3	4	5	6	7	8	9	10	11
Показатель 1	68	59	60	75	78	53	52	55	59	80	80
Показатель 2	53	65	55	78	89	63	64	65	51	75	75

Вариант 2

№ проекта	1	2	3	4	5	6	7	8	9	10
Показатель 1	50	66	39	60	75	78	53	52	55	59
Показатель 2	63	53	35	55	68	89	63	64	65	51

Вариант 3

№ проекта	1	2	3	4	5	6	7	8	9
Показатель 1	63	66	89	60	75	78	53	52	59
Показатель 2	53	53	80	55	78	89	43	61	65

Вариант 4

№ проекта	1	2	3	4	5	6	7	8	9	10	11
Показатель 1	50	68	59	60	75	78	53	52	55	59	70
Показатель 2	53	53	65	55	78	89	63	64	65	51	75

Вариант 5

№ проекта	1	2	3	4	5	6	7	8	9	10
Показатель 1	59	58	63	57	60	61	54	52	55	56
Показатель 2	60	49	62	50	61	65	53	54	58	48

Вариант 6

№ проекта	1	2	3	4	5	6	7	8	9
Показатель 1	43	66	89	60	75	78	53	52	59
Показатель 2	43	53	80	50	78	89	43	61	65

Вариант 7

№ проекта	1	2	3	4	5	6	7	8	9	10
Показатель 1	60	85	78	53	52	75	59	75	59	88
Показатель 2	55	78	89	63	64	65	51	65	51	76

Вариант 8

№ проекта	1	2	3	4	5	6	7	8	9	10
Показатель 1	59	58	63	57	60	61	54	52	55	56
Показатель 2	60	49	62	50	61	65	53	54	58	48

Вариант 9

№ проекта	1	2	3	4	5	6	7	8	9
Показатель 1	43	66	89	60	75	78	53	52	59
Показатель 2	43	53	80	50	78	89	43	61	65

Вариант 10

№ проекта	1	2	3	4	5	6	7	8	9	10
Показатель 1	96	89	60	85	78	53	52	75	59	90
Показатель 2	83	85	55	78	89	63	64	65	51	75

Вариант 11

№ проекта	1	2	3	4	5	6	7	8	9	10
Показатель 1	67	68	59	60	75	78	53	52	55	80
Показатель 2	73	53	65	55	78	89	63	64	65	75

Вариант 12

№ проекта	1	2	3	4	5	6	7	8	9	10
Показатель 1	50	66	39	60	75	78	53	52	55	59
Показатель 2	63	53	35	55	68	89	63	64	65	51

Вариант 13

№ проекта	1	2	3	4	5	6	7	8	9
Показатель 1	63	66	89	60	75	78	53	52	59
Показатель 2	53	53	80	55	78	89	43	61	65

Вариант 14

№ проекта	1	2	3	4	5	6	7	8	9	10
Показатель 1	50	68	59	60	75	78	53	52	55	59
Показатель 2	53	53	65	55	78	89	63	64	65	51

Вариант 15

№ проекта	1	2	3	4	5	6	7	8	9	10
Показатель 1	59	58	63	57	60	61	54	52	55	56
Показатель 2	60	49	62	50	61	65	53	54	58	48

Вариант 16

№ проекта	1	2	3	4	5	6	7	8	9
Показатель 1	63	66	89	60	75	78	53	52	59
Показатель 2	53	53	80	55	78	89	43	61	65

Вариант 17

№ проекта	1	2	3	4	5	6	7	8	9	10
Показатель 1	47	68	59	60	75	78	53	52	55	59
Показатель 2	53	53	65	55	78	89	63	64	65	51

Вариант 18

№ проекта	1	2	3	4	5	6	7	8	9	10
Показатель 1	59	58	63	57	60	61	54	52	55	56
Показатель 2	59	49	62	50	61	65	53	54	58	48

Вариант 19

№ проекта	1	2	3	4	5	6	7	8	9
Показатель 1	71	66	89	60	75	78	53	52	59
Показатель 2	53	53	80	55	78	89	43	61	65

Вариант 20

№ проекта	1	2	3	4	5	6	7	8	9	10	11
Показатель 1	50	68	59	60	75	78	53	52	55	59	50
Показатель 2	53	53	65	55	78	89	63	64	65	51	69

Вариант 21

№ проекта	1	2	3	4	5	6	7	8	9	10
Показатель 1	59	58	63	57	60	61	54	52	55	56
Показатель 2	68	49	62	50	61	65	53	54	58	48

