

Ю. М. ЖУЧЕНКО

**ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ
В БИОЛОГИИ И ХИМИИ**

ЛАБОРАТОРНЫЙ ПРАКТИКУМ

**Гомель
УО «ГГУ им. Ф. Скорины»
2010**

**Министерство образования Республики Беларусь
Учреждение образования
«Гомельский государственный университет
имени Франциска Скорины»**

Ю. М. ЖУЧЕНКО

**ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ
В БИОЛОГИИ И ХИМИИ**

ЛАБОРАТОРНЫЙ ПРАКТИКУМ

Рекомендовано УМО вузов Республики Беларусь по образованию в области информатики и радиоэлектроники в качестве учебно-методического пособия для студентов учреждений, обеспечивающих получение высшего образования по специальности 1-31 01 01 “Биология”

**Гомель
УО «ГГУ им. Ф. Скорины»
2010**

УДК 57; 004.9+54; 004.9 (075.8)

ББК 28с51+24с51я73

Ж 834

Рецензенты:

С. В. Буга профессор, доктор биол. наук,
заведующий кафедрой зоологии Белорусского государственного
университета;

А. К. Баев доктор хим. наук, профессор кафедры биохимии и
биофизики учреждения образования «Международный
государственный экологический университет им. А. Д. Сахарова

Рекомендовано к изданию научно-методическим советом
учреждения образования «Гомельский государственный
университет имени Франциска Скорины»

Жученко, Ю. М.

Ж 834 Информационные технологии в биологии и химии:
лабораторный практикум для студентов вузов по специальности 1-
31 01 01 «Биология» / Ю. М. Жученко; М-во образования РБ,
Гомельский гос. ун-т им. Ф. Скорины. – Гомель: ГГУ
им. Ф. Скорины, 2010.– 148 с.

ISBN

Целью лабораторного практикума является оказание помощи
студентам в усвоении основ курса информационных технологий в
биологии и химии, в частности математической статистики,
обработки результатов экспериментов с применением
возможностей персональных компьютеров.

Лабораторный практикум адресован студентам специальности 1-
31 01 01 «Биология».

УДК 57; 004.9+54; 004.9 (075.8)

ББК 28с51+24с51я73

ISBN

© Жученко Ю. М., 2010

© УО «Гомельский государственный
университет им Ф. Скорины», 2010

СОДЕРЖАНИЕ

Введение	6
Лабораторная работа 1	
Дискретные распределения	7
Краткие теоретические сведения	7
Задания для выполнения	16
Лабораторная работа 2	
Основные модели теоретических распределений (Statistica 6)	18
Краткие теоретические сведения	18
Задания для выполнения	25
Лабораторная работа 3	
Распределение параметров выборки	26
Краткие теоретические сведения	26
Задания для выполнения	30
Лабораторная работа 4	
Разнообразие значений признака (MS Excel).....	32
Краткие теоретические сведения	32
Задания для выполнения	42
Лабораторная работа 5	
Разнообразие значений признака (Statistica 6).....	43
Краткие теоретические сведения	43
Задания для выполнения	50
Лабораторная работа 6	
Парная корреляция (MS Excel)	52
Краткие теоретические сведения	52
Задания для выполнения	59
Лабораторная работа 7	
Парная корреляция (Statistica 6)	60
Краткие теоретические сведения	60
Задания для выполнения	64
Лабораторная работа 8	
Частная и множественная линейные корреляции и регрессия (Statistica 6)	65
Краткие теоретические сведения	65
Задания для выполнения	75

Лабораторная работа 9	
Криволинейная корреляция и регрессия (MS Excel).....	76
Краткие теоретические сведения	76
Задания для выполнения	80
Лабораторная работа 10	
Криволинейная корреляция и регрессия (Statistica 6)	Ошибка! Закладка не
Краткие теоретические сведения ...	Ошибка! Закладка не определена.
Задания для выполнения	Ошибка! Закладка не определена.
Лабораторная работа 11	
Однофакторный дисперсионный анализ (однофакторный комплекс в MS Excel).....	Ошибка! Закладка не определена.
Краткие теоретические сведения ...	Ошибка! Закладка не определена.
Задания для выполнения	Ошибка! Закладка не определена.
Лабораторная работа 12	
Однофакторный дисперсионный анализ (однофакторный комплекс в Statistica 6).....	Ошибка! Закладка не определена.
Краткие теоретические сведения ...	Ошибка! Закладка не определена.
Задания для выполнения	Ошибка! Закладка не определена.
Лабораторная работа 13	
Многофакторный дисперсионный анализ (двухфакторный комплекс в MS Excel).....	Ошибка! Закладка не определена.
Краткие теоретические сведения ...	Ошибка! Закладка не определена.
Задания для выполнения	Ошибка! Закладка не определена.
Лабораторная работа 14	
Многофакторный дисперсионный анализ (многофакторный комплекс в Statistica 6).....	Ошибка! Закладка не определена.
Краткие теоретические сведения ...	Ошибка! Закладка не определена.
Задания для выполнения	Ошибка! Закладка не определена.
Лабораторная работа 15	
Классификация (дискриминантный анализ в Statistica 6)	Ошибка! Закладка не определена.
Краткие теоретические сведения ...	Ошибка! Закладка не определена.
Задания для выполнения	Ошибка! Закладка не определена.
Лабораторная работа 16	
Классификация (кластерный анализ в Statistica 6)	Ошибка! Закладка не определена.
Краткие теоретические сведения ...	Ошибка! Закладка не определена.
Задания для выполнения	Ошибка! Закладка не определена.
Литература	81

Приложение А (обязательное).....	82
Приложение Б (обязательное)	83
Приложение В (обязательное).....	84
Приложение Г (обязательное)	85
Приложение Д (обязательное).....	86
Приложение Е (обязательное).....	88
Приложение Ж (обязательное).....	90
Приложение И (обязательное).....	91

Введение

Некоторые биологические вопросы не могут быть решены без применения специальных математических методов. К таким вопросам относятся: сравнение выборочных групп по изучаемым показателям и определение достоверности результатов такого сравнения с заданной вероятностью безошибочных прогнозов, определение достаточной численности подопытных объектов, измерение силы влияния различных факторов на биологические процессы и явления и т. д. Современный этап развития науки характеризуется широким применением средств вычислительной техники. Применительно к современным персональным компьютерам, в арсенале математической обработки информации пользователем существует огромное количество программных продуктов. Среди них важное место занимают табличный процессор **MS Excel** и программа **Statistica 6.0** для статистического анализа данных в среде **WINDOWS**.

Целью лабораторного практикума является ознакомление с указанными программами, в которых материал расположен таким образом, что пользователь может повторить все описанные действия на своем компьютере. Самостоятельная работа в форме упражнений и задач позволит углубить понимание разбираемых примеров. Ранее каждый шаг исследований, начиная от представления данных, перевода их в нужный формат, проверки, группировки, сортировки, графической интерпретации, подготовки программ обработки до просмотра результатов, был трудной задачей. Теперь достаточно двух-трех щелчков мыши, чтобы огромные объемы данных чрезвычайно быстро преобразовались, обработались и появились на экране в виде графиков, диаграмм и таблиц.

На простых и доступных примерах, взятых из различных сфер жизни, показаны возможности этих систем по статистической обработке данных.

В основу лабораторного практикума положены материалы, опубликованные в работах (1–14) и адаптированы к современным версиям табличного процессора **MS Excel** и пакета статистического анализа **Statistica 6.0**.

Лабораторный практикум может быть адресован не только студентам специальности 1-31 01 01 «Биология», но и студентам других специальностей, работающих на персональных компьютерах.

Лабораторная работа 1

Дискретные распределения

Цель работы: научиться вычислять отдельные биномиальные вероятности и суммировать их далее по определенному множеству целых чисел на вероятностном калькуляторе в **Statistica 6**.

Краткие теоретические сведения

Классический подход к определению вероятностей событий лежит в основе теории анализа случайных событий и теоретических (модельных) распределений исходов испытаний. В свою очередь, теория математического анализа случайных событий и модели распределений исходов испытаний являются базой статистических методов, в частности, базой статистических заключений.

Дискретно варьирующие признаки, распределяются так, что вероятные численности их появления могут быть найдены по формуле бинома Ньютона:

$$N(p+q)^n = N \left(p^n + np^{n-1}q + \frac{n(n-1)p^{n-2}q^2}{1 \cdot 2} + \frac{n(n-1)(n-2)p^{n-3}q^3}{1 \cdot 2 \cdot 3} + \dots + q^n \right),$$

где n – число независимых исходов в одном испытании; p – вероятность благоприятного исхода одного случая; q – вероятность неблагоприятного исхода; N – общее число испытаний (исходов).

Когда вероятности альтернатив неравны, т. е. $p \neq q$, биномиальное распределение асимметрично. При очень малой вероятности ожидаемого события, исчисляемой сотыми или тысячными долями единицы, по сравнению с вероятностью q противоположного события распределение вероятности или частоты таких событий описывается формулой Пуассона:

$$y = \frac{\lambda^m}{m!} \cdot e^{-\lambda},$$

В этом выражении m – частота ожидаемого события в n испытаниях, $e = 2,7183$; параметр $\lambda = np$ равен математическому ожиданию или наиболее вероятной частоте события $\bar{\mu}$, а также дисперсии.

Биномиальное распределение и игровые задачи

На практике часто нужно вычислять отдельные биномиальные вероятности и суммировать их далее по определенному множеству

целых чисел. Это достаточно трудоемкая процедура (представьте, что n и m – большие числа). В литературе существуют обширные таблицы.

В нынешней реализации вероятностного калькулятора в **Statistica** нет биномиального распределения, однако биномиальное распределение реализовано в языке **Statistica Basic**, и им легко воспользоваться.

Создайте пустую электронную таблицу **Spreadsheet.sta** из 20 строк.

Дважды щелкните на имени переменной **var1** и откройте диалоговое окно спецификации переменной **var1**.

В нижней части окна в поле **Long name** запишите формулу, как показано на рисунке 1.1.

Нажмите кнопку **OK** в правом верхнем углу окна (рисунок 1.1).

