

Министерство образования Республики Беларусь
Учреждение образования
Международный государственный экологический
университет имени А. Д. Сахарова



Факультет экологической медицины
Кафедра радиационной гигиены и эпидемиологии

ЖИВИЦКАЯ Е. П.

**СТАТИСТИЧЕСКИЙ АНАЛИЗ
И ОБРАБОТКА ДАННЫХ
С ПРИМЕНЕНИЕМ SPSS**

Методические указания по проведению
лабораторных работ по дисциплине
«Статистические пакеты прикладных программ»

Минск
2010

УДК 311:002
ББК 60.6:73
Ж66

*Рекомендовано к изданию НМС МГЭУ им. А.Д.Сахарова
(протокол № 6 от 16 марта 2010 г.)*

Автор:

Е. П. Живицкая, ст. преподаватель кафедры радиационной гигиены
и эпидемиологии МГЭУ им. А. Д. Сахарова

Рецензенты:

профессор кафедры «Экология» БНТУ, д.т.н., профессор *В. П. Бубнов*;
в.н.с. НИЛ антропоэкологии и эпидемиологии
МГЭУ им. А. Д. Сахарова, к.м.н. *С. В. Петренко*

Живицкая, Е. П.

Ж66 Статистический анализ и обработка данных с применением SPSS :
метод. указания / Е. П. Живицкая. – Минск : МГЭУ им. А. Д. Сахарова,
2010. – 52 с.

ISBN 978-985-6931-46-1.

Методические указания составлены в соответствии с базовой программой по модулю специализации «Статистические пакеты прикладных программ» и предназначены для использования при проведении лабораторных работ по данной дисциплине. Пособие включает методические указания по выполнению лабораторных работ, контрольные вопросы для проверки усвоения учебного материала и задания для самоконтроля, список литературы, которую студенты могут использовать для углубленного изучения данной статистической программы.

УДК 311:002
ББК 60.6:73

ISBN 978-985-6931-46-1

© Живицкая Е. П., 2010
© Международный государственный
экологический университет
имени А. Д. Сахарова, 2010

Введение

В настоящее время существует множество программных средств для статистической обработки данных. Одним из наиболее популярных среди таких пакетов является пакет SPSS (аббревиатура SPSS расшифровывается как Statistical Package for Social Science – Статистический пакет для социальных наук), который предоставляет большой выбор функций для анализа и обработки данных. SPSS – самый часто используемый пакет статистической обработки. Пакет программ SPSS широко используется специалистами медико-биологического профиля, предоставляет пользователю широкие возможности по статистической обработке эмпирических данных, по формированию баз данных (файлов данных SPSS с возможностью импорта/экспорта в файлы данных других форматов), по их модификации по мере необходимости, а также по созданию отчетов, предоставляя широкие возможности по представлению результатов статистической обработки в текстовой, табличной и графической формах (диаграммы, гистограммы и т. п.). Интерфейс программы доброжелателен к пользователю любого уровня. Использование простого меню делает возможным выполнение сложных видов статистического анализа без использования командной строки синтаксиса.

Основным достоинством программного комплекса SPSS является возможность анализа и обработки данных социологических опросов, анкетирования и др. Кроме того, SPSS предоставляет широкий спектр существующих статистических методов, который удачно сочетается с большим количеством удобных средств визуализации результатов обработки.

В данном пособии представлены методические указания по проведению лабораторных работ по дисциплине «Статистические пакеты прикладных программ» с применением программного продукта SPSS.

Цель курса «Статистические пакеты прикладных программ» – освоение студентами основных приемов и методов анализа данных медико-биологических исследований, обучение специфике использования методов многомерной статистики при работе с данными с применением современных компьютерных программ статистического анализа данных.

Поскольку курс ориентирован на использование современных программных средств обработки данных на персональных компьютерах, важной задачей является обучение студентов работе с программной системой SPSS.

Задание № 1

Создание базы данных и основы работы с медико-биологическими данными

Цель: ознакомиться с функциональными возможностями пакета программ SPSS на примере медико-биологических данных.

Ход работы

1.1. Создание базы данных. Ввод данных

Методические указания: при выполнении работы используйте ключ к лабораторной работе (Приложение 1, таблица 1); после выполнения каждого задания сохраняйте матрицу данных в своей папке.

Определите переменные, как предложено в таблице 1.1.

Таблица 1.1

Определяемые переменные

Имя (name)	Тип (type)	Ширина (width)	Десятичные знаки (decimals)	Метка переменной (label)	Метки значений (values)
ID	numeric	2	0	'Личный номер'	
Name	string	20	0	'Имя больного'	
Gender	string	1	0	'Пол больного'	м 'мужской' ж 'женский'
Protein	numeric	4	1	'Содержание общего белка в крови, г/л'	
Glucose	numeric	3	1	'Содержание глюкозы в крови, ммоль/л'	
Chole	numeric	3	1	'Содержание холестерина в крови, ммоль/л'	

Вид в SPSS представлен на рис. 1.1.

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align
1	id	Numeric	3	0	Личный номер	None	None	8	Right
2	name	String	20	0	Имя больного	None	None	11	Right
3	gender	String	1	0	Пол больного	(ж, женский)...	None	8	Right
4	protein	Numeric	3	1	Содержание общего белка в крови, г/л	None	None	8	Right
5	glucose	Numeric	3	1	Содержание глюкозы в крови, ммоль/л	None	None	8	Right
6	chole	Numeric	3	1	Содержание холестерина в крови, ммоль/л	None	None	8	Right

Рис. 1.1. Вкладка Variable View

Создайте базу данных, состоящую из 25 человек, на основе предложенной (Приложение 1, таблица 2).

Добавьте в список еще одного больного:

– вставьте пустую строку после строки № 20;

– присвойте больному № 26;

– введите для него произвольные имя и пол и следующие результаты: 'Total protein' = 85,4; 'Glucose' = 4,1; 'Cholesterol' = 4,7.

Создайте переменную 'Result' 'отношение к норме':

– вставьте новую переменную после переменной 'Total protein';

– определите и задайте все параметры новой переменной;

– вычислите значения новой переменной через перекодировку результатов 'Total protein' (recode).

Проверьте правильность вычисления переменной 'Result':

– упорядочьте больных по убыванию переменной 'Total protein';

– убедитесь, что уровни содержания соответствуют результатам;

– при необходимости исправьте ошибки в правилах перекодировки.

Проверьте качество ввода данных по одномерным распределениям (рис. 1.2.) переменных 'Gender', 'Total protein', 'Result', 'Glucose', 'Cholesterol'.

При необходимости исправьте ошибки ввода.

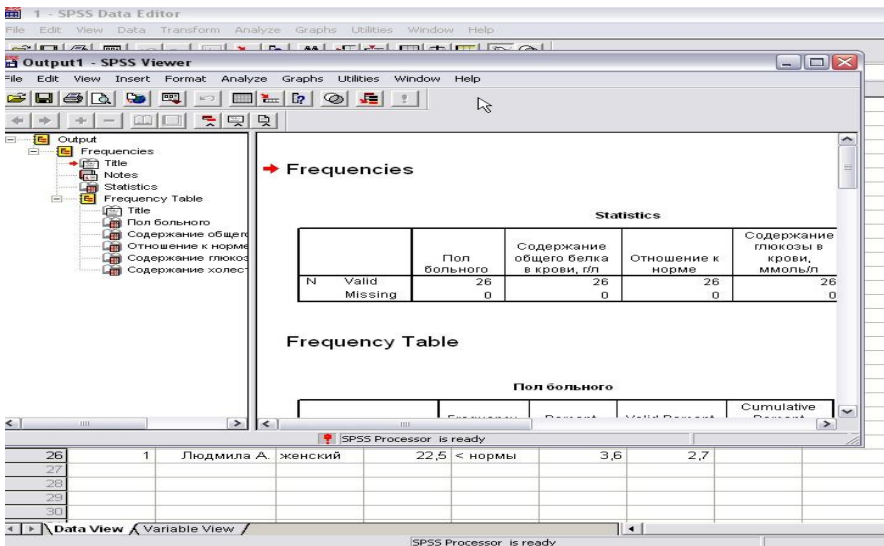


Рис. 1.2. Вид одномерных распределений в SPSS

Постройте таблицы сопряженности (рис. 1.3) между переменными 'Total protein' и 'Glucose'; 'Total protein' и 'Cholesterol'.

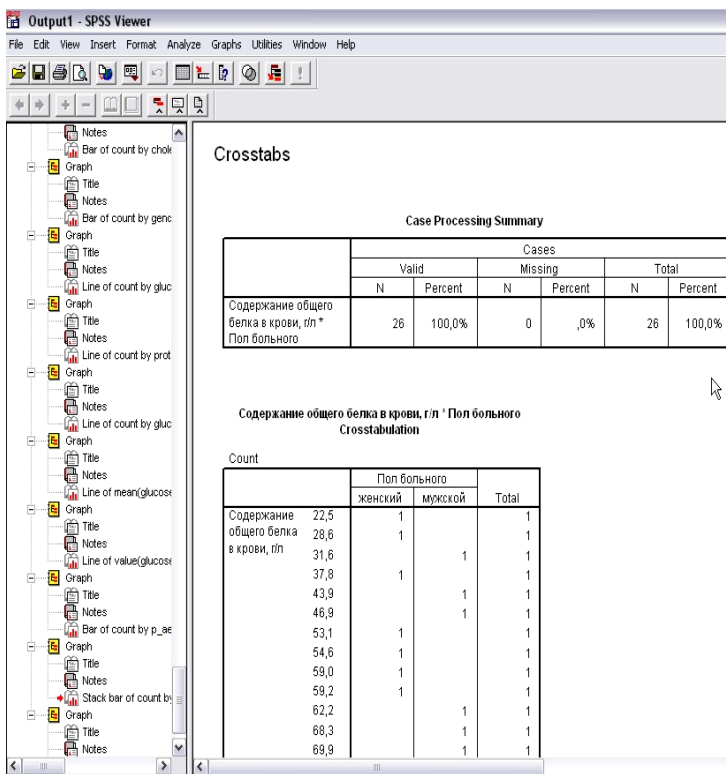


Рис. 1.3. Вид таблицы сопряженности в SPSS

Создайте переменную Glucose 1 'содержание глюкозы в крови при повторном анализе':

- определите новую переменную, задайте для нее необходимые параметры (рис. 1.4);

- вычислите $Glucose\ 1 = Glucose + 0,2$;

- постройте одномерное распределение переменной 'Glucose 1'.

Осуществите «набор» из числа больных, у которых содержание белка в крови находится выше нормы:

- упорядочьте матрицу данных по убыванию переменной 'Total protein';

- определите верхнее значение нормы;

- создайте новую дихотомическую переменную Proteinaemia (0 – 'нет', 1 – 'да');

- вычислите переменную 'Proteinaemia', используя опцию **if**.

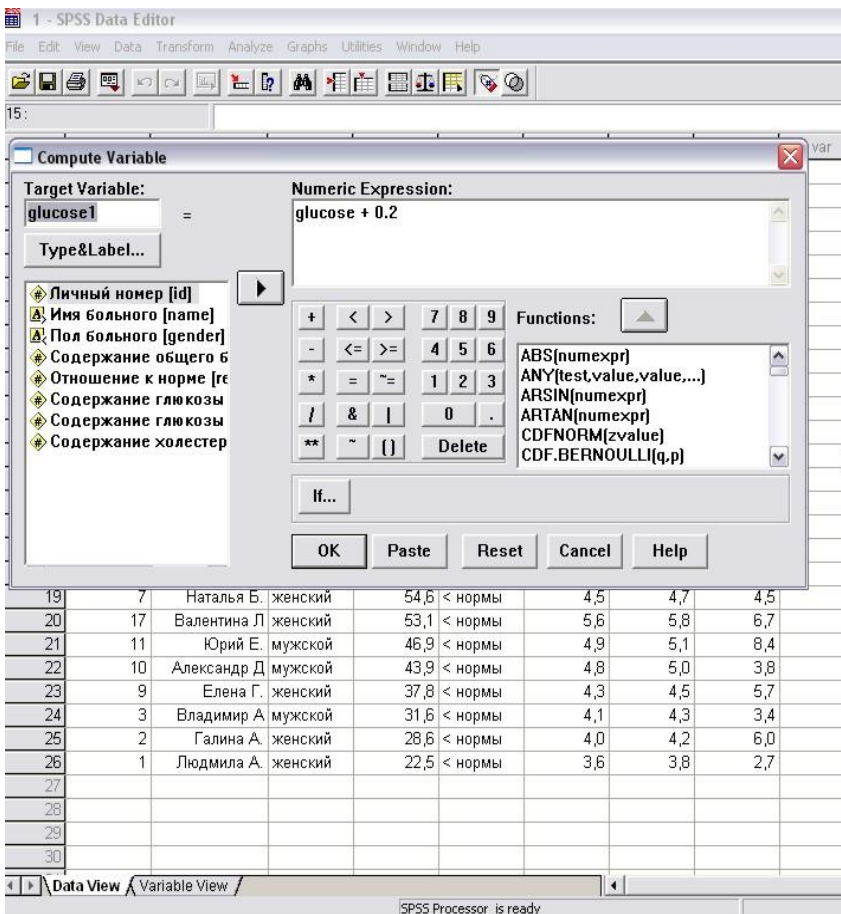


Рис. 1.4. Вкладка Compute Variable

1.2. Статистическая обработка данных и построение графиков

Посчитайте описательную статистику (рис. 1.5) для переменных 'Glucose' и 'Chole':

- посчитайте среднее значение переменной, моду, медиану, дисперсию, минимум и максимум выборки, размах вариационного ряда, накопленную частоту;
- удалите несколько значений, посмотрите, как изменились показатели выборки;
- обратите внимание на значения валидного процента.