Вариант 22

№ проекта	1	2	3	4	5	6	7	8	9
Показатель 1	38	66	89	60	75	78	53	52	59
Показатель 2	43	53	80	50	78	89	43	61	65

Вариант 23

№ проекта	1	2	3	4	5	6	7	8	9
Показатель 1	63	66	89	60	75	78	53	52	59
Показатель 2	53	53	80	55	78	89	43	61	65

Вариант 24

№ проекта	1	2	3	4	5	6	7	8	9	10
Показатель 1	50	68	59	60	75	78	53	52	55	59
Показатель 2	53	53	65	55	78	89	63	64	65	51

Вариант 25

№ проекта	1	2	3	4	5	6	7	8	9	10
Показатель 1	59	58	63	57	60	61	54	52	55	56

Показатель 2	60	49	62	50	61	65	53	54	58	48
--------------	----	----	----	----	----	----	----	----	----	----

Вариант 26

№ проекта	1	2	3	4	5	6	7	8	9
Показатель 1	63	66	89	60	75	78	53	52	59
Показатель 2	53	53	80	55	78	89	43	61	65

Вариант 27

№ проекта	1	2	3	4	5	6	7	8	9	10
Показатель 1	47	68	59	60	75	78	53	52	55	59
Показатель 2	53	53	65	55	78	89	63	64	65	51

Вариант 28

№ проекта	1	2	3	4	5	6	7	8	9	10
Показатель 1	59	58	63	57	60	61	54	52	55	56
Показатель 2	59	49	62	50	61	65	53	54	58	48

Вариант 29

№ проекта	1	2	3	4	5	6	7	8	9
Показатель 1	71	66	89	60	75	78	53	52	59
Показатель 2	53	53	80	55	78	89	43	61	65

Вариант 30

№ проекта	1	2	3	4	5	6	7	8	9
Показатель 1	68	66	89	60	75	78	53	52	88
Показатель 2	43	53	84	50	78	89	43	61	65

Задание 9.

Группа из нескольких экспертов оценивает качество однотипной продукции, выпускаемой на нескольких предприятиях.

Предпочтения экспертов (их ранги) показателя представлены в таблице.

Взаимосвязаны (согласуются) ли мнения экспертов?

Рассчитать коэффициент конкордации и оценить его значимость.

Варианты задания

Вариант 1

Эксперт	Предприятия					
	1	2	3	4	5	6
1	3	1	4	2	5	6
2	4	2	5	3	6	1
3	2	1	4	5	6	3
4	2	3	4	6	1	5
5	3	6	1	4	2	5

Вариант 2

Эксперт	Предприятия						
	1	2	3	4	5	6	7
1	5	3	1	5	4	2	6

2	1	2	5	3	6	4	7
3	2	1	7	5	6	4	3
4	1	2	4	6	3	5	7

Вариант 3

Эксперт	Предприятия						
	1	2	3	4	5	6	7
1	4	3	1	2	6	7	5
2	2	1	5	3	6	4	7
3	2	1	7	5	6	4	3
4	1	2	4	6	3	5	7

Вариант 4

Эксперт	Предприятия						
	1	2	3	4	5	6	7
1	3	1	4	2	6	7	5
2	1	2	5	3	6	4	7
3	2	1	7	5	6	4	3
4	3	1	5	4	2	6	7

Вариант 5

Эксперт	Предприятия						
	1	2	3	4	5	6	7
1	4	3	1	2	6	7	5
2	4	2	5	3	1	6	7
3	2	1	7	5	6	4	3
4	2	1	4	6	3	5	7
5	3	1	5	4	2	7	6

Вариант 6

Эксперт	Предприятия						
	1	2	3	4	5	6	7
1	2	1	5	7	6	3	4
2	2	1	7	5	6	4	3
3	2	1	4	6	3	5	7
4	3	1	5	4	2	7	6

Вариант 7

Эксперт	Предприятия						
	1	2	3	4	5	6	7
1	4	3	1	2	6	7	5
2	2	3	7	5	6	4	1
3	2	1	7	5	6	4	3
4	2	1	4	6	3	5	7
5	3	1	5	4	2	7	6

Вариант 8

Эксперт	Предприятия						
	1	2	3	4	5	6	7
1	5	3	4	2	6	7	1
2	1	2	3	4	6	5	7
3	2	1	6	5	7	4	3
4	1	2	4	6	3	5	7