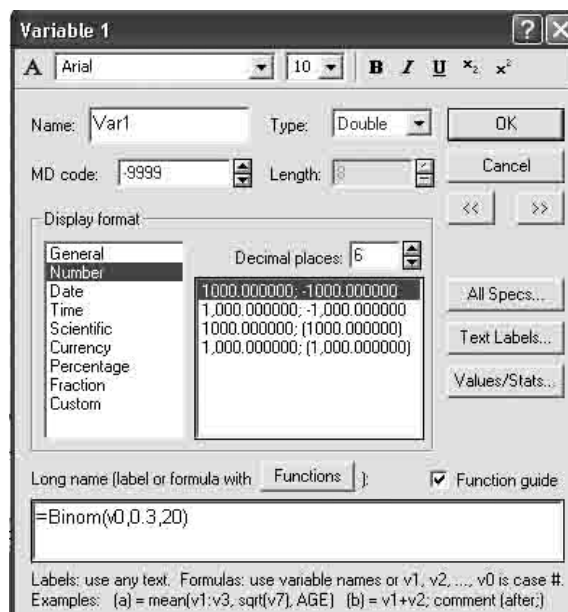


Рисунок 1.1 – Задание формулы вычисления биномиальных вероятностей

Согласно этой формуле программа вычислит вероятности успеха и занесет их в таблицу в значения первой переменной. Теперь таблица примет вид, представленный на рисунке 1.2 а.

В данной таблице вероятность успеха – выпадения герба – равна 0,3. Из таблицы вы видите, что вероятность выпадения ровно одного герба в 20 бросаниях – 0,006839, вероятность выпадения ровно двух гербов в 20 бросаниях – 0,027846 и т. д.

Вероятность успеха легко изменить, сделав ее равной, например, 0,5. Это означает, что бросается симметричная монета и вероятность успеха

равна вероятности неудачи (вероятность выпадения герба равна вероятности выпадения цифры).

Дважды щелкните на имени переменной **var1** и откройте окно спецификации переменной **var1**.

В нижней части окна в поле **Long name** измените формулу, вместо 0,3 запишите 0,5.

Нажмите кнопку **ОК** в правом верхнем углу. Программа вычислит новые биномиальные вероятности и занесет их в электронную таблицу (рисунок 1.2 б).

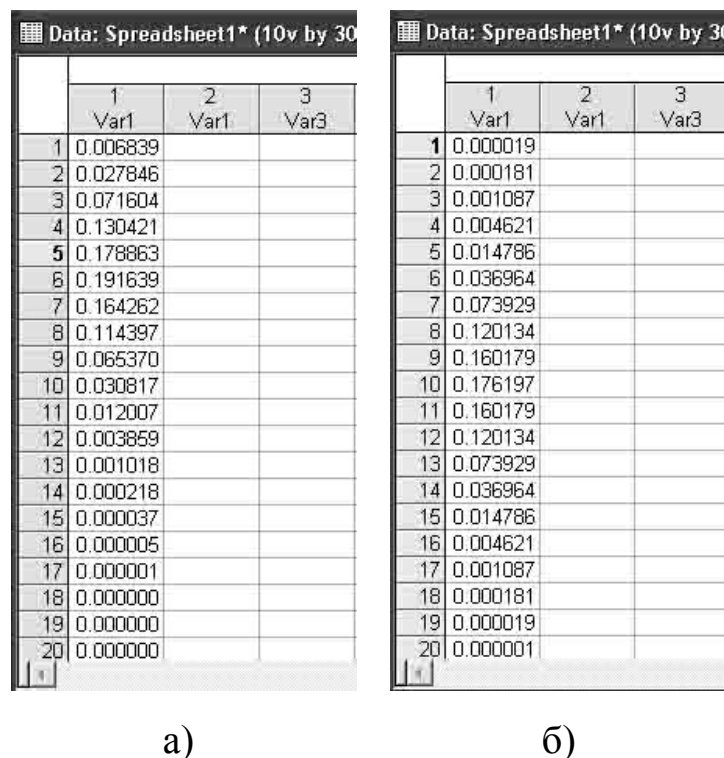


Рисунок 1.2 – Электронная таблица с биномиальными вероятностями:

- а) – вероятность успеха 0,3, число испытаний 20;
- б) – вероятность успеха 0,5, число испытаний 20

Заметьте, что максимальная вероятность в этой таблице приходится на значения $10 \mu = p \cdot n = 0,5 \cdot 20 = 10$, что и понятно из соображений симметрии, а таблице 1.2 а) – на значении 6 ($\mu = p \cdot n = 0,3 \cdot 20 = 6$).

Если вы забудете функцию, которая вычисляет биномиальные вероятности в системе, то воспользуйтесь средством **Function Browser**.

Нажав кнопку **Functions** в окне спецификации переменной, вы откроете диалоговое окно **Function Browser**, в котором легко выбрать нужную функцию биномиального распределения.

Для графического отображения результатов расчетов из модуля **Statistica** выберите **Graphs_2DGraphs_Bar/Column Plots** (рисунок 1.3).

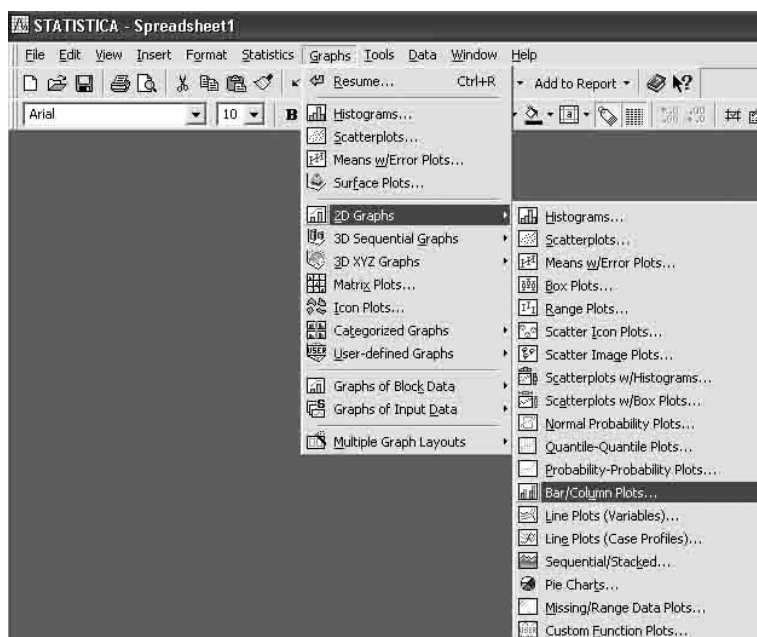


Рисунок 1.3 – Меню запуска графика

Шаг 1. Выполните установки, как показано на рисунке 1.4.

Шаг 2. Выберите столбец с данными, нажав кнопку **Variables**.
Нажмите **ОК**.

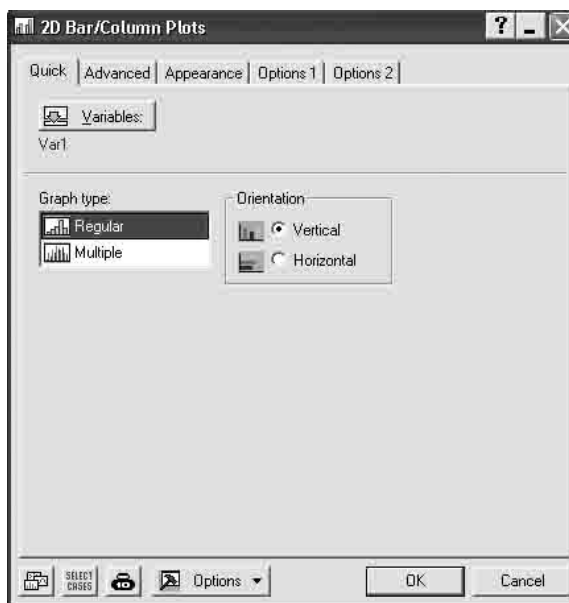


Рисунок 1.4 – Окно выбора графика

В появившемся окне представлен график биномиального распределения $p = 0,3$, число испытаний $n = 20$ (рисунок 1.5).

Аналогичную процедуру выполните для биномиального распределения $p = 0,5$, число испытаний $n = 20$ (рисунок 1.6).
Отредактируйте рисунки.

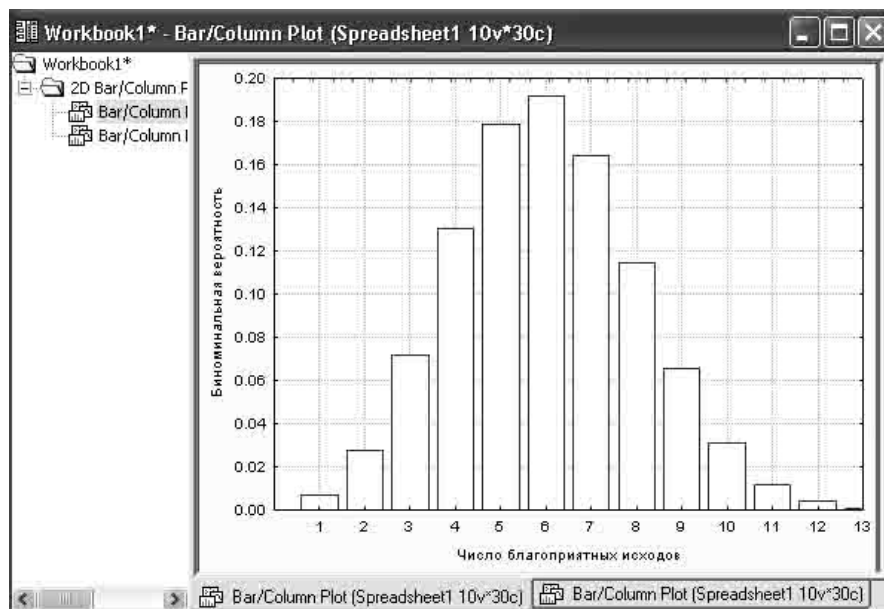


Рисунок 1.5 – График биномиального распределения $p = 0,3, n = 20$

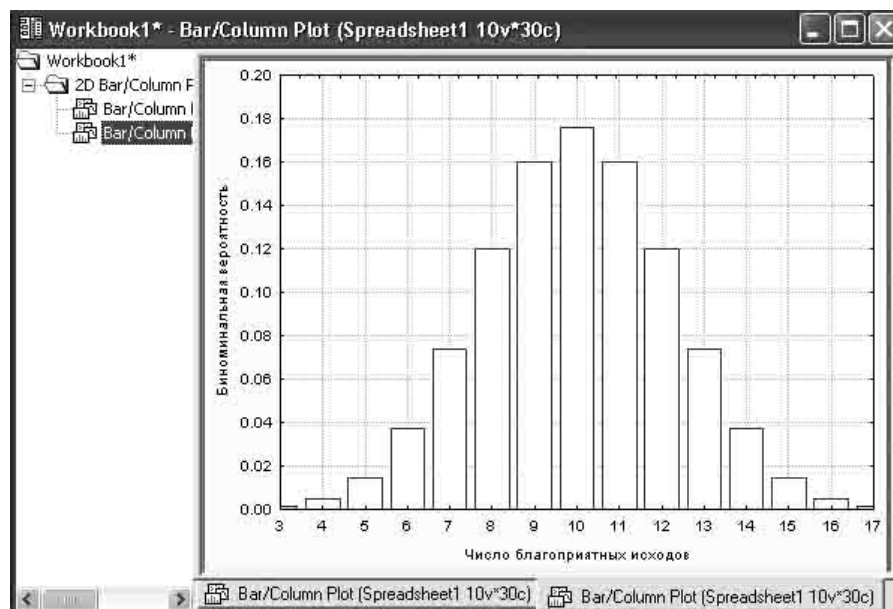


Рисунок 1.6 – График биномиального распределения $p = 0,5, n = 20$

Сравнение двух полученных распределений свидетельствует об асимметрии в первом варианте. При существенном увеличении числа испытаний асимметрия нивелируется.