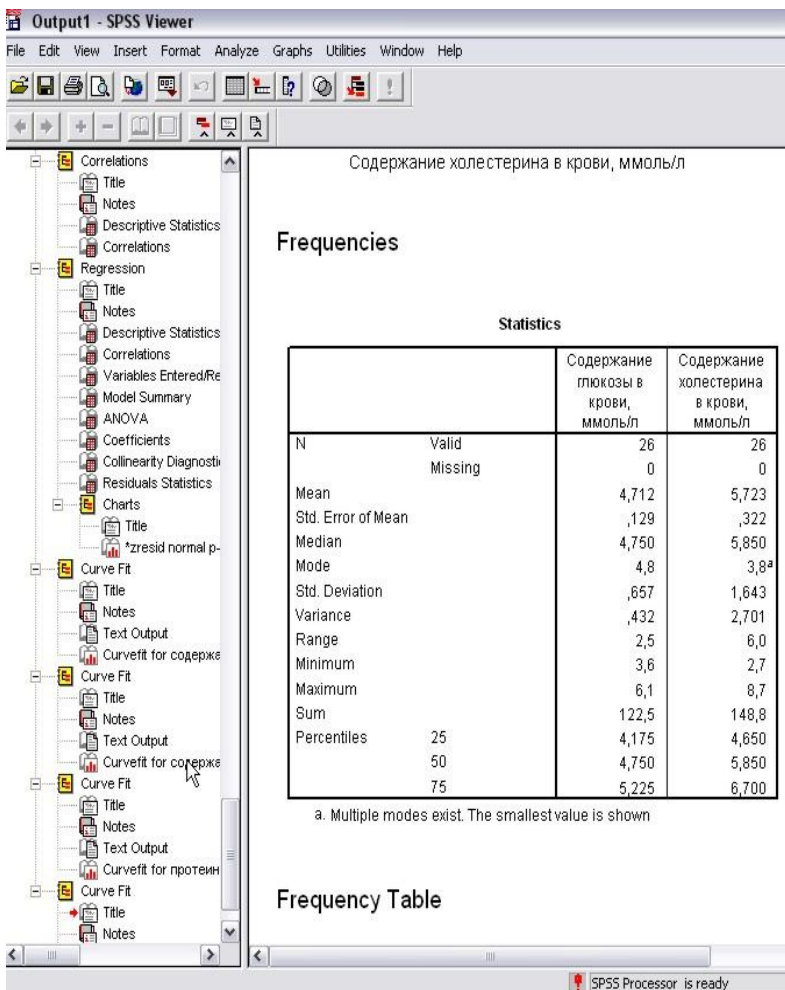


Рис. 1.5. Вид таблицы с описательной статистикой в SPSS

Проведите корреляционный анализ (рис. 1.6) между переменными Glucose и Glucose 1.

Постройте:

- круговую диаграмму по переменной 'Result';
- диаграмму столбцов по переменной 'P_aemia';
- кластеризованную диаграмму (различия по полу больных протеинемией);
- ленточную диаграмму (половая структура по отношению к переменной 'Result').

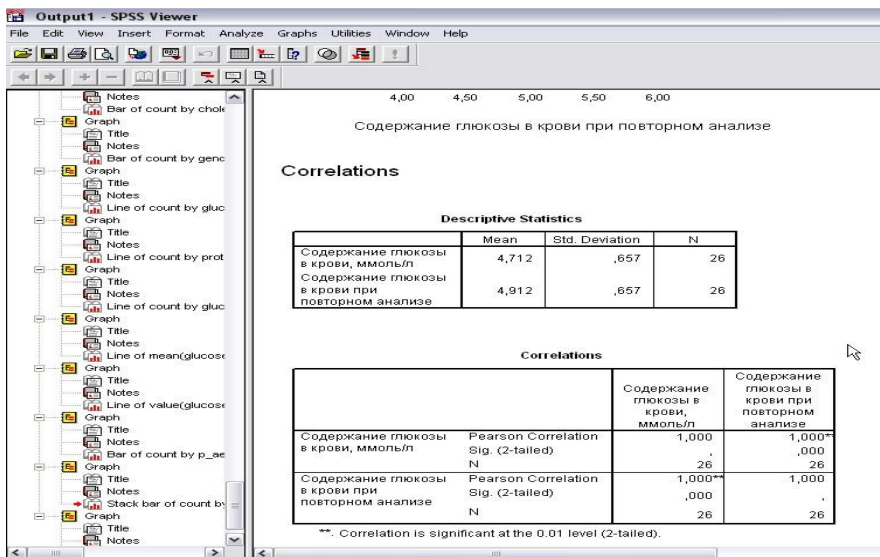


Рис. 1.6. Вид корреляционной таблицы в SPSS

1.3. Подготовка отчета

Сохраните результаты работы:

- отсортируйте матрицу данных по возрастанию переменной ID, сохраните ее;
- отредактируйте файл результатов, сохранив только правильно выполненные задания, сохраните его в своей папке.

Подготовьте в MS Word отчет «врачебной комиссии», проиллюстрируйте его соответствующими таблицами и сохраните в своей папке.

Отчет должен содержать следующие разделы:

1. Цель работы.
2. Краткие теоретические сведения, описание программного продукта.
3. Созданную базу данных на основе предложенной.
4. Результаты исследований и их анализ.
5. Выводы по работе.

Задание № 2

Статистический анализ и обработка медико-биологических данных

Цель работы: закрепить навыки работы и освоить распространенные тесты для проверки гипотез о среднем.

Ход работы

2.1. Создание базы данных. Ввод данных

Определите переменные и создайте базу данных на основе предложенной (Приложение 2, таблица 1).

2.2. Описательная статистика

Перейдите к исследованию данных, выбрав команды меню Analyze (Анализ) Descriptive Statistics (Дескриптивные статистики) Explore... (Исследовать). Откроется диалоговое окно Explore (рис. 2.1).

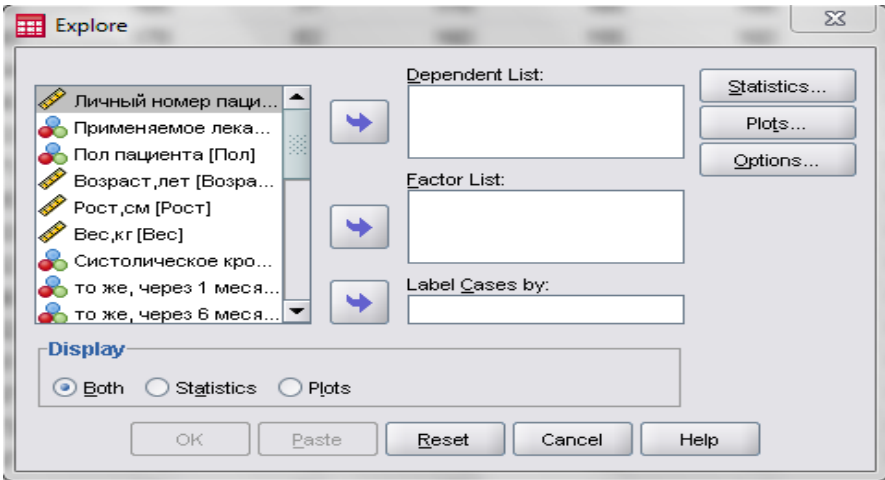


Рис. 2.1. Диалоговое окно Explore

В этом диалоговом окне проводится различие между зависимыми переменными и факторами. Это означает, что можно выполнять анализ отдельно по группам наблюдений. В этом случае анализируемой переменной будет зависимая переменная, а группирующей переменной – фактор. Если же такой отдельный анализ проводить не требуется, список факторов не используется.

Проведите анализ возраста пациентов (анализ без группирующей переменной):

– выясните, какие методы анализа выполняются по умолчанию, поэтому не вносите никаких изменений в настройки;

– перенесите переменную 'Возраст' в список зависимых переменных (Dependent List);

– запустите вычисление, щелкнув на кнопке ОК.

Будут созданы таблицы 2.1–2.3:

Таблица 2.1

Case Processing Summary (Обработанные наблюдения)

	Cases (Случаи)					
	Valid (Допустимые)		Missing (Отсутствующие)		Total (Всего)	
	N	Percent	N	Percent	N	Percent
Возраст, лет	174	100,0 %	0	0 %	174	100,0 %

Таблица 2.2

Descriptives (Описательная статистика)

			Statistic	Std. Error
Возраст, лет	Mean (Среднее)		62,11	,875
	95 % Confidence Interval for Mean (95 % доверительный интервал среднего)	Lower Bound (Нижняя граница)	60,38	
		Upper Bound (Верхняя граница)	63,84	
	5 % Trimmed Mean (5 % усеченное среднее)		62,25	
	Median (Медиана)		63,00	
	Variance (Дисперсия)		133,358	
	Std. Deviation (Стандартное отклонение)		11,548	
	Minimum (Минимум)		36	
	Maximum (Максимум)		87	
	Range (Размах)		51	
	Interquartile Range (Межквартильная ширина)		17	
	Skewness (Асимметрия)		-,143	,184
Kurtosis (Коэффициент вариации)		-,635	,366	

Stem-and-Leaf Plot (диаграмма ветвей и листьев)

Возраст, лет

Frequency Stem & Leaf

```

6,00 3 . 677999
7,00 4 . 0223333
14,00 4 . 66677788888999
23,00 5 . 0111111112222333333444
20,00 5 . 556677777888888899
27,00 6 . 0000111112223333333444444
27,00 6 . 55555666666677888888999999
24,00 7 . 000000011111122233333444
13,00 7 . 5566666788899
11,00 8 . 00001111224
2,00 8 . 67

```

Stem width: 10

Each leaf: 1 case(s)

В этом случае окно вывода результатов содержит:

- статистические характеристики;
- диаграмму stem-and-leaf (ветвей и листьев);
- коробчатую диаграмму (box plot).

Новые характеристики:

- 5 % усеченное среднее: среднее значение, вычисленное без учета 5 % наименьших и 5 % наибольших значений;
- 95 % доверительный интервал: доверительный интервал, в котором находится среднее значение с вероятностью 95 %;
- межквартильная широта: расстояние между первым и третьим квартилями.

Диаграмма ветвей и листьев представляет собой комбинацию гистограммы и табличного списка. Как на гистограмме, длина каждой строки соответствует количеству наблюдений, попадающих в определенный интервал. Но, сверх этого, на данной диаграмме выводится также наблюдаемое численное значение для каждого наблюдения. Для этой цели численное значение разбивается на два компонента: ветвь, представляющую собой первую цифру или группу цифр, и лист – последующие цифры. Ветвь соответствует тем разрядам численного значения наблюдаемой переменной, которые не изменяются, а листья – разрядам, которые изменяются в пределах избранного интервала. В рассматриваемом примере ветви разбиты на две части: одну – для листьев с 0 по 4 и другую – для листьев с 5 по 9.

Коробчатая диаграмма состоит из прямоугольника, занимающего пространство от первого до третьего квартиля (т. е. от 25 до 75 процентиля). Линия внутри этого прямоугольника соответствует медиане. Кроме того, на коробчатой диаграмме отмечаются максимальное и минимальное значения, если только они не являются выбросами (рис. 2.2).

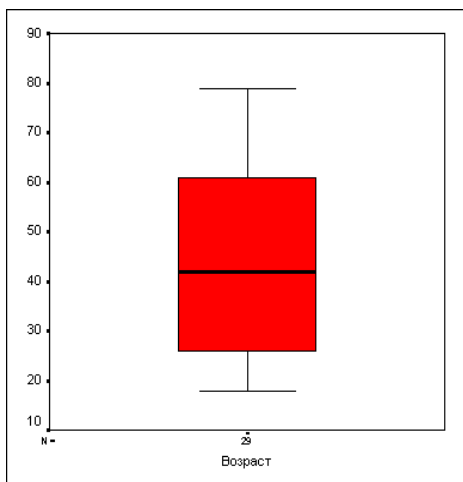


Рис. 2.2. Коробчатая диаграмма

Проведите анализ для групп наблюдений. Проанализируйте исходное содержание холестерина (переменная 'Хол0') для четырех возрастных классов (переменная 'Класс').