5	3	1	5	4	2	6	7
---	---	---	---	---	---	---	---

Вариант 9

Эксперт	Предприятия						
	1	2	3	4	5	6	7
1	6	3	2	4	1	7	5
2	3	2	5	1	6	4	7
3	2	1	7	5	6	4	3
4	1	2	4	6	3	5	7
5	3	1	5	4	2	6	7

Вариант 10

Эксперт	Предприятия						
	1	2	3	4	5	6	7
1	2	1	5	7	6	4	3
2	1	2	5	3	6	4	7
3	2	3	7	5	6	4	1
4	1	2	4	6	7	5	3
5	3	1	5	4	2	6	7

Вариант 11

Эксперт	Предприятия					
	1	2	3	4	5	6
1	3	1	4	2	5	6
2	4	2	5	3	6	1
3	1	2	4	5	6	3
4	2	3	4	6	1	5
5	3	1	6	4	2	5

Вариант 12

Эксперт	Предприятия						
	1	2	3	4	5	6	7
1	4	3	1	2	6	7	5
2	2	3	5	7	4	6	1
3	3	1	7	5	6	4	2
4	2	4	1	6	3	5	7
5	3	5	4	1	2	7	6

Вариант 13

Эксперт	Предприятия						
	1	2	3	4	5	6	7
1	2	1	5	7	6	4	3
2	4	2	5	6	3	1	7
3	2	1	7	5	6	4	3
4	1	2	4	6	3	5	7
5	3	1	5	4	2	7	6

Вариант 14

Эксперт	Предприятия					
	1	2	3	4	5	6
1	4	1	3	2	6	5

2	4	2	5	3	6	1
3	2	1	4	5	6	3
4	2	4	3	6	1	5
5	3	1	6	4	2	5

Вариант 15

Эксперт	Предприятия					
	1	2	3	4	5	6
1	3	1	4	2	5	6
2	4	2	5	3	6	1
3	2	1	4	5	6	3
4	3	6	4	2	1	5

Вариант 16

Эксперт	Предприятия					
	1	2	3	4	5	6
1	4	1	3	2	6	5
2	4	2	5	3	6	1
3	2	1	4	5	6	3
4	2	4	3	6	1	5
5	3	6	1	4	2	5

Вариант 17

Эксперт	Предприятия					
	1	2	3	4	5	6
1	3	1	4	2	5	6
2	4	2	5	3	6	1
3	2	1	4	5	6	3
4	2	3	4	6	1	5
5	3	1	6	4	2	5

Вариант 18

Эксперт	Предприятия					
	1	2	3	4	5	6
1	3	1	4	2	5	6
2	5	2	4	3	6	1
3	2	1	4	5	6	3
4	2	3	4	6	1	5
5	3	1	6	4	2	5

Вариант 19

Эксперт	Предприятия					
	1	2	3	4	5	6
1	1	4	3	2	6	5
2	4	2	5	3	6	1
3	2	1	4	5	6	3
4	2	4	3	6	1	5
5	3	1	6	4	2	5

Вариант 20

Эксперт	Предприятия					
	1	2	3	4	5	6
1	3	1	4	2	5	6
2	4	2	5	3	6	1

3	2	1	4	5	6	3
4	3	2	4	6	1	5

Вариант 21

Эксперт	Предприятия						
	1	2	3	4	5	6	7
1	6	3	2	4	1	7	5
2	3	2	5	1	6	4	7
3	2	1	7	5	6	4	3
4	6	2	4	1	3	5	7
5	3	1	5	4	2	6	7

Вариант 22

Эксперт	Предприятия						
	1	2	3	4	5	6	7
1	5	3	4	2	6	7	1
2	1	2	3	4	6	5	7
3	2	1	6	5	7	4	3
4	4	2	1	6	3	5	7
5	3	1	5	4	2	6	7

Вариант 23

Эксперт	Предприятия						
	1	2	3	4	5	6	7
1	6	4	2	3	1	7	5
2	3	2	5	1	6	4	7
3	2	1	7	5	6	4	3
4	1	2	4	6	3	5	7
5	3	5	1	4	2	6	7

Вариант 24

Эксперт	Предприятия					
	1	2	3	4	5	6
1	3	1	2	4	5	6
2	4	2	5	3	6	1
3	2	1	4	5	6	3
4	2	3	4	6	1	5
5	3	6	1	4	2	5