Задача шевалье де Мере

Классическим и вместе с тем забавным является пример шевалье де Мере, когда ставший известным в веках, благодаря своей любознательности азартный игрок спросил себя: стоит ему ставить на выпадение двух шестерок одновременно при бросании двух костей 24 раза или нет? Его собственные вычисления показали, что стоит, так как вероятность данного события при 24 бросаниях костей больше $1/2$. Как же он удивился, когда с течением времени обнаружил, что постоянно оказывается в проигрыше! Де Мере во всем обвинил статистику. И только знаменитый Паскаль указал ему на заблуждение: оказывается, вероятность данного события 0,49, следовательно, в длинной серии игр, состоящих из 24 подбрасываний двух костей, выигрыш происходит лишь в 49%, а не более чем в 50% играх, как рассчитывал де Мере. Шевалье обычно играл всю ночь, и для него было важно, чтобы более чем в половине игр он был в выигрыше.

Можно «мгновенно» решить эту задачу с помощью самых простых средств **Statistica**.

Создайте рабочий файл **Spreadsheet.sta**. Дважды щелкните на имени переменной и откройте окно спецификации переменной **var1**.

В нижней части окна в поле **Long name** запишите формулу, как показано на рисунок 1.7 а. Число испытаний в задаче шевалье 24. Вероятность успеха равна $1/36$, потому что с такой вероятностью при бросании двух костей выпадают шестерки.

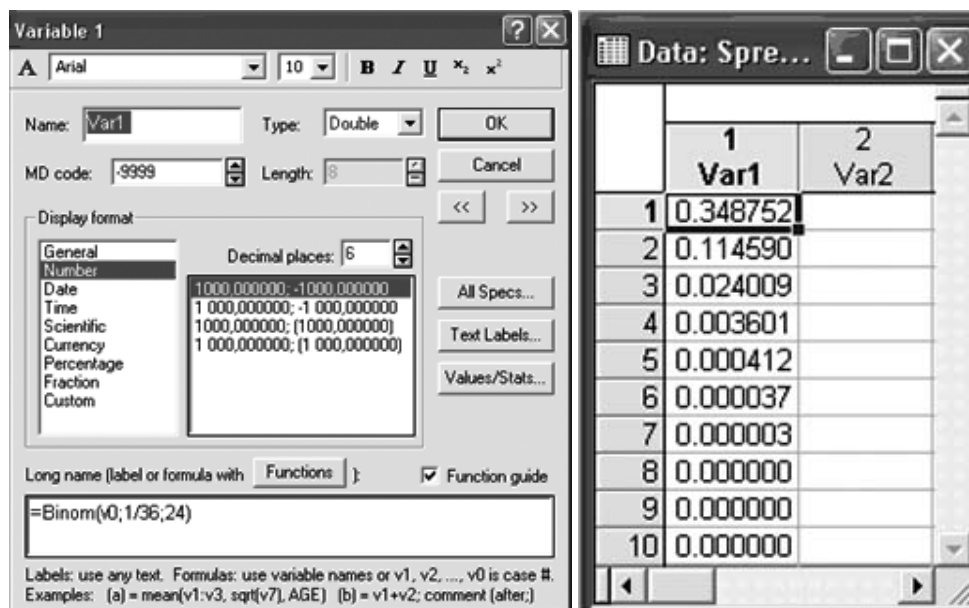
Нажмите кнопку **ОК** в правом верхнем углу. Программа вычислит биномиальные вероятности. Результат представлен на рисунке 1.7 б.

В первом столбце этой таблицы даны последовательно вероятности выпадения двух шестерок один раз, два раза, три раза и т. д.

Шевалье де Мере спросил, стоит ли ему ставить на выпадение двух шестерок одновременно при бросании двух костей 24 раза или нет?

Нужно вычислить вероятность выпадения, по крайней мере, одной пары шестерок. Следовательно, все эти вероятности нужно сложить. Сделав это, вы получите ответ к классической задаче. Вероятность выпадения, по крайней мере, одной пары шестерок при 24 бросаниях пары костей равна 0,49140 (сумма биномиальных вероятностей на рисунке 1.7 б).

Таким образом, в длинной серии игр, состоящих из 24 бросаний пары костей, игрок, ставящий на выпадение двух шестерок одновременно, в среднем устойчиво проигрывает.



а)

б)

Рисунок 1.7 – Вероятности выпадения шестерок при 24 бросаниях двух костей: а) задание формулы; б) результат

Но вот вопрос: как изменить условия игры, чтобы находиться в выигрыше?

Будет ли выигрыш, если игра состоит из 25 бросаний, то есть ставка на выпадение пары шестерок в 25 бросаниях.

Итак, предположим, что шевалье де Мере изменил условия игры и стал ставить на выпадение пары шестерок в 25 бросаниях. Оказывается, увеличение числа бросков всего на 1 делает игру уже выигрышной.

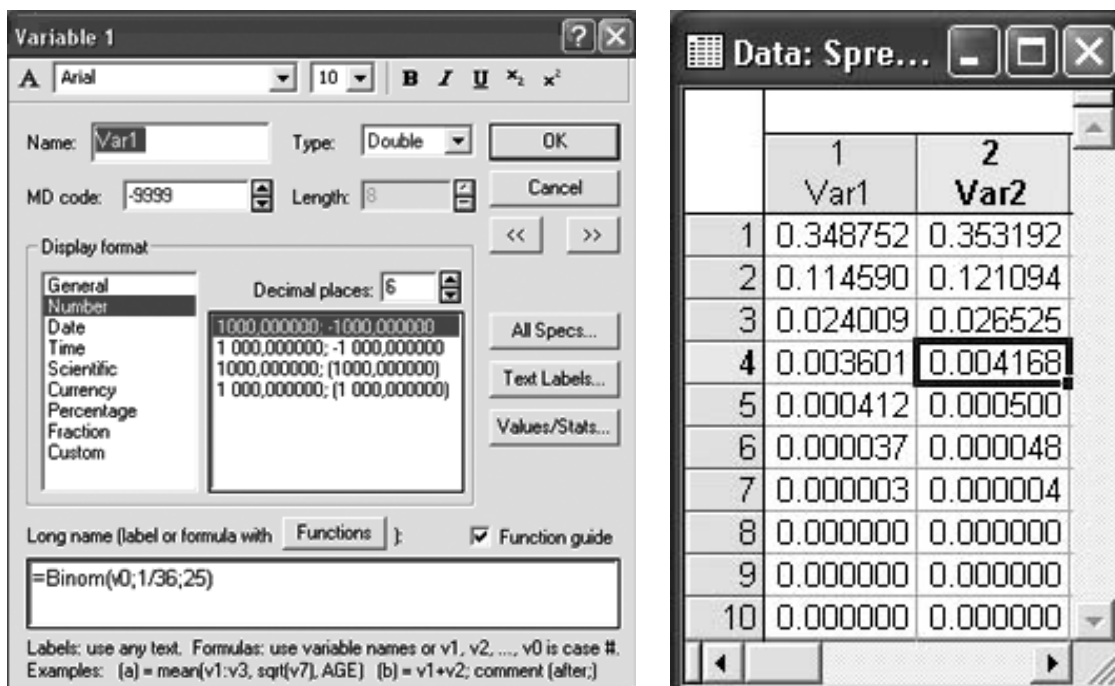
В этом можно убедиться, по-прежнему работая с файлом **Spreadsheet.sta**.

Повторите все действия предыдущей задачи с переменной **var2**. Дважды щелкните на имени переменной и откройте окно спецификации переменной **var2** (рисунок 1.8 а).

Заметьте, единственное отличие этой формулы от формулы в исходной задаче шевалье в том, что сделана ставка на число испытаний 25 вместо 24.

Нажмите кнопку **ОК** в правом верхнем углу. Программа вычислит новые биномиальные вероятности и занесет их в значения переменной **var2**.

Теперь файл **Spreadsheet.sta** будет выглядеть следующим образом (рисунок 1.8 б).



а)

б)

Рисунок 1.8 – Вероятности выпадения шестерок при 25 бросаниях двух костей: а) задание формулы; б) результат

Складывая значения в столбце, легко найти, что вероятность выпадения, по крайней мере, одной пары шестерок в 25 подбрасываниях пары костей больше 0,5.

Если бы шевалье де Мере играл в такую игру, он находился бы в среднем в выигрыше, так как более чем в 50% игр, состоящих из 25 подбрасываний пары костей, по крайней мере, один раз выпадали бы шестерки.

Распределение редких событий (Пуассона)

Создайте рабочий файл **Spreadsheet.sta**. (10 колонок и 30 строк). Дважды щелкните на имени переменной и откройте окно спецификации переменной **var1**.

В нижней части окна в поле **Long name** запишите формулу, как показано на рисунке 1.9. Положим, что вероятность события мала и равна $p = 0,001$, и поэтому для получения полноценного распределения редких событий число испытаний должно быть достаточно велико. В данном случае достаточно 10000 испытаний.

Результат приведен на рисунке 1.10.