В диалоговом окне Explore кнопкой Reset (Сброс) восстановите настройки по умолчанию и перенесите переменную Хол0 в список зависимых переменных (Dependent List), а переменную 'Класс' – в список факторов (Factor List). Щелкните на кнопке ОК.

В результате будут вычислены характеристики описательной статистики и построена диаграмма ветвей и листьев отдельно по четырем возрастным классам. На коробчатой диаграмме соответственно появятся четыре прямоугольника.

Остальные статистические параметры также можно вычислить отдельно по разным значениям группирующей переменной (в данном случае по возрастным классам). Это относится и к выводу гистограмм и диаграмм нормального распределения в окне просмотра.

2.3. Построение диаграмм

В диалоговом окне Explore щелкните на кнопке Plots... (Диаграммы). Откроется диалоговое окно Explore: Plots (рис. 2.3).

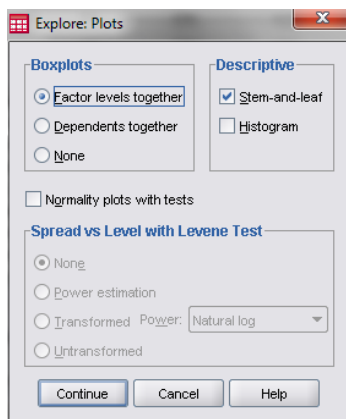


Рис. 2.3. Диалоговое окно Explore: Plots

С коробчатой диаграммой и диаграммой ветвей и листьев уже ознакомились, поэтому в поле **Boxplots** (Коробчатые диаграммы) выберите опцию **None** (Нет) и снимите флажок **Stem-and-leaf**; вместо него установите флажок **Histogram** (Гистограмма).

Щелкните на кнопке **Continue**, а затем на **OK**. В окне просмотра появится гистограмма (рис. 2.4).

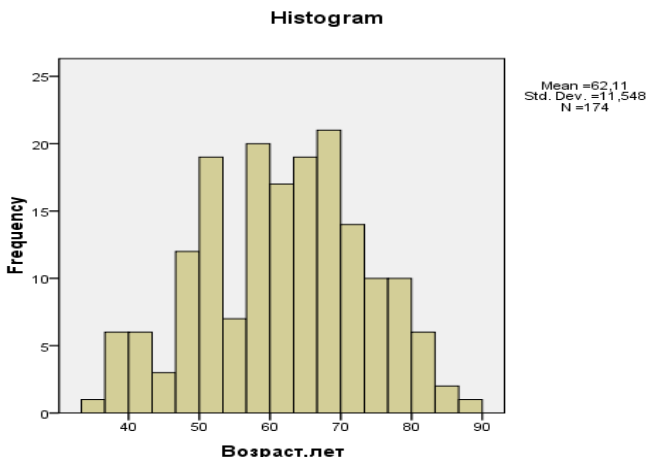


Рис. 2.4. Гистограмма возрастной структуры

Далее посмотрите, какие результаты можно получить, если установить в диалоговом окне **Explore: Plots** флажок **Normality plots with tests** (Диаграмма нормального распределения с тестами).

Установите этот флажок и подтвердите настройку кнопкой **OK** (рис. 2.5).

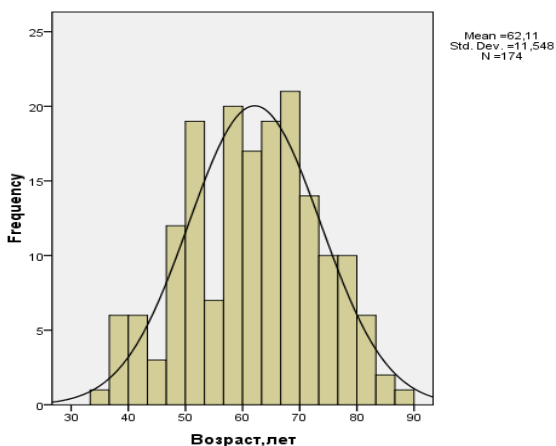


Рис. 2.5. Кривая нормального распределения

В окне просмотра будет показан результат теста Лиллифора (модификации теста Колмогорова-Смирнова) на нормальное распределение.

Таблица 2.4

Tests of Normality (Тест на нормальное распределение)

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Возраст, лет	,059	174	,200*	,987	174	,094

a. Lilliefors Significance Correction (Это нижняя граница истинной значимости)

*. This is a lower bound of the true significance (Коррекция значимости по Лиллифору)

Если в результате получена вероятность ошибки p менее 0,05, то данное распределение значимо отличается от нормального. В данном примере при $p = 0,200$ распределение можно считать нормальным.

При объеме выборки менее 50 наблюдений проводится также тест Шапиро-Уилкса.

В окне просмотра будут показаны две диаграммы:

- диаграмма нормального распределения;
- диаграмма с исключенным трендом.

По диаграмме нормального распределения (также называемой диаграммой Q-Q) можно визуальнo определить, достаточно ли близко заданное распределение приближается к нормальному. Здесь каждое наблюдаемое значение сравнивается со значением, ожидаемым при нормальном распределении. При условии точного выполнения нормального распределения все точки лежат на прямой. Наблюдаемые значения откладываются по оси X, а ожидаемые – по оси Y, при этом все значения подвергаются стандартизации (z -преобразованию). В данном примере (см. рис. 2.6) наблюдаемые значения достаточно близки к прямой.

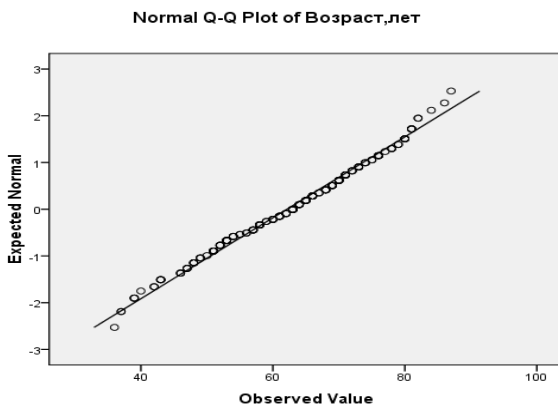


Рис. 2.6. Диаграмма нормального распределения

На диаграмме с исключенным трендом отклонения наблюдаемых значений от ожидаемых при нормальном распределении представлены в зависимости от наблюдаемых значений. В случае нормального распределения все точки лежат на горизонтальной прямой, проходящей через нуль. Явное отклонение от прямой указывает на отличие распределения от нормального. На этой диаграмме все значения также подвергаются стандартизации (z-преобразованию) (рис. 2.7).

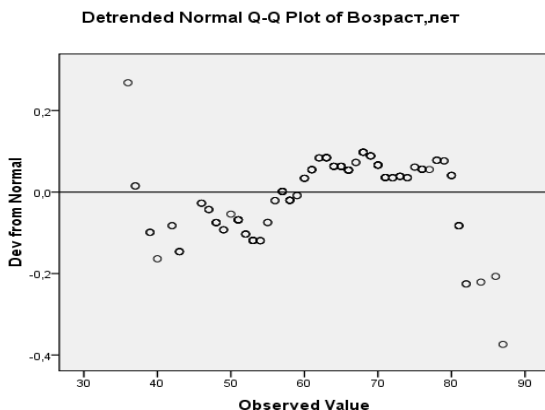


Рис. 2.7. Диаграмма с исключенным трендом

2.4. Сравнение средних

При сравнении средних значений выборок предполагается, что обе выборки подчиняются нормальному распределению. Если это не так, то вычисляются медианы и для сравнения выборок используется непараметрический тест.

При сравнении средних значений выборок выделяют четыре различные тестовые ситуации:

- сравнение двух независимых выборок;
- сравнение двух зависимых (спаренных) выборок;
- сравнение более двух независимых выборок;
- сравнение более двух зависимых выборок.

В этих ситуациях соответственно применяются следующие статистические тесты:

- t-тест для независимых выборок (тест Стьюдента);
- t-тест для зависимых выборок;
- однофакторный дисперсионный анализ;
- однофакторный дисперсионный анализ с повторными измерениями.

Выберите в меню команды Analyze (Анализ) Compare Means (Сравнение средних).

Необходимо проверить, значительно ли различается действие двух групп медикаментов на людей в зависимости от их возраста (Сравнение двух независимых выборок).

Выберите в подменю команду Independent-Samples T Test... (t-тест для независимых выборок).

Откроется диалоговое окно Independent-Samples T Test (рис. 2.8).

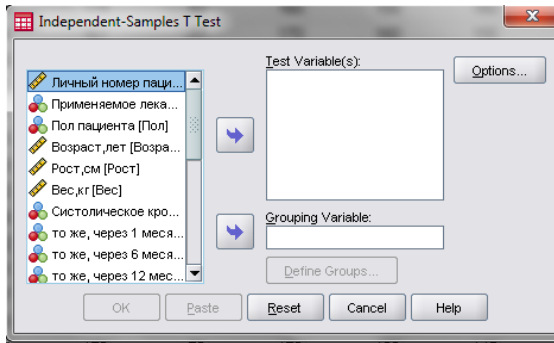


Рис. 2.8. Диалоговое окно Independent-Samples T Test

В списке исходных переменных щелкните на переменной 'Возраст' и щелчком на кнопке со стрелкой перенесите ее в список тестируемых переменных (Test Variable(s)).

Таким же способом перенесите переменную 'Лекарство' в поле Grouping Variable (Группирующая переменная).

Щелчком на кнопке Define Groups... (Определить группы) открывается окно, в котором можно ввести значения двух категорий для группирующей переменной. Нужно сравнить две группы, удовлетворяющие условиям соответственно «Лекарство = 1» и «Лекарство = 2». Поэтому внесите в поле Group 1 (Группа 1) значение 1, а в поле Group 2 – значение 2.

Щелчком на кнопке Continue вернуться в основное диалоговое окно.

Теперь следует выяснить, какие параметры установлены по умолчанию. Щелкните для этого на кнопке Options... (Параметры). Не изменяя настроек, щелкните на кнопке Continue и вернуться в основное диалоговое окно.

Запустите t-тест, щелкнув на ОК. В окне просмотра появятся следующие результаты (табл. 2.5, 2.6):

Таблица 2.5

Group Statistics (Статистика групп)

Применяемое лекарство		N	Mean (Среднее)	Std. Deviation (Стандартное отклонение)	Std. Error Mean (Стандартная ошибка среднего)
Возраст, лет	Alphasan	87	62,24	11,191	1,200
	Betasan	87	61,98	11,958	1,282

Таблица 2.6.

Independent Samples Test (Тест для независимых выборок)

		Levene's Test for Equality of Variances (Тест Левена на равенство дисперсий)		t-test for Equality of Means (Тест Стьюдента на равенство средних)						
Lower (Нижняя граница)	Upper (Верхняя граница)	F	Sig. (Значимость)	T	df	Sig. (2-tailed) (Значимость (двусторонняя))	Mean Difference (Разность средних)	Std. Error Difference (Стандартная ошибка разницы)	95 % Confidence Interval of the Difference (Доверительный интервал разницы)	
Возраст	Equal variances assumed (Дисперсии равны)	,54	,462	,151	172	,880	,26	1,76	-3,20	3,73
	Equal variances not assumed (Дисперсии не равны)			,151	171,249	,880	,26	1,76	-3,20	3,73

Выведенные результаты содержат:

– количество наблюдений, средние значения, стандартные отклонения и стандартные ошибки средних в обеих группах;

– результаты теста Левена на равенство дисперсий.

Как правило, гипотеза о равенстве (гомогенности) дисперсий не принимается, если тест Левена дает значение $p < 0,05$ (гетерогенность дисперсий). Для случаев как гомогенности (равенства), так и гетерогенности (неравенства) выводятся следующие характеристики:

– результаты t-теста: значение распределения t, количество степеней свободы df, вероятность ошибки p (под обозначением «Значимость (2-сторонняя)»);

– разница средних значений, ее стандартная ошибка и доверительный интервал.

В данном примере не наблюдается значимого различия воздействия двух групп лекарств по возрасту ($p = 0,880$).

В следующем t-тесте необходимо проверить, различается ли действие двух групп лекарств по так называемому индексу Брока. Этот индекс, разработанный одним парижским хирургом, предусматривает, что нормальный вес человека можно определить из следующего уравнения:

$$\text{Нормальный вес (кг)} = \text{Рост (см)} - 100.$$

Если взять отношение фактического веса человека к нормальному весу по этой формуле, то получается процентный показатель, который у людей с нормальным весом равен 100, у людей с избытком веса > 100 и у людей с недостатком веса < 100 .

$$\text{Индекс Брока} = ((\text{Вес в кг}) / (\text{Рост в см} - 100)) \cdot 100.$$

Определите на основе существующих переменных новую переменную, для чего выберите команды меню Transform (Преобразовать) Compute... (Вычислить).