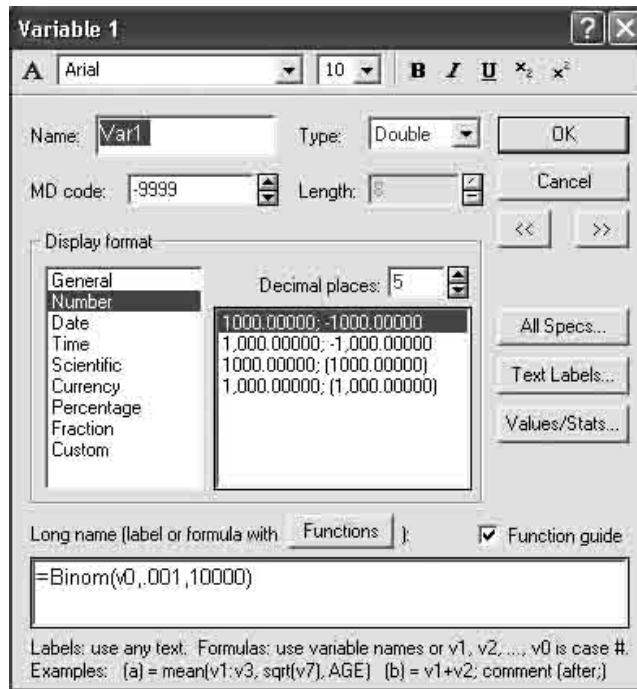


Рисунок 1.9 – Пусковая панель

Data: Spreadsheet1* (10v by 100000c)

	1 Var1	2 Var2	3 Var3
1	0.000161		
2	0.001076		
3	0.004308		
4	0.012319		
5	0.027393		
6	0.049833		
7	0.076700		
8	0.102303		
9	0.120387		
10	0.126742		
11	0.120712		
12	0.104962		
13	0.083957		
14	0.062175		
15	0.042864		
16	0.027642		
17	0.016743		
18	0.009561		
19	0.005164		
20	0.002646		
21	0.001290		
22	0.000599		
23	0.000266		
24	0.000113		
25	0.000046		
26	0.000018		
27	0.000007		
28	0.000002		
29	0.000001		

Рисунок 1.10 – Результат при вероятности события $p = 0,001$ в 10000 испытаниях

Из таблицы (рисунок 1.10) видно, что максимальная биномиальная вероятность находится на значении $10 \mu = p \cdot n = 0,001 \cdot 10000 = 10$.

Для графического отображения результатов расчетов выполните аналогичные процедуры, описанные для биномиального распределения (рисунки 1.3 и 1.4).

После выполнения указанных действий и последующего редактирования будет получен график распределения Пуассона (рисунок 1.11).

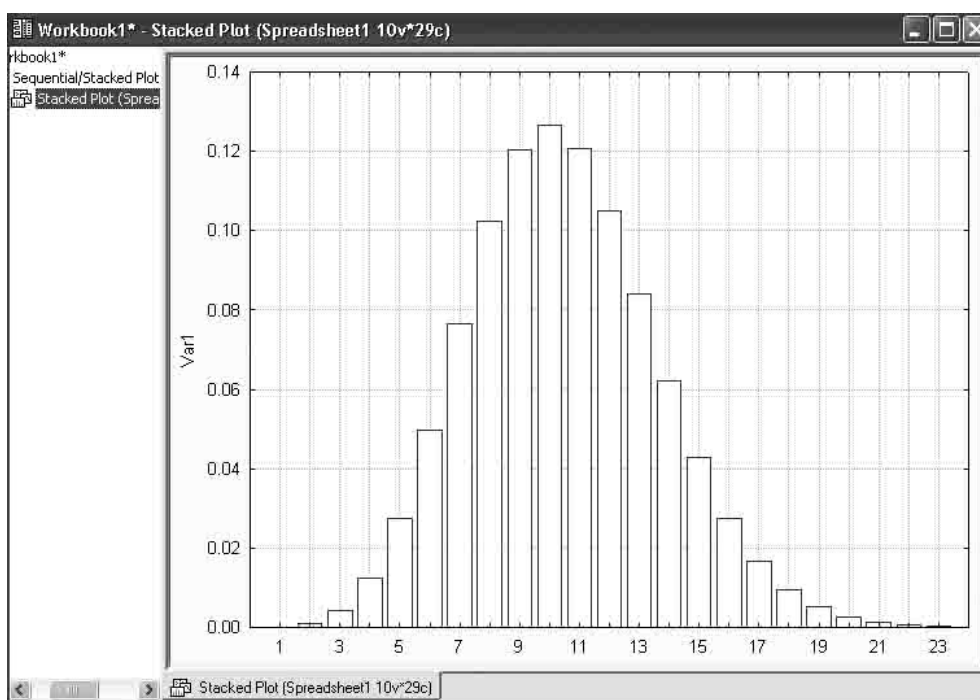


Рисунок 1.11 – График распределения Пуассона

Рисунок 1.11 иллюстрирует, что при достаточно большом числе испытаний биномиальное распределение даже при малой вероятности ожидаемого события становится практически симметричным.

Задания для выполнения

1 С помощью вероятностного калькулятора вычислите отдельные биномиальные вероятности при различных значениях вероятности успеха (таблица 1.1).

2 При задании формулы вычисления биномиальных вероятностей определите число испытаний (n) таким образом, чтобы получить полноценное биномиальное распределение для каждого варианта.

3 Выполните графическое отображение результатов расчетов.

4 Дайте объяснение полученным результатам.

Таблица 1.1 – Задания для выполнения расчетов

№ варианта	Процедура	Вероятность успеха
1	Бросание монеты	0,05
2	Бросание монеты	0,1
3	Бросание монеты	0,25
4	Бросание монеты	0,75
5	Бросание монеты	0,9
6	Бросание монеты	0,99
7	Бросание монеты	0,0004
8	Бросание монеты	0,00005
9	Бросание 1 кубика	Выпадение больше или равно 4 очкам
10	Бросание 1 кубика	Выпадение меньше 4 очков
11	Бросание 2 кубиков	Выпадение больше или равно 8 очкам
12	Бросание 2 кубиков	Выпадение меньше 5 очков

Лабораторная работа 2

Основные модели теоретических распределений (Statistica 6)

Цель работы: научиться работе с вероятностным калькулятором в программном продукте **Statistica 6** на примерах нормального и логарифмически нормального распределений.

Краткие теоретические сведения

Наиболее часто встречающееся в статистике и в теории вероятностей распределение – это нормальное распределение.

Известно, что случайные ошибки в экономических рядах, рядах, возникающих в природе, имеют приблизительно нормальное распределение. Рост взрослых людей также можно приближенно описать нормальным распределением.

Нормальное распределение имеет два параметра:

mean – среднее;

standard deviation – стандартное отклонение.

Эти параметры задаются в окне вероятностного калькулятора.

Иногда стандартное отклонение называют среднеквадратическим отклонением.

Перечислим некоторые признаки нормального распределения.

Плотность нормального распределения симметрична относительно среднего. Среднее значение определяет меру расположения плотности. Среднее значение нормального распределения совпадает с медианой и модой.

Напомним факт, известный из курса элементарной теории вероятностей, что дисперсия дает меру рассеяния плотности вероятности.

Корень квадратный из дисперсии равен стандартному отклонению. Дисперсию часто обозначают сигмой в квадрате (σ^2), а стандартное отклонение – просто сигмой (σ). Дисперсия и стандартное отклонение – положительны. Дисперсия может сколь угодно приближаться к нулю, но может принимать и сколь угодно большое значение, при этом, очевидно, изменяется и распределение вероятности.

Говорят, что случайная величина X имеет логарифмически-нормальное распределение, если величина $\ln(X)$ является нормальной. Это можно выразить так: логарифм логарифмически-нормальной

величины является нормальной величиной. Так как нормальное распределение описывается двумя параметрами, то и логарифмически-нормальное распределение также имеет два параметра.

Вероятностный калькулятор

Запустите модуль **Basic Statistics/Tables** (Основные статистики / таблицы) из переключателя модулей. Высветите в стартовой панели модуля **Basic Statistics/Tables** (Основные статистики/таблицы) строку **Probability calculator** (Вероятностный калькулятор) (рисунки 2.1).

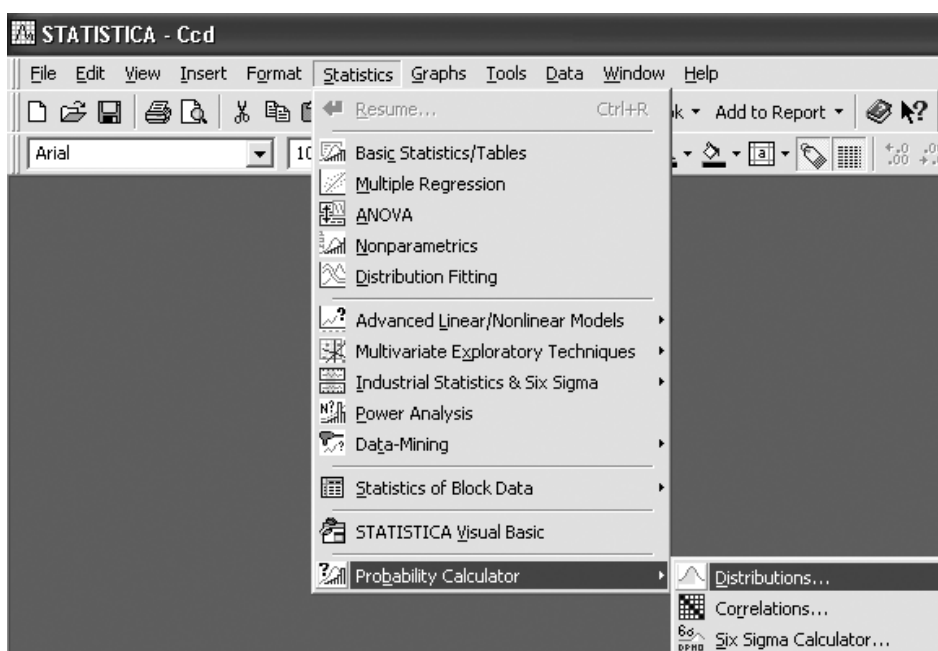


Рисунок 2.1 – Запуск модуля

Нажмите кнопку **ОК**. Откроется окно **Probability Distribution Calculator** (Калькулятор вероятностных распределений) (рисунки 2.2).

Окно имеет следующую структуру – в левой части список распределений **Distribution** (распределение). Многие стандартные распределения в этом окне можно выбрать, высвечивая их названия в списке слева: Бета, Коши, χ -квадрат (хи-квадрат), нормальное, логнормальное, распределение Стьюдента и т. д. Выберем, например, в списке строчку **Z(Normal)** (Нормальное распределение). Автоматически справа появляются поля, где можно задать параметры нормального распределения: **mean** (среднее) и **st. dev** (стандартное отклонение) (рисунки 2.2).

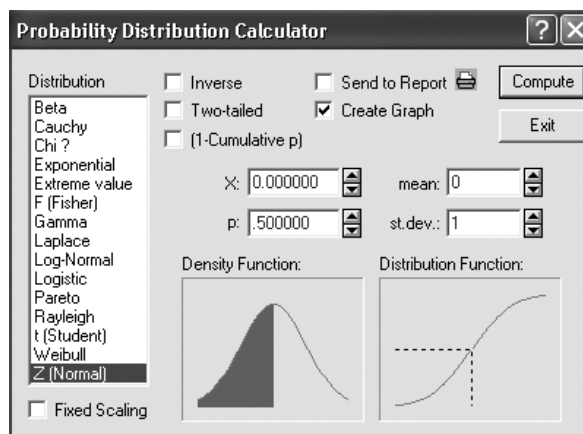


Рисунок 2.2 – Калькулятор вероятностных распределений

Система по умолчанию запишет в них стандартные значения: **среднее = 0**, **стандартное отклонение = 1**. Эти значения можно изменить, поместив курсор мыши в эти поля, щелкнуть левой кнопкой и ввести с клавиатуры нужные величины.

Одновременно с выбором распределения в левом списке справа в калькуляторе появляются графики нормальной плотности и функции распределения: **Density Function** (Функция плотности), **Distribution Function** (Функция распределения).

В поле p задается уровень вероятности. Поместите курсор мыши в это поле и щелкните левой кнопкой. Наберите далее любое значение в интервале от 0 до 1. После нажатия на кнопку **Compute** (Вычислить) (в правом верхнем углу калькулятора) в строке X появится соответствующий квантиль.

То же можно сделать и в обратную сторону – по заданному значению X вычислить уровень вероятности p . Задав какое-либо значение, щелкните по кнопке **Compute** (Вычислить) в правом верхнем углу. В строке p появится уровень для данного значения X .

Опции в верхней части окна имеют следующее назначение: **Inverse** (Обратная функция распределения), **Two-tailed** (Двухсторонний), **1 - Cumulative p** (1 – Кумулята p), **Send to Print** (Отправить на печать), **Create graph** (Создать график).

Если пометить опцию **Create graph** (Создать график) и нажать далее кнопку **Compute** (Вычислить), то на экране появится график плотности и функции распределения (задайте в строке p какое-либо значение, например: $p = 0$ (рисунок 2.3). Таким образом, вероятностный калькулятор заменяет многие таблицы. Теперь, вместо того чтобы использовать таблицы распределений, вы можете использовать данный калькулятор.

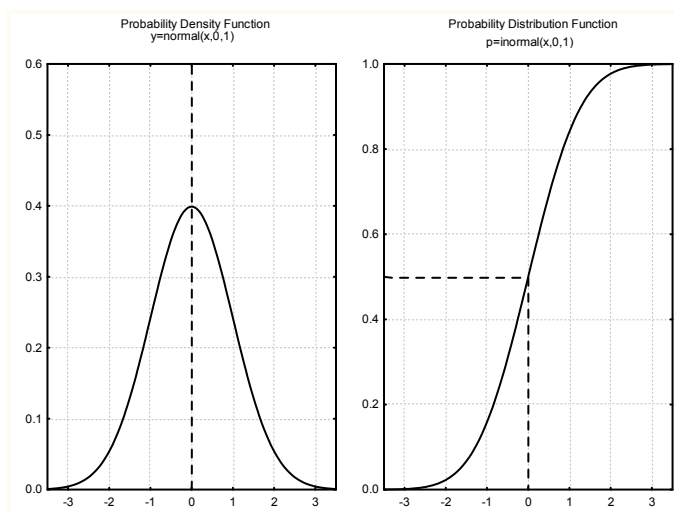


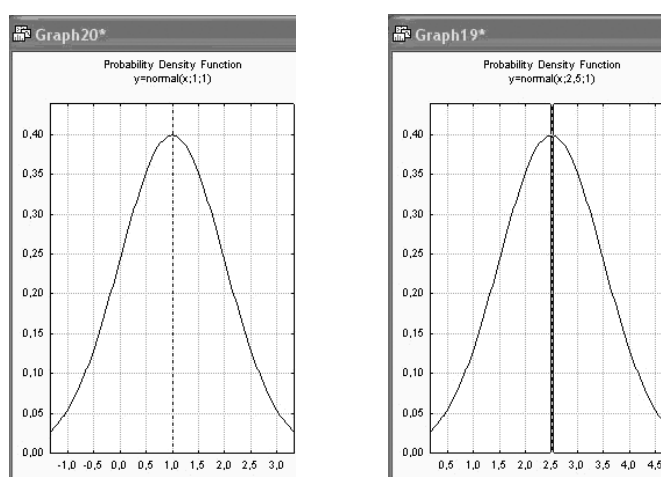
Рисунок 2.3 – Плотность и функция распределения стандартной нормальной величины

Нормальное распределение

Зададим различные значения среднего, оставив пока без изменения стандартное отклонение. Будем считать, что оно равно 1.

Откройте вероятностный калькулятор в поле **mean** (среднее), задав вначале 1. В поле *p* зададим значение 0,5 (в данном примере это чисто техническая установка).

Выберите опцию **Create graph** (Создать график) и нажмите далее кнопку **Compute** (Вычислить), на экране появится график плотности (рисунок 2.4 а). Повторите те же действия, задав в поле **mean** (среднее) значение 2,5. Вы увидите следующий график (рисунок 2.4 б).



а)

б)

Рисунок 2.4 – Плотность нормального распределения со средним: а) 1; б) 2,5

Посмотрите внимательно на эти графики. Вы видите, что график плотности нормального распределения сдвигается по оси ординат при изменении среднего. Можно сказать и более точно: при возрастании среднего графики сдвигаются вправо.

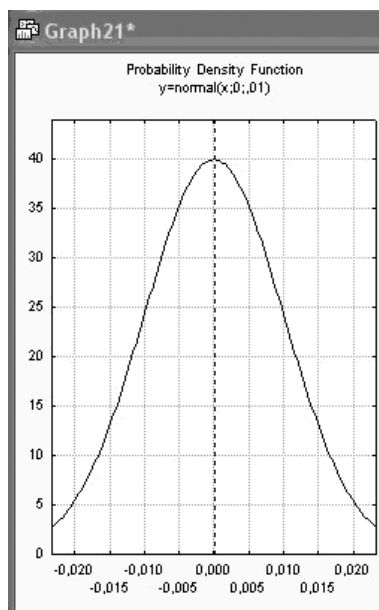
Пик плотности нормального распределения находится в точке с ординатой, равной среднему значению, а плотность симметрична относительно этого значения.

Это значение задается в поле **mean** (среднее).

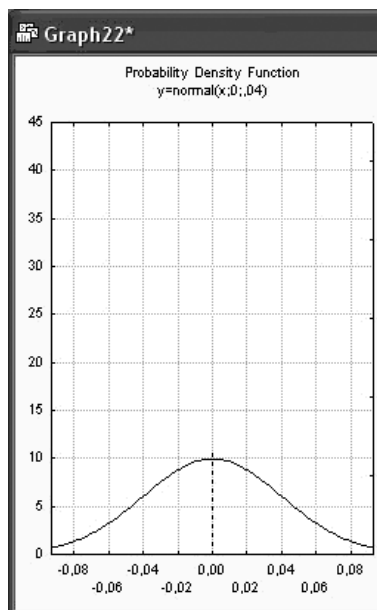
Посмотрим, как меняется плотность распределения при изменении другого параметра – стандартного отклонения.

Зададим различные значения стандартного отклонения, считая, что среднее фиксировано и равно нулю.

Покажем на графиках, как изменяется плотность нормального распределения при уменьшении и увеличении дисперсии (рисунки 2.5 а, 2.5 б).



а)



б)

Рисунок 2.5 – Плотность нормального распределения со средним = 0 и дисперсией: а) 0,01 и б) 0,04

Итак, при увеличении дисперсии плотность нормального распределения расплывается или рассеивается относительно среднего значения, при уменьшении дисперсии она наоборот сжимается, концентрируясь возле одной точки – точки максимального значения.

Рассмотрим пример использования нормального распределения.

Известно, что в некоторой стране рост взрослых мужчин

приближенно имеет нормальное распределение со средним 176,6 см и стандартным отклонением 7,63 см.

Какова вероятность того, что рост наугад выбранного мужчины не больше 185 см и не меньше 175 см?

Шаг 1. Откройте вероятностный калькулятор. Выберите в списке распределений **Z(Normal)** (Нормальное распределение).

Шаг 2. Задайте:

- в поле **mean** – среднее 175,6;
- в поле **st.dev.** – стандартное отклонение 7,63.

Шаг 3. В поле X задайте 185. Нажмите кнопку **Compute** (Вычислить).

В поле p появилось значение 0,891022. Запомните это значение как p_1 .

Шаг 4. В поле X задайте 175. Нажмите кнопку **Compute** (Вычислить).

В поле p появилось значение 0,468661. Запомните это значение как p_2 .

Шаг 5. Вычтите p_2 из p_1 . Вы получите 0,422361.

Итак, с вероятностью 0,422361 встреченный вами мужчина имеет рост не ниже 175 и не выше 185 сантиметров.

Правила 2 и 3 сигма

Правила 2 и 3 сигма полезно знать. Они часто используются на практике. Смысл этих правил состоит в том, что если от точки среднего или от точки максимума плотности нормального распределения отложить вправо и влево соответственно два и три стандартных отклонения (2 и 3 сигма), то площадь под графиком нормальной плотности, подсчитанная по этому промежутку, будет соответственно равна 95,45% и 99,73% всей площади под графиком.

Другими словами, это можно выразить следующим образом: 95,45% и 99,73% всех независимых наблюдений из нормального распределения лежат в пределах двух и трех стандартных отклонений от среднего значения.

Это правило также легко проверить с помощью вероятностного калькулятора. Выберите нормальное распределение в списке распределений, задайте, например, стандартные параметры: среднее = 0, стандартное отклонение = 1, пометьте опцию **Two-tailed** (Двухсторонний), в строке X задайте 2 (два стандартных отклонения), нажмите **Compute**, в строке p появится значение 0,9545 (рисунок 2.6).