В поле выходной переменной (Target Variable) задайте новое имя «Брок», а в поле численного выражения (Numeric Expression) введите выражение: $\text{Вес} / (\text{Рост} - 100) * 100$.

Щелкните на кнопке ОК. Теперь можно командами меню Analyze (Анализ) Compare Means (Сравнение средних) Independent Samples T Test... (t-тест для независимых выборок) описанным выше способом провести t-тест для новой переменной 'Брок'.

И этот тест показывает, что между двумя группами лекарств не наблюдается значимого различия по индексу Брока ($p = 0,233$).

Необходимо выяснить, значимо ли изменяется содержание холестерина через месяц после начала приема лекарств. Для этого сравните переменные 'Хол0' и 'Хол1' при помощи t-теста для зависимых выборок. В этом тесте будут участвовать данные всех пациентов, независимо от группы принимаемых лекарств.

Выберите в соответствующем подменю команду Paired-Samples T Test... (t-тест для парных выборок).

Откроется диалоговое окно Paired-Samples T Test (рис. 2.9).

Перенесите переменные 'Хол0' и 'Хол1' из списка исходных переменных в поле парных переменных (Paired Variables).

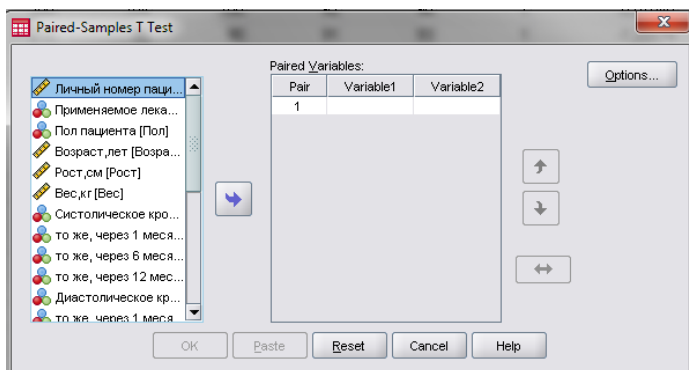


Рис. 2.9. Диалоговое окно Paired-Samples T Test

Щелкните на ОК, чтобы начать вычисления. В окне просмотра появятся результаты, которые представлены в таблицах 2.7–2.9.

Результаты, выведенные в окне просмотра, содержат:

- средние значения, количество наблюдений, стандартные отклонения и стандартные ошибки средних для обеих переменных;
- результаты t-теста: тестовая величина, полученная из распределения Стьюдента, количество степеней свободы df, вероятность ошибки p, обозначенная «Sig. (2-tailed)».

Значимого изменения содержания холестерина за один месяц после начала приема лекарств не наблюдается ($p = 0,332$).

Таблица 2.7

Paired Samples Statistics (Статистика для парных выборок)

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Холестерин, исходное значение	237,27	174	49,421	3,747
	то же, через 1 месяц	239,20	174	49,510	3,753

Таблица 2.8

Paired Samples Correlations (Корреляции для парных выборок)

		N	Correlation (Корреляция)	Sig. (Значимость)
Pair 1	Холестерин, исходное значение & то же через 1 месяц	174	,861	,000

Таблица 2.9

Paired Samples Test (Тест для парных выборок)

		Paired Differences (Парные разницы)					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95 % Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Холестерин, исходное значение – то же, через 1 месяц	-1,925	26,085	1,978	-5,828	1,978	-,974	173	,332

Повторите вычисления, но теперь только для пациентов, принимавших альфасан (переменная 'Лекарство' имеет значение 1; условие «Лекарство = 1»):

- выберите в меню команды Data (Данные) Select Cases... (Выбрать наблюдения);

- выберите опцию If condition is satisfied... (Если выполняется условие). Щелчком на кнопке If... (Если) откройте диалоговое окно, в котором можно сформулировать условие. Введите в соответствующем поле условие «Лекарство = 1»;

- щелкните на кнопке Continue, а в основном диалоге – на кнопке ОК;

- снова запустите t-тест. Теперь он будет выполнен только для наблюдений (N = 87), относящихся к первой группе лекарств. Мы снова получим незначимый результат (p = 0,666);

- чтобы последующий анализ снова можно было проводить с использованием всех наблюдений, откройте диалоговое окно Select Cases и выберите в нем опцию All cases (Все наблюдения).

Сравните более двух независимых выборок. Для этого необходимо исследовать, существует ли значимое различие веса (переменная 'Вес') между четырьмя разными возрастными группами (переменная 'Класс').

Выберите в подменю команду One-Way ANOVA... (Однофакторный дисперсионный анализ).

Подобная возможность есть и в первом пункте подменю (Means...), но она дает значительно более ограниченные возможности для анализа. Появится диалоговое окно One-Way ANOVA (рис. 2.10).

Перенесите переменную 'Вес' в список зависимых переменных (Dependent List), а переменную 'Класс' – в поле Factor (Фактор).

Посмотрите, какие параметры можно задать для этого теста (кнопка Options...). Задайте вывод описательной статистики (флажок Descriptive) и проверку на гомогенность дисперсий (флажок Homogeneity-of-variance).

Запустите тест, щелкнув на ОК.

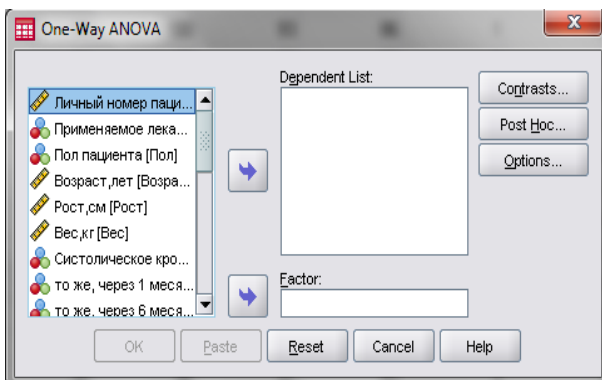


Рис. 2.10. Диалоговое окно One-Way ANOVA

В окне просмотра появятся результаты, которые представлены в таблицах 2.10–2.12.

Таблица 2.10

Descriptives (Описательная статистика)

Вес, кг								
	N	Mean	Std. Deviation	Std. Error	95 % Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
до 55 лет	52	79,96	13,949	1,934	76,08	83,85	54	129
56–65 лет	51	72,96	10,664	1,493	69,96	75,96	55	107
66–75 лет	47	70,43	11,263	1,643	67,12	73,73	50	97
> 75 лет	24	68,75	11,479	2,343	63,90	73,60	48	93
Total	174	73,79	12,638	,958	71,90	75,68	48	129

Таблица 2.11

Test of Homogeneity of Variances (Тест гомогенности дисперсий)

Вес, кг	Levene Statistic	df1	df2	Sig.
	,313	3	170	,816

Выведенные результаты содержат:

- количество наблюдений, средние значения, стандартные отклонения и стандартные ошибки средних, 95 % доверительные интервалы, минимумы и максимумы для всех слоев фактора;
- результаты теста Левена на гомогенность дисперсий;
- типовую схему дисперсионного анализа, включая вероятность ошибки p (значимость) для оценки общей значимости.

ANOVA (Дисперсионный анализ)

Вес, кг	Sum of Squares (Сумма квадратов)	df	Mean Square (Средний квадрат)	F	Sig. (Значимость)
Between Groups (Между группами)	3157,298	3	1052,433	7,310	,000
Within Groups (В группах)	24475,834	170	143,975		
Total	27633,132	173			

В этом примере дисперсионный анализ дает максимально значимый результат ($p < 0,001$).

Уменьшение веса с увеличением возраста может быть связано с тем, что в старших возрастных классах преобладают женщины, вес которых мал по сравнению с мужчинами, что и вызывает данный эффект. Повторите этот анализ для категорий пола.

Непараметрические (не основанные на каком-либо распределении вероятности) тесты применяются там, где выборки из переменных, принадлежащих к интервальной шкале, не подчиняются нормальному распределению. В этих тестах обрабатывается не само измеренное значение, а его ранг (положение внутри выборки), поэтому эти тесты нечувствительны к выбросам. Непараметрические тесты применяются также в тех случаях, когда переменные относятся к порядковой, а не к интервальной шкале. Открыть эти тесты можно в меню Analyze (Анализ) Nonparametric Tests (Непараметрические тесты).

Для сравнения двух независимых выборок (U-тест по методу Манна и Уитни) проверьте, отличаются ли показатели сахара в крови для мужчин и женщин в примере об исследовании гипертонии.

Постройте гистограмму показателя сахара в крови (переменная 'Сах0'). Обратите внимание на явную деформацию распределения в левую сторону. Тест Колмогорова-Смирнова (проведите его) также покажет очень значительное отклонение от нормального распределения. Стало быть, для сравнения обеих выборок следует вместо t-теста Стьюдента применить U-тест по методу Манна и Уитни.

Выберите в меню Analyze (Анализ) Nonparametric Tests (Непараметрические тесты) 2 Independent Samples... (Две независимые выборки).

Появится диалоговое окно Two Independent Samples Tests (Тесты для двух независимых выборок) (рис. 2.11).

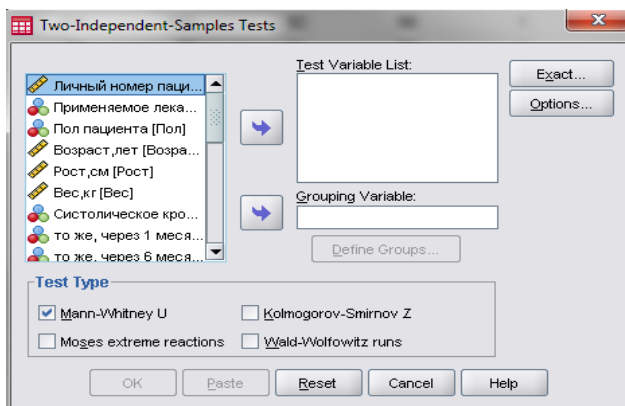


Рис. 2.11. Диалоговое окно Two Independent Samples Tests (Тесты для двух независимых выборок)

U-тест по методу Манна и Уитни является предварительной установкой. Перенесите переменную 'Сах0' из поля исходных переменных в поле тестируемых переменных, а переменную 'Пол' в поле групповых переменных.

После щелчка на кнопке Define Groups (Определить группы) вы получите возможность внести коды (1 – мужской и 2 – женский), необходимые для идентификации групп.

После щелчка на Continue вы опять попадаете в исходное диалоговое окно. Запустите программу вычисления, нажав на ОК.

В окне просмотра появятся следующие результаты (табл. 2.13, 2.14):

Таблица 2.13

Ranks (Ранги)

	Пол пациента	N	Mean Rank (Усредненный ранг)	Sum of Ranks (Ранговая сумма)
Сахар в крови, исходное значение	мужской	59	81,66	4818,00
	женский	115	90,50	10407,00
	Total	174		

Таблица 2.14

Test Statistics^a (Статистика теста)

	Сахар в крови, исходное значение
Mann-Whitney U (U-тест по Манну и Уитни)	3048,000
Wilcoxon W (Уилкоксона)	4818,000
Z	-1,096
Asymp. Sig. (2-tailed)	,273

a. Grouping Variable (Групповая переменная): Пол пациента

Выведенные результаты включают следующие показатели:

- количество наблюдений, усредненные ранги и ранговая сумма для двух выборок (причем большим значениям присваиваются низшие ранговые места);

- тестовую величину U , определенную с помощью теста Манна и Уитни;

- наименьшее значение из обеих ранговых сумм (W -тест Уилкоксона);

- точное значение вероятности ошибки p при количестве наблюдений менее 30;

- тестовую величину z , определенную по тесту Колмогорова-Смирнова, а также относящуюся к ней вероятность ошибки p , которую следует использовать при количестве наблюдений более 30.

Выясняется, что в рассматриваемом примере разница показателей сахара в крови между полами не является статистически значимой ($p = 0,273$).

Нажав кнопку Options, вы можете выбрать дополнительные возможности вывода данных, относящихся к рассмотренному и к другим непараметрическим тестам. Наряду с обычной обработкой пропущенных значений, можно организовать расчет дескриптивных статистик (среднее значение, минимум, максимум, стандартное отклонение, количество наблюдений) и квартилей (25, 50 и 75 перцентилей).

Тест Уилкоксона является традиционным непараметрическим тестом для сравнения двух зависимых выборок. Он основан на построении ранговой последовательности абсолютных разностей пар значений.

Переменные 'Сис0' и 'Сис1' (начальный уровень систолического давления и уровень через месяц после начала приема медикамента) представляют собой типичный пример связанных (зависимых) выборок.