В поле **Density Function** (Функция плотности) вероятностного

калькулятора показана заштрихованная площадь под графиком плотности, в поле p показано значение 0,9545. Переходя к процентам, имеем 95,45%. Заштрихованная площадь составляет 95,45% всей площади под графиком.

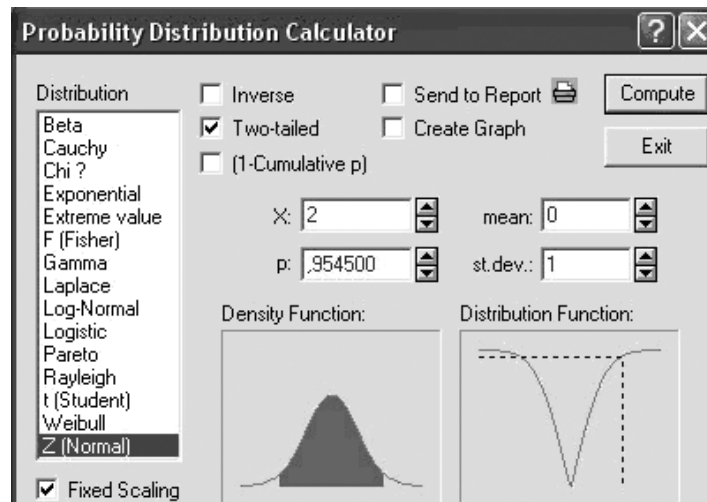


Рисунок 2.6 – Иллюстрация к правилу 2 сигм

Выберите нормальное распределение, задайте стандартные параметры: среднее = 0, стандартное отклонение = 1, пометьте опцию **Two-tailed** в строке X задайте 3 (три стандартных отклонения), нажмите **Compute**, в строке p появится значение 0,9973 (рисунок 2.7).

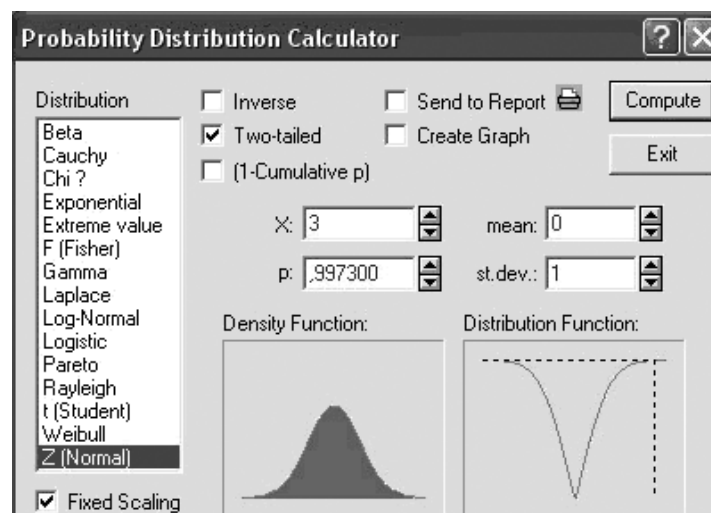


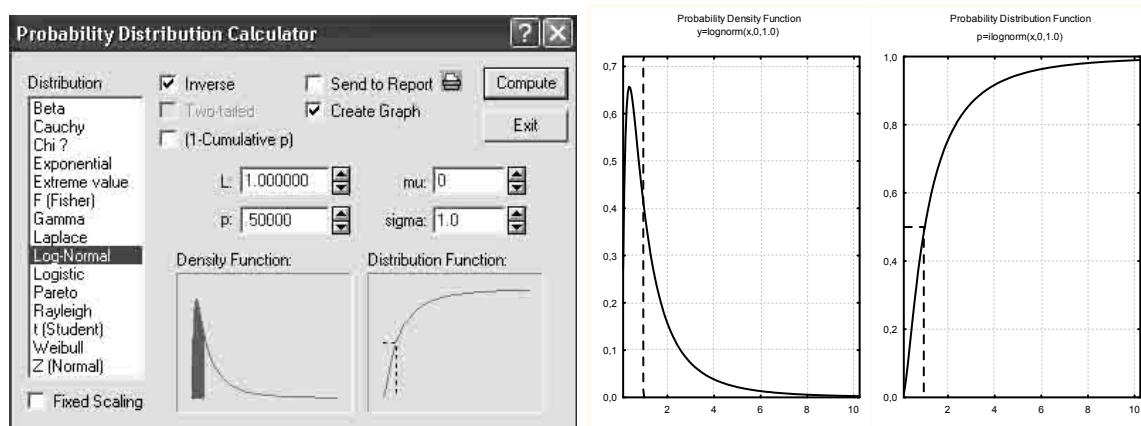
Рисунок 2.7 – Иллюстрация к правилу 3 сигм

Данные правила действуют при любых значениях среднего и стандартного отклонения нормального закона.

Логарифмически-нормальное распределение

Плотность распределения имеет одно максимальное значение и несимметрично. График плотности логарифмически-нормального распределения показан на рисунке 2.8 а.

В списке распределений вероятностного калькулятора выберите **Log-Normal** (Логнормальное распределение), рисунок 2.8 б).



а)

б)

Рисунок 2.8 – Логарифмически-нормальное распределение в вероятностном калькуляторе: а) задание; б) плотность распределения

Задания для выполнения

1 С помощью вероятностного калькулятора вычислите вероятности при заданном среднем и стандартном отклонении (таблица 2.1).

2 Дайте объяснение полученным результатам.

Таблица 2.1 – Задание для выполнения расчетов

№ варианта	Среднее	Стандартное отклонение	Минимум	Максимум
1	10	8	8	12
2	10	5	8	12
3	10	3	8	12
4	10	1	8	12
5	10	0,5	8	12
6	20	5	8	12
7	15	5	8	12
8	10	5	8	12
9	5	5	8	12
10	1	5	8	12

Лабораторная работа 3

Распределение параметров выборки

Цель работы: научиться работе с вероятностным калькулятором в программном продукте **Statistica 6** на примерах распределений: **F-распределение** Фишера, **t-распределение** Стьюдента и распределение **χ -квадрат**.

Краткие теоретические сведения

F-распределение возникает в регрессионном, дисперсионном и дискриминантном анализе, а также в других видах многомерного анализа данных. Далее оно будет неоднократно встречаться в таблицах вывода системы **Statistica 6**.

Случайная величина, имеющая **F-распределение** с парой степеней свободы m , n , определяется как отношение двух независимых случайных величин, имеющих распределение **χ -квадрат** (хи-квадрат) со степенями свободы m и n с умножением на нормировочный сомножитель n/m .

F-распределение сосредоточено на положительной полуоси. Это распределение в отличие от нормального несимметрично. Покажем, как построить график **F-распределения** и вычислить его процентные точки.

t-распределение важно в тех случаях, когда рассматриваются оценки среднего и неизвестна дисперсия выборки. В этом случае используют выборочную дисперсию и **t-распределение**.

t-распределение возникает в таблицах вывода регрессионного анализа. Это одно из важнейших распределений, наряду с нормальным и распределением **χ -квадрат** (хи-квадрат).

t-распределение с k -степенями свободы сосредоточено на всей действительной оси, симметрично относительно нуля. Среднее **t-распределения** равно нулю, дисперсия равна $k/(k-2)$.

Случайная величина, имеющая распределение **χ -квадрат**, определяется как сумма квадратов k независимых стандартных нормальных величин. Нормальные случайные величины – это величины, имеющие нормальное распределение. Число k в определении **χ -квадрата** называется числом степеней свободы. В частном случае, когда $k = 1$ случайная величина **χ -квадрат** равна квадрату стандартной нормальной величины. Итак, это распределение имеет только один параметр – число степеней свободы, являющийся целым

положительным числом.

F-распределение Фишера

В списке распределений вероятностного калькулятора выберите **F (Fisher)** (*F*-распределение) (рисунок 3.1 а).

Задайте в поле **df1** (степень свободы 1) значение 10, в поле **df2** (степень свободы 2) — значение 11. Пометьте опцию **Create Graph** (Создать график).

В поле *p* задайте 0,5. Нажав кнопку **Compute** (Вычислить), на экране вы увидите следующий график (рисунок 3.1 б).

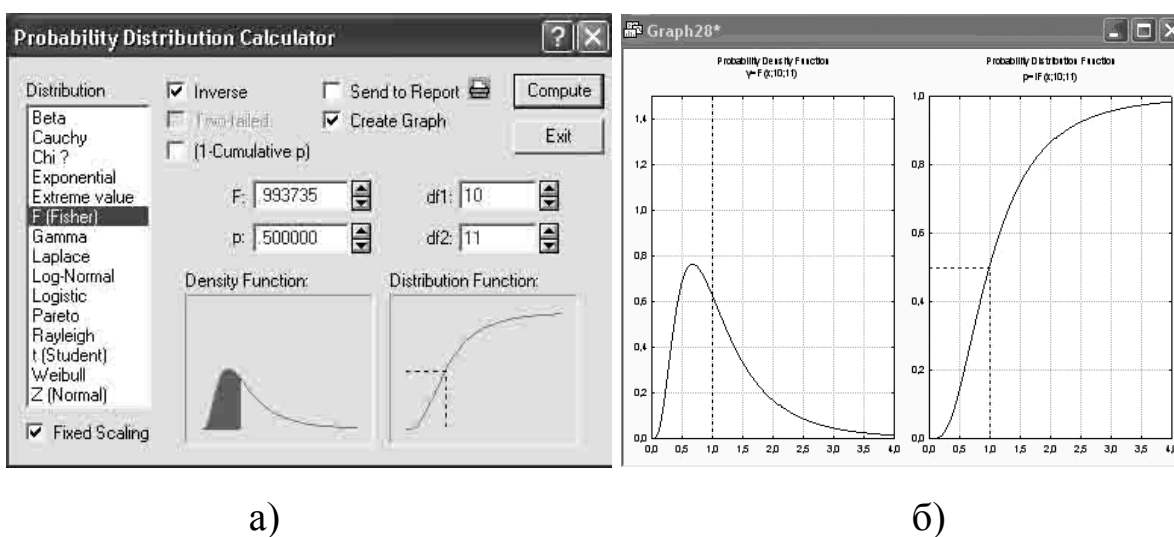


Рисунок 3.1 – *F*-распределение в вероятностном калькуляторе:
а) задание; б) плотность и функция *F*-распределения со степенями свободы 10, 11

t-распределение Стьюдента

В списке распределений вероятностного калькулятора выберите **t (Student)** (*t*-распределение Стьюдента) (рисунок 3.2 а).

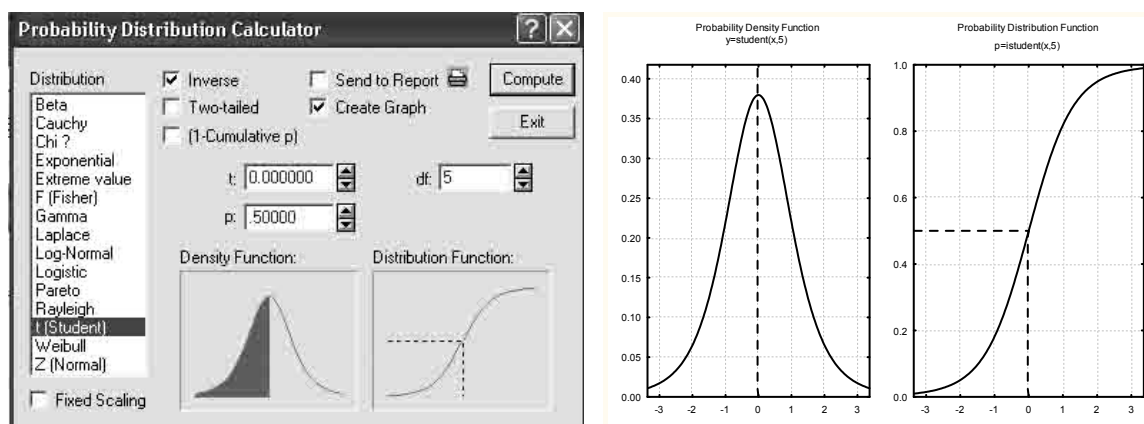
В строке **df** задайте 5 – число степеней свободы. Пометьте опцию **Create Graph** (Создать график).

В поле *p* задайте 0,5. Нажмите кнопку **Compute** (Вычислить), на экране вы увидите следующий график (рисунок 3.2 б).

При больших степенях свободы (больших 30) **t-распределение** практически совпадает со стандартным нормальным распределением.

Плотность **t-распределения** деформируется при возрастании числа степеней свободы следующим образом: пик увеличивается, хвосты более круто идут к нулю, кажется, как будто плотность сжимается с

БОКОВ.



а)

б)

Рисунок 3.2 – t-распределения Стьюдента в вероятностном калькуляторе: а) задание; б) плотность и функция t-распределения Стьюдента с 5 степенями свободы

В такой деформации плотности легко убедиться с помощью вероятностного калькулятора. Задайте в поле **df** (степень свободы) значение 50. Нажав кнопку **Compute** (Вычислить), на экране вы увидите следующий график (рисунок 3.3).

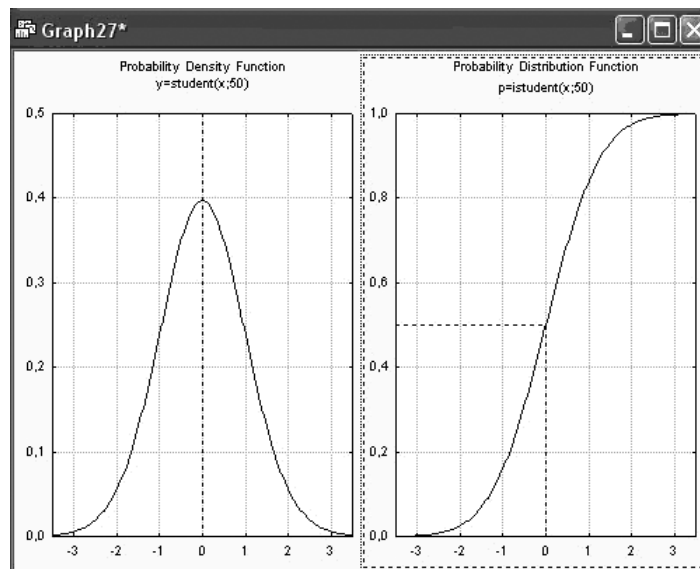


Рисунок 3.3 – Плотность и функция распределения Стьюдента с 50 степенями свободы

Сравнив график плотности **распределения Стьюдента** с большим числом степеней свободы, например 50, и график плотности стандартного нормального распределения, вы убедитесь, что они очень

ПОХОЖИ.

Распределение χ -квадрат (хи-квадрат)

В списке распределений вероятностного калькулятора выберите **Chi ?** (χ -квадрат) (рисунок 3.4).

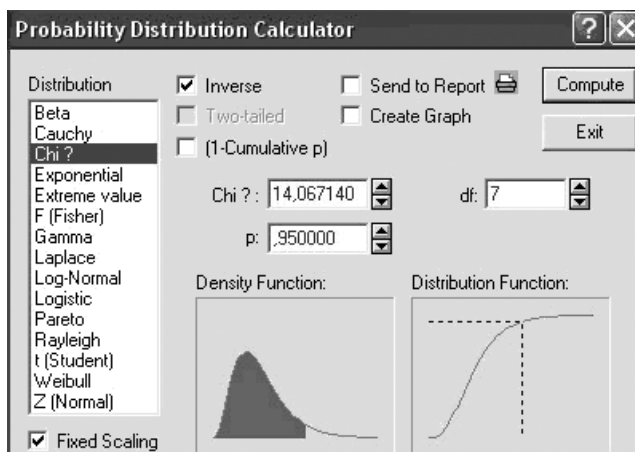


Рисунок 3.4 – Задание χ -квадрат-распределения в вероятностном калькуляторе

В строке **df** задайте 7 – число степеней свободы.

В поле **p** задайте 0,95. Нажмите кнопку **Compute** (Вычислить), в строке **Chi ?** вы увидите 0,95 – квантиль χ -квадрат-распределения с 7 степенями свободы (14,06714).

Выберите далее опцию **Create Graph** (Создать график) и вновь щелкните на кнопке **Compute** (Вычислить) либо просто нажмите **Enter** на клавиатуре, вы увидите график плотности и функции распределения χ -квадрат с 7 степенями свободы (рисунок 3.5).

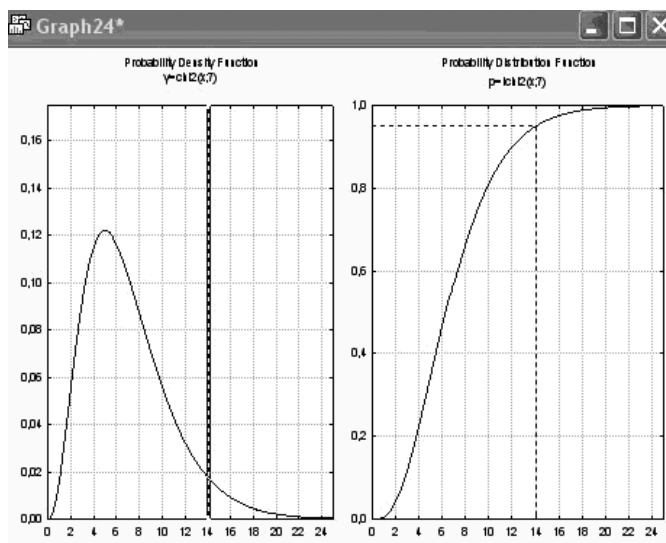


Рисунок 3.5 – График плотности и функции распределения случайной величины χ -квадрат с 7 степенями свободы

Обратите внимание на то, что это распределение несимметрично и сосредоточено только на положительной полуоси.

Распределение χ -квадрат играет важную роль при исследовании оценки дисперсии нормальной выборки, а также при проверке зависимостей в таблицах сопряженности и в критериях согласия.

Задания для выполнения

1 С помощью вероятностного калькулятора вычислите значения распределений **F-Фишера**, **t-Стьюдента** и **χ -квадрат** при различных значениях степеней свободы для уровня надежности 0,95.

2 Дайте объяснение полученным результатам.

Таблица 3.1 – Варианты заданий

№ задания	Распределение	Степени свободы	
		df ₁	df ₂
1	F-Фишера	1	10
2	F-Фишера	2	10
3	F-Фишера	3	10
4	F-Фишера	4	10
5	F-Фишера	5	10
6	F-Фишера	10	1
7	F-Фишера	10	2
8	F-Фишера	10	3
9	F-Фишера	10	4
10	F-Фишера	10	5
11	F-Фишера	10	10
12	t-Стьюдента	1	
13	t-Стьюдента	2	
14	t-Стьюдента	3	
15	t-Стьюдента	10	
16	t-Стьюдента	20	
17	t-Стьюдента	50	
18	χ -квадрат	1	
19	χ -квадрат	3	
20	χ -квадрат	5	
21	χ -квадрат	10	
22	χ -квадрат	20	
23	χ -квадрат	100	

Лабораторная работа 4

Разнообразие значений признака (MS Excel)

Цель работы: Научить выполнять первичную обработку данных в табличном редакторе MS Excel.

Краткие теоретические сведения

Всякая группа состоит из особей или объектов, отличающихся друг от друга по каждому из признаков.

Средняя величина какого-нибудь признака определяется для того, чтобы получить характеристику этого признака для всей изучаемой группы в целом:

$$\mu = \sum_i \frac{X_i}{n}.$$

(4.1)

Степень разнообразия особей в группе по изучаемому признаку измеряется несколькими показателями, из которых наибольшее значение имеет стандартное отклонение:

$$\sigma = \sqrt{\frac{x^2}{n-1}}.$$

(4.2)

Скошенность кривой называется асимметрией:

$$A = \frac{\sum x^3}{n \cdot \sigma^3}.$$

(4.3)

Правосторонняя асимметрия – отрицательна, левосторонняя – положительна.

Отклонение крутизны называют эксцессом:

$$E = \frac{\sum x^4}{n \cdot \sigma^4} - 3.$$

(4.4)

Эксцесс положителен при островершинной кривой, отрицателен при плосковершинной.