Выберите в меню Analyze (Анализ) Nonparametric Tests... (Непараметрические тесты) 2 Related Samples... (Две связанные выборки). Вы сможете убедиться в том, что предварительно по умолчанию установлен тест Уилкоксона (рис. 2.12).

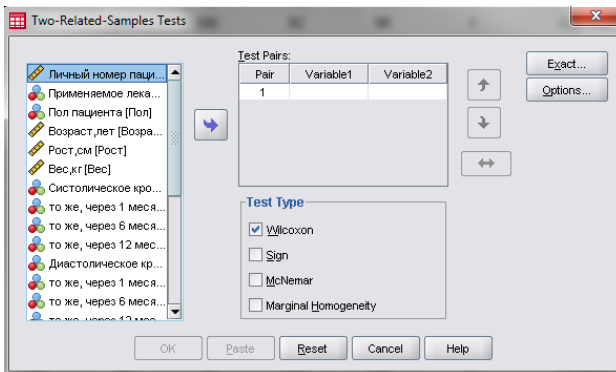


Рис. 2.12. Диалоговое окно Two-Related-Samples Tests (Тесты для двух связанных выборок)

Теперь в поле тестируемых переменных нужно выделить две необходимые переменные и эту пару перенести в поле для спаренных переменных. В нашем примере такими переменными являются 'Сис0' и 'Сис1'.

Запустите тест Уилкоксона на исполнение нажатием кнопки ОК. В окне просмотра появятся результаты расчета (табл. 2.15, 2.16).

Результаты расчета включают следующие данные:

- количества, средние ранги и ранговые суммы для отрицательных и положительных разностей (причем большим абсолютным разностям присваивается более высокое ранговое место);
- количество нулевых разностей;
- контрольную величину z с соответствующей вероятностью ошибки p .

Полученная в приведенном примере величина $p = 0,000$ свидетельствует об очень значимой разнице.

Таблица 2.15

Ranks (Ранги)

		N	Mean Rank (Средний ранг)	Sum of Ranks (Ранговая сумма)
то же, через 1 месяц – Систолическое кровяное давление, исходное значение	Negative Ranks (Отрицательные ранги)	144 ^a	77,81	11204,00
	Positive Ranks (Положительные ранги)	8 ^b	53,00	424,00
	Ties (Связи)	22 ^c		
	Total	174		

a. то же, через 1 месяц < Систолическое кровяное давление, исходное значение

b. то же, через 1 месяц > Систолическое кровяное давление, исходное значение

c. то же, через 1 месяц = Систолическое кровяное давление, исходное значение

Таблица 2.16

Test Statistics^b (Статистика теста)

	то же, через 1 месяц – Систолическое кровяное давление, исходное значение
Z	-9,970 ^a
Asymp. Sig. (2-tailed)	,000

a. Based on positive ranks (Основано на положительных рангах)

b. Wilcoxon Signed Ranks Test (Тест Уилкоксона)

Теперь повторите тест, но отдельно для каждого медикамента. Это значит, что один раз расчет нужно произвести с условием «Лекарство = 1», а второй – с условием «Лекарство = 2».

Для расчета можно применить метод «Выбрать наблюдения», однако метод «Разделить файл» является более быстрым:

- выберите в меню Data (Данные) Split file... (Разделить файл);
- активируйте опцию Organize output by groups (Выводить результаты по группам) и перенесите переменную 'Лекарство' в поле Groups based on (Группы основываются на);
- данные не сортированы по групповым признакам, поэтому оставьте опцию Sort the File by grouping variables... (Файл сортировать по групповым признакам) включенной и щелкните на ОК;
- проведите еще раз тест Уилкоксона, как было описано в начале раздела. Теперь он производится отдельно для каждого медикамента.

2.5. Подготовка отчета

Сохраните результаты работы:

- отсортируйте матрицу данных по возрастанию переменной ID, сохраните ее;
- отредактируйте файл результатов, сохранив только правильно выполненные задания, сохраните его в своей папке.

Подготовьте в MS Word отчет «врачебной комиссии», проиллюстрируйте его соответствующими таблицами и сохраните в своей папке.

Отчет должен содержать следующие разделы:

1. Цель работы.
2. Краткие теоретические сведения, описание программного продукта.
3. Созданную базу данных на основе предложенной.
4. Результаты исследований и их анализ.
5. Выводы по работе.

Задание № 3

Корреляционный и регрессионный анализ

Цель работы: закрепить навыки работы и освоить методы корреляционного и регрессионного анализа.

Ход работы

3.1. Корреляционный анализ

Метод вычисления коэффициента корреляции зависит от вида шкалы, к которой относятся переменные:

- переменные с интервальной и с номинальной шкалой: коэффициент корреляции Пирсона (корреляция моментов произведений);
- по меньшей мере, одна из двух переменных имеет порядковую шкалу либо не является нормально распределенной: ранговая корреляция по Спирману или t (тау-гро-соая) Кендала;
- одна из двух переменных является дихотомической: точная двухрядная корреляция. Эта возможность в SPSS отсутствует. Вместо этого может быть применен расчет ранговой корреляции;
- обе переменные являются дихотомическими: четырехполевая корреляция. Данный вид корреляции рассчитывается в SPSS на основании определения мер расстояния и мер сходства.

На основании данных исследования гипертонии нам нужно рассчитать коэффициент корреляции по Пирсону попарно для переменных 'Хол0', 'Хол1', 'Хол6', 'Хол12' (т. е. сформировать для этих переменных корреляционную матрицу).

Выберите в меню Analyze... (Анализ) Correlate... (Корреляция) Bivariate... (Парные). Появится диалоговое окно Bivariate Correlations (Парные корреляции) (рис. 3.1).

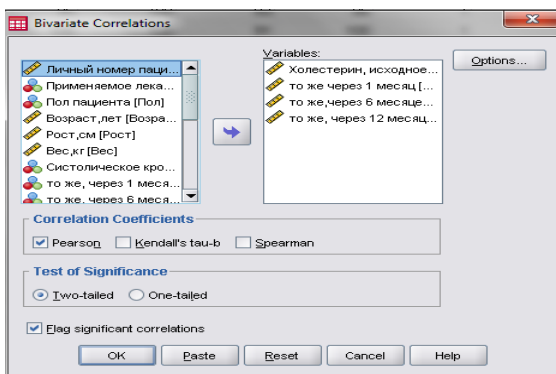


Рис. 3.1. Диалоговое окно Bivariate Correlations (Двумерные корреляции)

Переменные 'Хол0', 'Хол1', 'Холб', 'Хол12' перенесите по очереди в поле тестируемых переменных. Расчет коэффициента корреляции по Пирсону является предварительной установкой, также как двусторонняя проверка значимости и маркировка значимых корреляций. Начните расчет путем нажатия кнопки ОК.

В окне просмотра появятся следующие результаты (табл. 3.1):

Таблица 3.1

Correlations

	Холестерин, исходное зна- чение	то же, через 1 месяц	то же, через 6 месяцев	то же, через 12 месяцев
Холестерин, исходное значение	1	,861**	,775**	,802**
Pearson Cor- relation		,000	,000	,000
Sig. (2-tailed)		174	174	174
N				
то же, через 1 месяц	,861**	1	,852**	,813**
Pearson Cor- relation			,000	,000
Sig. (2-tailed)			174	174
N				
то же, через 6 месяцев	,775**	,852**	1	,892**
Pearson Cor- relation		,000		,000
Sig. (2-tailed)				174
N				
то же, через 12 месяцев	,802**	,813**	,892**	1
Pearson Cor- relation		,000	,000	
Sig. (2-tailed)				174
N				

** Correlation is significant at the 0.01 level (2-tailed). (Корреляция является значимой на уровне 0,01 (2-сторонняя))

Полученные результаты содержат: корреляционный коэффициент Пирсона r , количество использованных пар значений переменных и вероятность ошибки p , которая соответствует предположению о ненулевой корреляции. Сделайте вывод о силе корреляции.

Чтобы выявить ранговую корреляцию по Спирману и Кендалу, уберите в диалоговом окне Bivariate Correlations... (Парные корреляции) метку для расчета корреляции по Пирсону, установленную по умолчанию. Вместо этого активируйте расчет корреляции Спирмана. Этот расчет даст следующие результаты (табл. 3.2):

Таблица 3.2

Correlations (Корреляции)

			Cholesterin, Ausgangswert (Холестерин, исходная величина)	Cholesterin, nach 1 Monat (Холестерин, через 1 месяц)	Cholesterin, nach 6 Monaten (Холестерин, через 6 месяцев)	Cholesterin, nach 12 Monaten (Холестерин, через 12 месяцев)
Spearman's rho (рСпирмана)	Cholesterin, Ausgangswert (Холестерин, исходная величина)	Correlation Coefficient (Коэффициент корреляции) Sig. (2-tailed) (Значимость (2-сторонняя)) N	1 ,000 174	,877" ,000 174	,791" ,000 174	,792" ,000 174
	Cholesterin, nach 1 Monat (Холестерин, через 1 месяц)	Correlation Coefficient (Коэффициент корреляции) Sig. (2-tailed) (Значимость (2-сторонняя)) N	,877" ,000 174	1 ,000 174	,874** ,000 174	,834" ,000 174
	Cholesterin, nach 6 Monaten (Холестерин, через 6 месяцев)	Correlation Coefficient (Коэффициент корреляции) Sig. (2-tailed) (Значимость (2-сторонняя)) N	,791** ,000 174	,874** ,000 174	1 ,000 174	,879" ,000 174
	Cholesterin, nach 12 Monaten (Холестерин, через 12 месяцев)	Correlation Coefficient (Коэффициент корреляции) Sig. (2-tailed) (Значимость (2-сторонняя)) N	,792** ,000 174	,834" ,000 174	,879" ,000 174	1 ,000 174

** Correlation is significant at the .01 level (2-tailed). (Корреляция является значимой на уровне 0,01 (2-сторонняя))

Коэффициенты ранговой корреляции весьма близки к соответствующим значениям коэффициентов Пирсона (исходные переменные имеют нормальное распределение).

Еще одним вариантом ранговых коэффициентов корреляции являются коэффициенты Кендала (tb Кендала), расчет которых можно вызвать в диалоговом окне *Bivariate Correlations...* (Парные корреляции). Применение коэффициента Кендала является предпочтительным, если в исходных данных встречаются выбросы.

Рассчитайте корреляционную матрицу Кендала. Обратите внимание, что в данном случае коэффициенты значительно ниже корреляционных коэффициентов Спирмана.

3.2. Регрессионный анализ

Если расчет корреляции характеризует силу связи между двумя переменными, то регрессионный анализ служит для определения вида этой связи и дает возможность для прогнозирования значения одной (зависимой) переменной, отталкиваясь от значения другой (независимой) переменной.

Чтобы вызвать регрессионный анализ в SPSS, выберите в меню *Analyze...* (Анализ) *Regression...* (Регрессия). Откроется соответствующее подменю.

Простая линейная регрессия является простейшей и применяется чаще всех остальных видов.

Если рассматривать показатель холестерина через один месяц (переменная 'Хол1') как зависимую переменную (y), а исходную величину как независимую переменную (x), то тогда для проведения регрессионного анализа нужно будет определить параметры соотношения

$$\text{Хол1} = b * \text{Хол0} + a.$$

После определения этих параметров, зная исходный показатель холестерина, можно спрогнозировать показатель, который будет через один месяц. Для расчета уравнения регрессии выберите в меню *Analyze...* (Анализ) *Regression...* (Регрессия) *Linear...* (Линейная). Появится диалоговое окно *Linear Regression* (Линейная регрессия) (рис. 3.2).

Перенесите переменную 'Хол1' в поле для зависимых переменных и присвойте переменной 'Хол0' статус независимой переменной.

Ничего больше не меняя, начните расчет нажатием ОК.

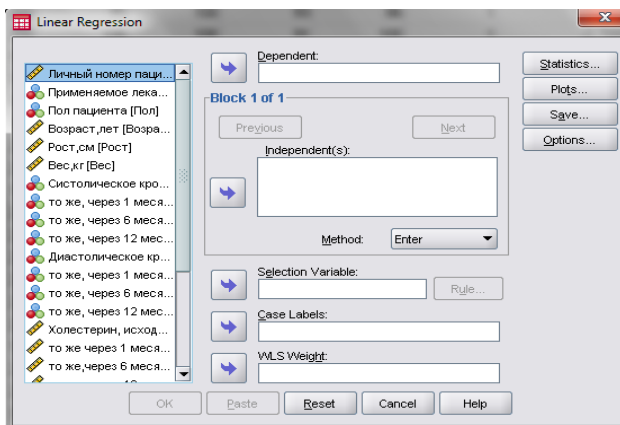


Рис. 3.2. Диалоговое окно Линейная регрессия

Вывод основных результатов выглядит следующим образом (табл. 3.3–3.5):

Таблица 3.3

Model Summary (Сводная таблица по модели)

Model (Модель)	R	R Square (R-квадрат)	Adjusted R Square (Смещенный R-квадрат)	Std. Error of the Estimate (Стандартная ошибка оценки)
1	,861a	,741	,740	25,26

a. Predictors: (Constant), Cholesterin, Ausgangswert (Влияющие переменные: (константа), холестерин, исходная величина)

Таблица 3.4

ANOVA^b

Model (Модель)		Sum of Squares (Сумма Квадратов)	df	Mean Square (Среднее значение квадрата)	F	Sig. (Значимость)
1	Regression (Регрессия)	314337,948	1	314337,9	492,722	,000a
	Residual (Остатки)	109729,408	172	637,962		
	Total (Сумма)	424067,356	173			

a. Predictors: (Constant), Cholesterin, Ausgangswert (Влияющие переменные: (константа), холестерин, исходная величина)

b. Dependent Variable: Cholesterin, nach 1 Monat (Зависимая переменная холестерин через 1 месяц)

Таблица 3.5

Coefficients (Коэффициенты) ^a

Model (Мо- дель)		Unstandardized Coefficients (Не стандартизиро- ванные коэффи- циенты)		Standardized Coefficients (Стандартизиро- ванные коэффи- циенты)		T	Sig. (Зна- чимость)
		B	Std. Error (Стан- дартная ошибка)	β (Beta)			
1	(Constant) (Констан- та)	34,5 46	9,416			3,66 9	,000
	Cholesterin , Ausgangs wert (хо- лестерин, исходная величина)	,863	,039		,861	22,1 97	,000

a. Dependent Variable (Зависимая переменная)

Проанализируйте таблицы. Здесь выводятся коэффициент регрессии b и смещение по оси координат a под именем «константа», т. е. уравнение регрессии выглядит следующим образом:

$$\text{chol1} = 0,863 - \text{chol0} + 34,546.$$

Если значение исходного показателя холестерина составляет, к примеру, 280, то через один месяц можно ожидать показатель равный 276.

Таблица также отражает два источника дисперсии: дисперсию, которая описывается уравнением регрессии (сумма квадратов, обусловленная регрессией), и дисперсию, которая не учитывается при записи уравнения (остаточная сумма квадратов). Частное от суммы квадратов, обусловленных регрессией, и остаточной суммы квадратов называется «коэффициентом детерминации». В таблице результатов это частное выводится под именем «R-квадрат». В данном примере мера определенности равна

$$314337,948 / 424067,356 = 0,741.$$

Эта величина характеризует качество регрессионной прямой, т. е. степень соответствия между регрессионной моделью и исходными данными. Мера определенности всегда лежит в диапазоне от 0 до 1.

Чтобы на диаграмме рассеяния изобразить регрессионную прямую, поступите следующим образом: выберите в меню опции Graphs ... (Графики) Scatter plots... (Диаграмма рассеяния). Откроется диалоговое окно Scatter plots... (Диаграмма рассеяния), как изображено на рисунке 3.3.

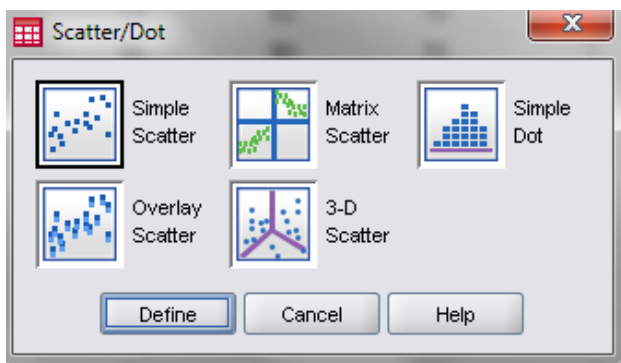


Рис. 3.3. Диалоговое окно Scatter plots... (Диаграмма рассеяния)

В диалоговом окне Scatter plots... (Диаграмма рассеяния) оставьте предварительную установку Simple (Простая) и щелкните на кнопке Define (Определить).

Откроется диалоговое окно Simple Scatterplot (Простая диаграмма рассеяния) (рис. 3.4).

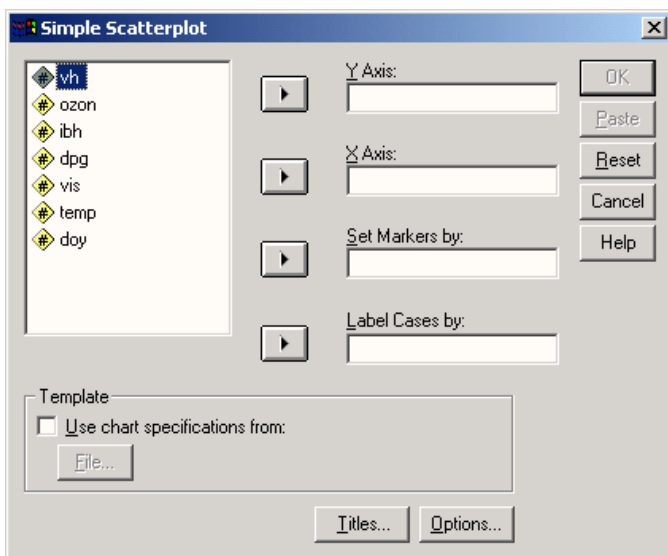


Рис. 3.4. Диалоговое окно Simple Scatterplot (Простая диаграмма рассеяния)

Перенесите переменную 'Хол1' в поле оси Y, а переменную 'Хол0' в поле оси X. Подтвердите щелчком на ОК.

В окне просмотра результатов появится диаграмма рассеяния (рис. 3.5).

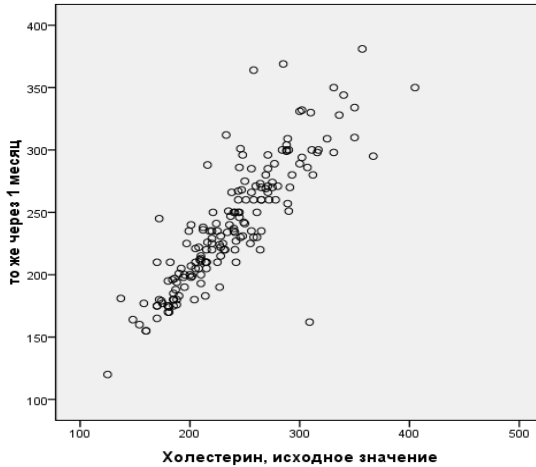


Рис. 3.5. Диаграмма рассеяния в окне просмотра

Щелкните дважды на этом графике, чтобы перенести его в редактор диаграмм. Выберите в редакторе диаграмм меню Elements... Fit Line at Total... .

Откроется диалоговое окно Properties (Возможности). Отметьте, как показано на рис. 3.6.

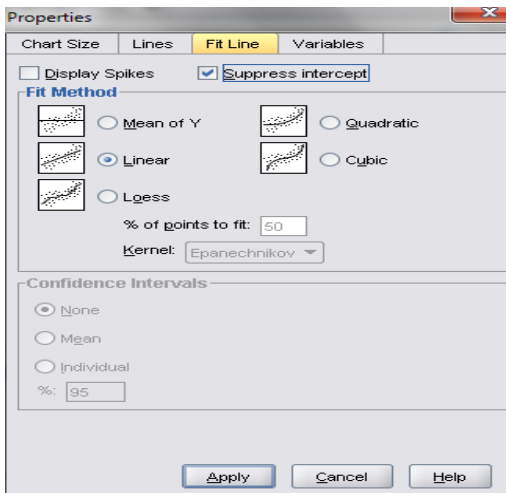


Рис. 3.6. Диалоговое окно Properties

Подтвердите предварительную установку (Apply) и затем на Close. Закройте редактор диаграмм и щелкните один раз где-нибудь вне графика.

Теперь в диаграмме рассеяния отображается регрессионная прямая (рис. 3.7).

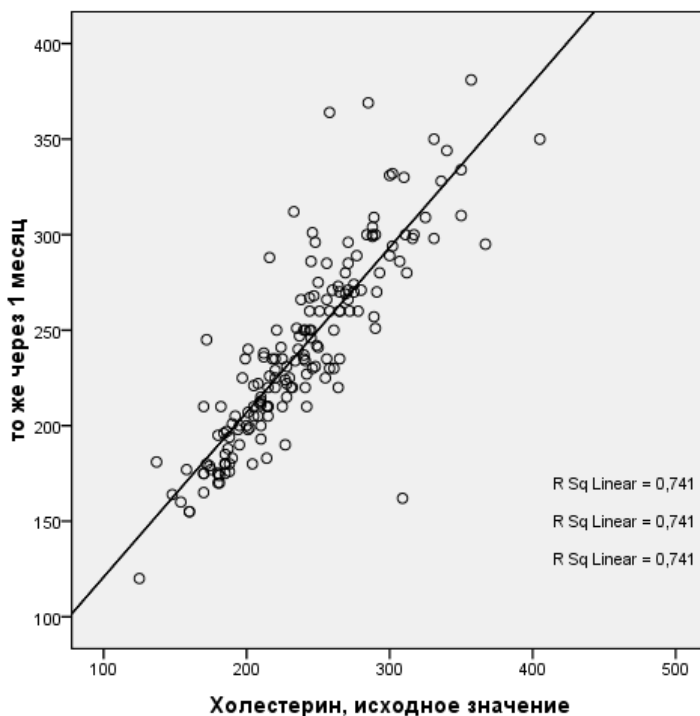


Рис. 3.7. Диаграмма рассеяния с регрессионной прямой

3.3. Подготовка отчета

Сохраните результаты работы и подготовьте в MS Word отчет, проиллюстрируйте его соответствующими таблицами и сохраните в своей папке.

Методические указания: при подготовке отчета опирайтесь на рекомендации из задания 1.3.

Контрольные вопросы

1. Каково назначение программного продукта SPSS?
2. Назовите основные возможности и отличительные черты программного продукта SPSS.
3. Дайте определения понятиям «мода», «медиана», «дисперсия».
4. Дайте определения понятиям «минимум и максимум» выборки, «размах вариационного ряда», «накопленная частота».
5. Поясните понятие «таблица сопряженности».
6. Опишите алгоритм построения диаграмм различного вида.
7. Какие предварительные оценки выполняются перед проведением любого статистического теста?
8. Назовите типы статистических шкал, приведите примеры.
9. Какие выборки являются зависимыми (независимыми)?
10. Дайте определение диаграммы ветвей и листьев.
11. Какие тестовые ситуации выделяют при сравнении средних значений выборок?
12. От чего зависит метод вычисления коэффициента корреляции?
13. Для чего служит регрессионный анализ?
14. Как интерпретируют уравнение регрессии?

Рекомендуемая литература

1. SPSS для Windows : руководство пользователя SPSS. – Кн. 1. – М. : Статистические системы и сервис, 1995.
2. SPSS BASE 8.0 : руководство пользователя SPSS. – М. : СПСС РУСЬ, 1998.
3. Handbook of Statistical Modeling for Social and Behavioral Sciences. – New York and London : Plenum press, 1995.
4. SPSS BASE 7.5. Syntax Reference Guide : руководство пользователя SPSS. – Chicago, 1997.
5. Ковалева, Г. Д. Анализ социологических данных с применением статистического пакета SPSS : уч.-метод. пособие / Г. Д. Ковалева, П. С. Ростовцев. – Новосибирск : НГУ, 2002. – 160 с.
6. Бююль, А. SPSS: искусство обработки информации. Анализ статистических данных и восстановление скрытых закономерностей / А. Бююль, П. Цефель. – СПб. : ДиаСофтЮП, 2002. – 608 с.

Приложение 1

Таблица 1

Ключ к работе

Действие	Реализация
Определение переменной (окно редактора данных, закладка Variable Viewer)	
Имя переменной	Столбец Name . Длина имени – до 8 символов. Запрещены слова: all, and, or, not, with, by, to, eq, ne, lt, le, gt, ge . Запрещены символы: пробел, !, ?, ' и *. Имя не может заканчиваться точкой. Каждое имя должно быть уникальным. Имена нечувствительны к регистру.
Тип переменной	Столбец Type . Основные типы: Numeric (числовой), String (строковый).
Формат переменной	Столбцы: Width (общее число позиций), Decimals (число позиций после десятичной точки). Рекомендуется начинать установку формата с Decimals .
Метка переменной	Столбец Label (любой поясняющий текст). Может быть на русском языке.
Метки значений переменной	Столбец Values (могут задаваться для каждого значения, особенно для номинальных и порядковых переменных). Могут быть на русском языке.
Пропущенные значения (пользовательские)	Столбец Missing . До 3-х дискретных значений либо интервал.
Пропущенные значения (системные)	Пустая клетка в окне редактора (закладка Data Viewer) или перекодировка (см. «Перекодировка»).
Уровень измерения	Столбец Measure: Scale (количественный), Ordinal (порядковый), Nominal (номинальный).
Ввод данных	
Ввод данных	Окно редактора данных (закладка Data View). Переход между клетками – клавишей Enter или клавишами управления курсором.
Проверка качества ввода и чистка данных	
Построение одномерных распределений	Analyze descriptive statistics frequencies... имена нужных переменных в окно Variable(s) OK .

Действие	Реализация
Построение таблиц сопряженности	Analyze descriptive statistics crossstabs... имя переменной, образующей строки таблицы, в окне Row(s) имя переменной, образующей столбцы, в окне Column(s) OK .
Чистка данных	Окно редактора данных (закладка Data View) или перекодировка (см. «Перекодировка»).
<i>Преобразование матрицы данных</i>	
Добавление переменной	Окно редактора данных (закладка Variable Viewer): выделить строку с переменной, перед которой должна быть вставка вызвать контекстное меню (правой кнопкой мыши) Insert Variables . Окно редактора данных (закладка Data Viewer): выделить столбец с переменной, перед которой должна быть вставка вызвать контекстное меню (правой кнопкой мыши) Insert Variables .
Удаление переменной	Окно редактора данных (закладка Variable Viewer): выделить строку с удаляемой переменной вызвать контекстное меню (правой кнопкой мыши) Clear . Окно редактора данных (закладка Data Viewer): выделить столбец с удаляемой переменной вызвать контекстное меню (правой кнопкой мыши) Cut или Clear .
Добавление случая	Окно редактора данных (закладка Data Viewer): выделить строку, перед которой должна быть вставка вызвать контекстное меню (правой кнопкой мыши) Insert Cases .
Удаление случая	Окно редактора данных (закладка Data Viewer): выделить удаляемую строку вызвать контекстное меню (правой кнопкой мыши) Cut или Clear .
Сортировка случаев	Окно редактора данных (закладка Data Viewer): выделить столбец с переменной, по которой производится сортировка вызвать контекстное меню (правой кнопкой мыши) Sort ascending (по возрастанию) или Sort descending (по убыванию). Окно редактора данных: Data Sort Cases... перенести в окно Sort by имена переменных, по которым осуществляется сортировка Sort Order : выбрать Ascending или Descending OK .

Действие	Реализация
Отбор случаев по заданным критериям с помощью фильтров	Окно редактора данных: Data Select Cases... в разделе Select выбрать If condition is satisfied кнопка If... в открывшемся окне задать критерии отбора, используя логические операторы Continue Unselected Cases Are выбрать Filtered OK . Для снятия фильтра в разделе Select выбрать All cases .
<i>Перекодировка, группировка, вычисление переменных</i>	
Перекодировка дискретных значений	Окно редактора данных: Data Transform Recode Into Different Variable(s)... перенести имя перекодируемой переменной в окно Input Variable → Output Variable ввести имя новой переменной в открывшееся окно Output Variable Name кнопка Old and New Values в разделе Old Value выбрать Value ввести старое значение в разделе New Value выбрать Value ввести новое значение кнопка Add (правило перекодировки появится в окне) задать последовательно все правила перекодировки, проверить их наличие в окне Continue Change OK .
Группировка (построение интервалов)	Окно редактора данных: Data Transform Recode Into Different Variable(s)... перенести имя перекодируемой переменной в окно Input Variable → Output Variable ввести имя новой переменной в открывшееся окно Output Variable Name кнопка Old and New Values в разделе Old Value выбрать один из трех вариантов интервалов Range (с двумя границами, с открытой верхней границей или с открытой нижней границей) задать границы интервалов в разделе New Value выбрать Value ввести новое значение кнопка Add (правило перекодировки появится в окне) задать последовательно все правила перекодировки, проверить их наличие в окне Continue Change OK .
Перцентильная группировка	Analyze descriptive statistics frequencies... имена группируемой переменной переместить в окно Variable(s) кнопка Statistics... в разделе Percentile Values выбрать один из трех вариантов группировки Continue OK полученные границы использовать для построения интервалов.

Действие	Реализация
Вычисление переменных с использованием арифметических операций и стандартных математических функций	Окно редактора данных: Data Transform Compute... ввести имя вычисляемой переменной в окно Target Variable в окне Numeric Expression задать формулу для вычисления с использованием арифметических выражений и стандартных математических функций OK .
Вычисление переменных в соответствии с некоторыми условиями	Окно редактора данных: Data Transform Compute... ввести имя вычисляемой переменной в окно Target Variable в окне Numeric Expression задать число или формулу для вычисления кнопка If... выбрать Include if case satisfies condition в окне задать условие Continue OK На вопрос Change existing variable? ответить OK . Повторять столько раз, сколько необходимо задать условий.
Вычисление переменных, подсчитывающих ответы с одинаковыми оценками	Окно редактора данных: Data Transform Count... в окне Target Variable задать имя новой переменной в окне Numeric Variables поместить имена переменных, содержащих оценки кнопка Define Values в разделе Values выбрать Value ввести значение подсчитываемых оценок кнопка Add Continue OK .
Стандартизация переменной	Окно редактора данных: Data Transform Rank Cases... поместить имена переменных в окно Variable(s) кнопка Rank Types... флажок Normal Scores... кнопка More >> Continue OK В результате в файле появится новая переменная <i>имя</i> , значения которой являются z-оценки значенной переменной <i>имя</i> .
Графики одномерных распределений	
Круговая диаграмма (только для номинальных шкал!)	Graphs Pie... выбрать Summaries for group of cases Define выбрать N of cases или % of cases поместить имя переменной в окно Define Slices by OK .
Диаграмма столбцов (для номинальных, порядковых, дискретных количественных шкал)	Graphs Bar... Simple Summaries for group of cases Define выбрать N of cases или % of cases для частот; Cum. n of cases или Cum. % of cases для накопленных частот поместить имя переменной в окно Category Axes OK .

Действие	Реализация
Кластеризованная диаграмма столбцов (для сравнительного анализа распределений переменных в двух или нескольких группах)	Graphs Bar... Clustered Summaries for group of cases Define выбрать N of cases или % of cases для частот; Cum. n of cases или Cum. % of cases для накопленных частот поместить имя группирующей переменной в окно Define clusters by поместить имя переменной в окно Category Axes OK .
Ленточная диаграмма (для исследований структуры группы объектов)	Graphs Bar... Stacked Summaries for group of cases Define выбрать N of cases или % of cases для частот; Cum. n of cases или Cum. % of cases для накопленных частот поместить имя переменной в окно Category Axes поместить имя структурирующей переменной в окно Define stacks by OK .
Редактирование графиков	При необходимости график можно отредактировать – изменить цвета, названия осей и т. п. Для редактирования графика необходимо открыть окно SPSS Chart Editor , щелкнув дважды мышью по графику в окне результатов.
<i>Чтение, сохранение, обмен файлами и результатами вычислений</i>	
Чтение файлов данных SPSS и других форматов	File Open Data... в открывшемся окне перейти в нужную папку выбрать нужный тип файла (*.sav, *.dbf, *.xls и т. п.) выбрать нужный файл кнопка открыть .
Сохранение файлов данных SPSS и других форматов	Окно редактора данных: File Save As... в открывшемся окне перейти в нужную папку выбрать нужный тип файла (*.sav, *.dbf, *.xls, *.dat и т. п.) ввести имя файла кнопка сохранить .
Использование результатов, полученных в SPSS, другими приложениями Windows (Word, Excel, PowerPoint)	Выделите в файле результатов нужную таблицу или рисунок откройте контекстное меню (правая кнопка мыши) Copy objects перейдите в нужное приложение поместите курсор в место, куда следует вставить объект вызовите контекстное меню вставить .

Таблица 2

Расширенная база данных биохимического анализа крови пациентов
поликлиники N.

Личный номер	Имя больного	Пол	Содержание общего белка в крови, г/л	Содержание глюкозы в крови, ммоль/л	Содержание холестерина в крови, ммоль/л
1	2	3	4	5	6
1	Людмила А.	ж	22,5	3,6	2,7
2	Галина А.	ж	28,6	4,0	6,0
3	Владимир А.	м	31,6	4,1	3,4
4	Геннадий Б.	м	68,3	4,5	5,4
5	Сергей Б.	м	95,2	3,9	5,8
6	Мария Б.	ж	88,9	4,4	6,2
7	Наталья Б.	ж	54,6	4,5	4,5
8	Нина В.	ж	89,3	4,9	4,8
9	Елена Г.	ж	37,8	4,3	5,7
10	Александр Д.	м	43,9	4,8	3,8
11	Юрий Е.	м	46,9	4,9	8,4
12	Виктор Ж.	м	83,6	5,3	4,8
13	Антон З.	м	81,5	5,2	7,6
14	Ольга К.	ж	104,2	5,3	8,1
15	Григорий К.	м	69,9	5,7	8,7
16	Ольга В.	ж	90,0	5,1	6,3
17	Валентина Л.	ж	53,1	5,6	6,7
18	Людмила П.	ж	59,2	5,7	8,0
19	Андрей Р.	м	62,2	6,1	5,3
20	Виктор Р.	м	98,9	4,8	3,0
21	Татьяна Ф.	ж	70,6	3,7	6,7
22	Валентин Ш.	м	93,3	4,2	3,8
23	Виктория Я.	ж	59,0	4,3	6,0
24	Сергей Я.	м	79,1	4,7	6,5
25	Светлана Я.	ж	42,2	4,6	6,9
26	Людмила Л.	ж	48,3	4,7	5,0
27	Галина Б.	ж	51,3	5,1	5,4
28	Владимир Б.	м	88,0	4,5	6,4
29	Геннадий В.	м	59,7	5,0	4,2
30	Сергей В.	м	82,4	5,1	2,4
31	Мария В.	ж	48,1	5,5	5,4

1	2	3	4	5	6
32	Наталья В.	ж	68,2	4,2	3,0
33	Нина Г.	ж	31,3	3,1	4,8
34	Елена Д.	ж	37,4	3,6	5,2
35	Александр Е.	м	77,1	4,8	5,6
36	Юрий Ж.	м	42,8	5,0	4,0
37	Виктор З.	м	62,9	5,3	4,3
38	Антон И.	м	26,0	5,6	5,1
39	Ольга Л.	ж	32,1	5,2	3,4
40	Григорий Л.	м	71,8	5,5	1,9
41	Ольга В.	ж	37,5	5,8	4,3
42	Валентина М.	ж	57,6	6,1	2,4
43	Людмила Н.	ж	20,7	5,6	3,9
44	Андрей С.	м	26,8	6,0	4,2
45	Виктор Т.	м	66,5	6,3	4,4
46	Татьяна Х.	ж	32,2	6,6	3,2
47	Валентин Ц.	м	52,3	6,3	4,6
48	Виктория Ю.	ж	36,2	6,6	4,9
49	Сергей Ю.	м	21,5	6,9	5,7
50	Светлана Ю.	ж	61,2	6,4	4,0
51	Людмила В.	ж	57,7	6,8	2,5
52	Галина В.	ж	81,8	7,1	4,9
53	Владимир В.	м	37,6	7,4	3,0
54	Геннадий Г.	м	44,9	6,1	4,5
55	Сергей Г.	м	92,5	5,0	4,8
56	Мария Г.	ж	51,4	4,8	5,0
57	Наталья Г.	ж	75,5	5,0	3,8
58	Нина Д.	ж	31,2	5,3	5,2
59	Елена Е.	ж	38,5	5,6	5,5
60	Александр Ж.	м	86,2	5,2	6,3
61	Юрий З.	м	45,0	5,5	4,6
62	Виктор Л.	м	69,1	5,8	3,1
63	Антон Л.	м	24,8	6,1	5,5
64	Ольга М.	ж	32,2	5,6	3,6
65	Григорий А.	м	79,8	6,0	5,1
66	Ольга Н.	ж	38,6	6,3	5,4
67	Валентина О.	ж	62,8	6,6	5,6

1	2	3	4	5	6
68	Людмила П.	ж	43,4	6,3	4,4
69	Андрей У.	м	70,9	6,6	5,8
70	Виктор С.	м	104,2	6,9	6,1
71	Татьяна Ц.	ж	43,1	6,4	6,9
72	Валентин Э.	м	53,2	6,8	2,9
73	Виктория А.	ж	118,9	7,1	3,2
74	Сергей А.	м	62,1	7,4	3,5
75	Светлана В.	ж	95,4	6,1	2,2
76	Людмила И.	ж	34,3	5,0	3,6
77	Галина Г.	ж	44,4	5,4	3,9
78	Владимир А.	м	110,1	5,7	4,7
79	Геннадий С.	м	53,3	6,0	3,0
80	Сергей Ч.	м	63,9	5,7	1,5
81	Мария К.	ж	59,9	6,0	3,9
82	Наталья М.	ж	97,8	6,3	2,0
83	Нина Р.	ж	73,4	5,8	3,5
84	Елена Ф.	ж	59,4	6,2	3,8
85	Александр З.	м	73,4	6,5	4,1
86	Юрий У.	м	69,0	6,8	2,8
87	Виктор И.	м	85,7	5,5	4,2
88	Антон К.	м	65,3	4,4	4,5
89	Ольга Д.	ж	47,3	4,8	5,3
90	Григорий С.	м	61,2	5,9	3,6
91	Ольга З.	ж	85,1	6,2	2,1
92	Валентина Н.	ж	73,6	6,4	4,5
93	Людмила П.	ж	73,4	6,7	2,6
94	Андрей Т.	м	82,7	6,4	4,1
95	Виктор У.	м	69,2	6,7	4,4
96	Татьяна Ж.	ж	63,0	7,0	4,7
97	Валентин Ю.	м	82,0	7,3	3,4
98	Виктория Б.	ж	81,4	7,0	4,8
99	Сергей Б.	м	67,0	7,3	5,1
100	Светлана Т.	ж	78,3	7,6	5,9

Приложение 2

Антропометрические и биохимические показатели группы пациентов поликлиники N.

ID	Лекарство	Пол	Возраст	Рост	Вес	Сис0	Сис1	Сис6	Сис12	Диас0	Диас1	Диас6	Диас12	Хол0	Хол1	Хол6	Хол12	Сах0	Сах1	Сах6	Сах12
1	2	2	66	166	70	170	135	155	160	100	90	75	75	265	235	245	205	90	85	84	90
2	2	2	58	163	63	160	155	150	155	95	95	85	85	185	185	175	195	85	80	87	92
3	2	1	47	173	70	165	155	145	145	100	100	90	90	225	235	250	260	95	90	98	100
4	1	1	80	150	48	170	165	145	145	95	90	80	80	160	155	135	136	86	80	76	77
5	2	2	52	165	77	170	165	155	150	105	100	75	85	265	260	245	250	104	100	92	98
6	1	1	48	178	82	160	155	160	150	100	95	95	85	195	200	185	210	95	90	91	93
7	1	2	54	161	77	170	160	155	150	100	100	95	80	230	225	235	200	95	100	93	86
8	2	1	46	178	75	160	150	160	155	100	85	95	85	185	180	195	155	90	95	87	88
9	1	1	53	174	83	160	155	155	155	100	100	90	90	220	220	235	200	105	100	91	100
10	1	1	71	158	62	165	155	150	150	100	95	95	80	180	175	175	195	105	100	93	95
11	2	2	55	172	88	170	170	150	155	105	110	100	90	265	260	250	230	145	185	145	125
12	2	1	63	171	73	160	150	150	155	95	90	90	85	185	180	185	170	90	85	78	82
13	2	1	52	181	88	155	150	150	145	100	100	90	85	170	175	175	185	100	95	91	91
14	2	2	66	160	66	165	160	150	150	105	100	90	90	195	190	175	185	120	125	100	105
15	1	1	52	176	78	160	150	150	140	100	85	80	80	185	180	195	170	87	82	78	80
16	2	2	51	150	67	170	170	155	150	105	100	90	85	245	250	245	250	105	110	100	101
17	2	1	57	171	70	160	160	140	145	100	95	85	90	180	170	185	160	75	87	78	75
18	2	2	73	158	55	170	165	160	150	100	95	90	85	205	210	200	200	95	96	97	80
19	1	1	47	173	70	170	160	145	150	105	100	85	85	215	220	200	195	81	88	72	70
20	2	2	43	155	70	170	170	145	145	110	105	90	85	220	225	235	240	96	97	100	105
21	1	1	42	176	72	160	150	150	145	105	95	90	90	185	175	180	175	76	73	82	78
22	2	2	48	163	75	160	155	150	140	100	95	90	85	180	175	180	170	78	86	90	80

ID	Лекарство	Пол	Возраст	Рост	Вес	Сис0	Сис1	Сис6	Сис12	Диас0	Диас1	Диас6	Диас12	Хол0	Хол1	Хол6	Хол12	Сах0	Сах1	Сах6	Сах12
23	2	2	37	172	70	160	160	155	160	100	100	95	90	170	175	180	170	74	76	87	87
24	1	1	51	176	76	155	150	135	145	100	95	90	85	210	215	220	220	93	96	90	78
25	2	2	46	163	72	155	150	135	135	105	100	85	85	215	210	175	185	73	77	76	73
26	1	1	53	171	76	150	150	135	135	105	95	95	85	220	235	240	245	86	78	92	93
27	2	2	51	156	61	160	155	150	130	100	95	85	80	215	210	190	205	86	72	78	86
28	1	1	43	181	83	165	160	140	145	95	90	85	85	160	155	180	190	76	71	76	83
29	2	2	49	161	129	140	150	165	130	105	110	100	95	205	205	190	180	84	75	89	73
30	2	2	76	157	83	150	125	150	140	70	80	80	90	215	205	205	215	100	100	99	100
31	2	2	43	160	83	160	170	140	140	100	100	90	85	225	210	195	200	87	75	77	92
32	2	2	63	152	69	160	140	150	150	80	80	95	80	170	165	175	175	75	86	71	70
33	1	2	78	156	48	160	150	175	130	105	90	100	80	210	200	205	280	250	89	186	180
34	2	1	51	179	124	210	200	190	190	120	105	105	100	261	250	241	234	127	120	119	116
35	1	2	62	154	79	190	150	130	135	100	85	70	75	256	266	239	224	98	103	104	106
36	2	2	71	152	85	180	180	170	180	105	100	100	95	210	211	210	201	102	121	110	99
37	1	2	74	145	84	190	150	150	145	105	90	90	95	271	271	270	260	123	109	100	99
38	2	1	48	167	80	160	150	155	145	100	80	90	80	271	266	250	200	99	101	100	98
39	2	1	81	177	77	185	140	150	150	90	80	80	80	210	213	201	194	121	124	108	122
40	2	2	70	174	82	180	140	160	140	100	75	85	80	201	198	194	196	99	109	100	101
41	1	2	53	162	94	180	140	140	145	130	80	80	85	302	294	270	261	135	124	120	119
42	2	1	64	164	85	165	140	150	140	100	80	95	80	208	222	211	199	124	121	115	102
43	1	2	70	160	68	180	170	170	155	100	90	95	85	250	275	269	268	104	113	100	139
44	1	2	77	166	80	160	150	150	140	100	90	80	90	200	200	261	268	110	99	109	102
45	2	2	39	167	80	180	150	150	150	120	80	90	90	190	201	264	240	99	101	109	105
46	1	2	63	174	86	180	140	160	160	100	100	85	95	194	198	186	190	110	104	102	104
47	1	1	58	165	72	190	160	150	145	110	90	95	80	275	274	248	248	140	136	118	110

ID	Лекарство	Пол	Возраст	Рост	Вес	Сис0	Сис1	Сис6	Сис12	Диас0	Диас1	Диас6	Диас12	Хол0	Хол1	Хол6	Хол12	Сах0	Сах1	Сах6	Сах12
48	1	2	80	164	74	195	180	160	160	90	80	85	80	209	212	263	240	97	102	127	96
49	2	1	42	178	74	180	160	140	135	100	100	90	80	184	196	190	180	114	114	117	120
50	1	2	63	165	62	175	145	150	150	95	70	70	70	201	240	332	300	101	100	98	100
51	1	1	58	174	82	170	160	130	150	110	105	75	90	188	194	212	257	100	101	99	98
52	2	2	49	157	100	200	200	190	170	120	140	110	95	236	240	232	235	169	172	180	154
53	2	2	86	162	67	170	145	150	150	120	90	95	90	290	300	290	287	128	154	130	121
54	1	2	50	160	75	180	180	160	180	110	120	110	110	246	230	221	211	112	110	108	109
55	2	2	70	153	97	200	190	180	160	100	100	90	90	256	235	225	237	149	186	211	184
56	1	2	60	159	79	150	150	160	150	110	110	100	95	275	270	275	314	115	119	123	114
57	2	2	74	148	76	180	180	160	160	110	110	100	100	214	210	189	180	177	142	223	210
58	1	1	53	172	77	180	145	145	140	110	110	105	100	208	205	190	194	109	104	103	110
59	2	2	53	155	96	190	160	160	150	115	100	105	95	285	369	325	294	119	128	126	121
60	1	2	65	146	73	200	190	165	170	95	100	105	85	350	334	363	395	97	98	101	100
61	2	2	66	157	68	180	190	145	150	100	90	70	80	258	230	202	226	132	170	186	136
62	2	1	82	162	75	160	150	150	150	100	95	90	90	220	229	233	228	89	89	86	100
63	2	2	64	155	55	175	170	170	165	100	90	90	95	258	364	270	228	106	102	94	89
64	2	1	37	185	104	150	140	145	140	110	90	95	90	289	257	212	236	141	115	117	128
65	1	2	64	173	93	205	165	160	150	105	95	95	90	224	241	246	243	112	104	98	106
66	2	1	51	180	110	195	195	180	200	130	125	110	120	405	350	281	347	289	301	321	394
67	1	2	80	159	80	175	140	145	150	110	100	100	100	241	234	240	251	105	108	114	124
68	1	2	36	164	68	170	125	140	130	100	90	100	95	269	269	291	276	111	116	113	124

ID	Лекарство	Пол	Возраст	Рост	Вес	Сис0	Сис1	Сис6	Сис12	Диас0	Диас1	Диас6	Диас12	Хол0	Хол1	Хол6	Хол12	Сах0	Сах1	Сах6	Сах12
69	1	2	69	160	75	180	180	150	170	100	95	90	90	204	180	208	172	352	258	310	236
70	2	1	81	169	72	180	160	145	130	110	100	80	80	264	273	289	302	68	77	79	103
71	2	1	47	184	74	170	140	130	120	100	90	90	70	231	220	210	183	78	83	88	77
72	1	1	59	168	73	170	170	170	140	120	105	95	95	190	183	184	201	122	102	102	78
73	1	2	54	171	76	180	160	130	140	100	80	75	85	310	330	270	294	98	83	103	90
74	1	1	63	179	79	160	160	140	125	120	105	80	70	284	300	304	310	80	80	81	84
75	2	1	68	170	70	180	160	140	145	80	80	85	85	232	220	198	200	94	99	103	100
76	2	2	63	163	87	170	150	140	140	115	100	100	80	316	298	278	241	69	74	86	94
77	1	2	65	168	85	190	180	180	150	110	90	90	80	290	251	258	305	88	94	78	88
78	1	2	58	167	85	150	150	120	120	100	105	80	80	272	260	249	250	107	140	178	160
79	2	2	46	160	69	150	140	130	110	110	100	100	95	188	180	174	176	103	100	99	101
80	2	2	61	159	68	160	150	130	140	100	90	75	75	312	280	263	284	85	87	94	91
81	1	2	56	159	70	160	130	130	130	110	80	80	85	245	246	247	228	97	94	90	100
82	1	1	52	176	80	150	150	150	150	105	105	100	80	240	250	253	210	71	75	88	70
83	1	2	56	159	70	180	140	130	135	95	90	90	85	248	231	270	257	93	84	87	104
84	1	2	54	165	92	210	140	160	140	110	90	90	90	242	210	185	210	97	98	105	106
85	2	1	69	175	82	160	140	155	130	100	85	90	90	251	260	271	278	81	82	84	85
86	1	1	82	178	73	155	145	150	145	100	80	85	80	265	270	296	290	100	98	94	96
87	2	2	63	158	78	180	170	165	165	95	90	90	80	240	235	240	214	90	85	80	111

Содержание

Введение	3
Задание № 1. Создание базы данных и основы работы с медико-биологическими данными.....	4
Задание № 2. Статистический анализ и обработка медико-биологических данных.....	10
Задание № 3. Корреляционный и регрессионный анализ.....	28
Контрольные вопросы.....	37
Рекомендуемая литература	38
Приложение 1.....	39
Приложение 2.....	47

Учебное издание

Живицкая Елена Петровна

**СТАТИСТИЧЕСКИЙ АНАЛИЗ И ОБРАБОТКА ДАННЫХ
С ПРИМЕНЕНИЕМ SPSS**

УЧЕБНО-МЕТОДИЧЕСКОЕ ПОСОБИЕ

Редактор *С. М. Курбыко, О. А. Кучинский*
Корректор *С. М. Курбыко*
Компьютерная верстка *С. М. Курбыко*

Подписано в печать 21.07.2010. Формат 60×90¹/₁₆.
Бумага офсетная. Гарнитура Times. Ризография.
Усл. печ. л. 3,25. Уч.-изд. л. 1,54.
Тираж 55 экз. Заказ № 132.

Издатель и полиграфическое исполнение
учреждение образования «Международный государственный
экологический университет имени А. Д. Сахарова»

ЛИ № 02330/0131580 от 28.07.2005 г.
Республика Беларусь, 220070, г. Минск, ул. Долгобродская, 23

E-mail: info@iseu.by
<http://www.iseu.by>