Ошибка средней арифметической:

$$s_\mu = \frac{\sigma}{\sqrt{n}}.$$

(4.5)

Ошибка стандартного (среднего квадратического) отклонения:

$$s_{\mu} = \frac{\sigma}{\sqrt{2 \cdot n}} = \frac{s_{\mu}}{\sqrt{2}} = 0.707 \times s_{\mu}.$$

(4.6)

Ошибка показателя асимметрии:

$$s_A = \sqrt{\frac{6}{n}}.$$

(4.7)

Ошибка показателя эксцесса:

$$s_E = \sqrt{\frac{24}{n}} = 2 \cdot s_A.$$

(4.8)

Проверка выбросов (выпадов, артефактов) должна проводиться всегда перед началом обработки полученных первичных данных. Если подтвердится, что резко выделяющееся значение действительно не может относиться к объектам данной группы, и попало в записи вследствие ошибок внимания, следует такой выброс исключить из обработки. Проверка выбросов может производиться по критерию, равному нормированному отклонению выброса:

$$T = \frac{\hat{X}_i - \mu}{\sigma} \geq T_{st},$$

(4.9)

где:

T – критерий выброса;

\hat{X}_i – выделяющееся значение признака (или очень большое или очень малое);

μ , σ – средняя и сигма, рассчитанные для группы, включающей артефакт;

T_{st} – стандартные значения критерия выбросов, определяемых по таблице 4.1.

Таблица 4.1 – Стандартные значения критерия выбросов (T_{st})

n	T_{st}	n	T_{st}	n	T_{st}	n	T_{st}
2	2,0	16 – 20	2,4	47 – 66	2,8	125 – 174	3,2
3 – 4	2,1	21 – 28	2,5	67 – 84	2,9	175 – 349	3,3
5 – 9	2,2	29 – 34	2,6	85 – 104	3,0	350 – 599	3,4
10 – 15	2,3	35 – 46	2,7	105 – 124	3,1	600 – 1500	3,5

Если $T \geq T_{st}$, то анализируемое значение признака является выбросом. Альтернатива $T < T_{st}$ не позволяет исключить из анализа значение признака.

Для того чтобы оценить генеральный параметр для количественных признаков в форме доверительных границ необходимо:

1 Проверить на нормальность распределения исходных данных.

2 Установить число степеней свободы по правилам, приведенным при описании оценки каждого параметра.

3 Установить, исходя из ответственности исследования, порог вероятности безошибочных прогнозов ($\beta_1 = 0,95, \beta_2 = 0,99, \beta_3 = 0,999$).

4 В соответствии с числом степеней свободы найти значение критерия надежности t по таблице стандартных значений критерия Стьюдента. При отсутствии таблицы показатель надежности для данного исследования можно приближенно определить по приведенным формулам. Если объем выборки превышает нижние пределы больших выборок ($n > 30, n > 100, n > 200$), то показатели надежности берутся постоянные для каждого порога вероятности: $t_1 = 2,0; t_2 = 2,6; t_3 = 3,3$.

5 Рассчитать ошибку выборочного показателя.

6 Определить возможную погрешность оценки генерального параметра, помножив критерий надежности на ошибку репрезентативности: $\Delta = t \cdot s_{\mu}$.

7 Установить доверительные границы генерального параметра; возможный максимум $\bar{\mu} = \mu + \Delta$ и гарантированный минимум $\bar{\mu} = \mu - \Delta$.

Подготовительные процедуры

Формат представления исходных данных для выполнения первичного статистического анализа выполняется записью информации в одну строку или один столбец (таблица 4.2).

Таблица 4.2 – Исходные данные

a	b	c
16,6	19,3	3,9
14	15,3	5,1
16,1	25,3	4,7
12,6	24	4,2
13,2	15,6	4,7
14,9	16,3	4,5
14,6	18,5	4,4
15,1	21,1	5,3

4,1	21,2	14,2
	20,2	

Первая строчка в исходных данных (нецифровая информация) – **Метка в первой строке.**

Первой подготовительной процедурой является сортировка исходных данных. Для этого выделите анализируемый массив (например: массив *a*), а затем в меню выберите **Сортировка по возрастанию** (рисунок 4.1) или **Сортировка по убыванию** и щелкните левой кнопкой мыши.



Рисунок 4.1 – Подготовка к процедуре Сортировка

В появившемся окне выберите действие: сортировать в пределах указанного выделения и нажмите кнопку **Сортировка** (рисунок 4.2). В результате столбец *a* будет отсортирован от минимального значения до максимального. Аналогичную операцию выполните для столбцов *b* и *c* (таблица 4.3).

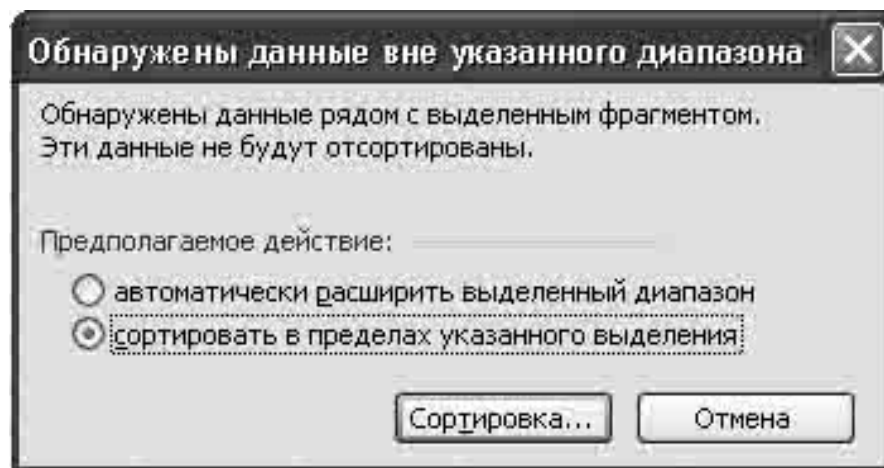


Рисунок 4.2 – Выполнение Сортировки

Таблица 4.3 – Отсортированные данные

a	b	c
4,1	15,3	3,9
12,6	15,6	4,2
13,2	16,3	4,4
14	18,5	4,5
14,6	19,3	4,7
14,9	20,2	4,7
15,1	21,1	5,1
16,1	21,2	5,3
16,6	24	14,2
	25,3	

Для активации модуля **Анализ данных** необходимо в меню **Сервис** левой кнопкой мышки щелкните **Настройка** и в появившемся окне установить флажок **Пакет анализа**.

Первичный анализ статистических данных

Шаг 1. В системе **Excel** в меню откройте модуль **Анализ данных** (рисунок 4.3).

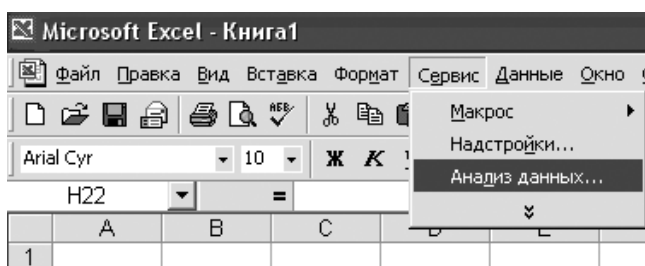


Рисунок 4.3 – Открытие модуля **Анализ данных**

Шаг 2. В модуле **Анализ данных** выберите **Описательная статистика**, после чего щелкните мышкой **ОК** (рисунок 4.4).

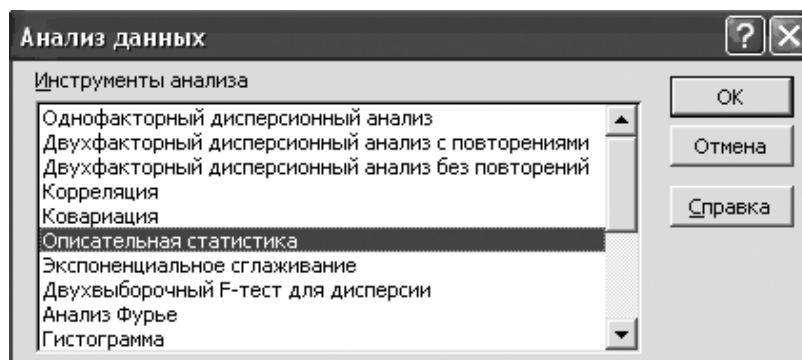


Рисунок 4.4 – Окно Анализа данных – описательная статистика

Шаг 3. В появившемся окне выполните операции и установки, как показано на рисунке 4.5. Щелкните мышкой **ОК**.

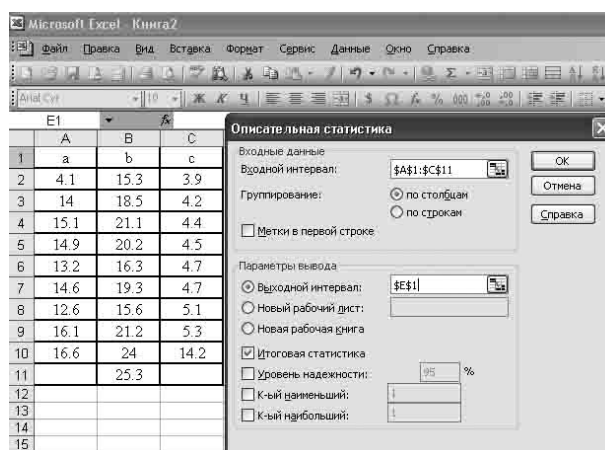


Рисунок 4.5 – Стартовая панель

Результат обработки появится в указанном поле (выходной интервал \$E\$1 – или любая свободная ячейка). На рисунке 4.6 показаны результаты статистической обработки.

	A	B	C		E	F	G	H
1	a	b	c			a	b	c
2	4.1	15.3	3.9		Среднее	13.46667	19.68	5.666667
3	14	18.5	4.2		Стандартн	1.24499	1.073085	1.07613
4	15.1	21.1	4.4		Медиана	14.6	19.75	4.7
5	14.9	20.2	4.5		Мода	#И/Д	#И/Д	4.7
6	13.2	16.3	4.7		Стандартн	3.73497	3.393392	3.22839
7	14.6	19.3	4.7		Дисперсия	13.95	11.51511	10.4225
8	12.6	15.6	5.1		Эксцесс	6.24106	-0.82847	8.554299
9	16.1	21.2	5.3		Асимметр	-2.3582	0.257282	2.899493
10	16.6	24	14.2		Интервал	12.5	10	10.3
11		25.3			Минимум	4.1	15.3	3.9
12					Максимум	16.6	25.3	14.2
13					Сумма	121.2	196.8	51
14					Счет	9	10	9

Рисунок 4.6 – Описательная статистика (результат обработки)

Примечание: эта опция позволяет обрабатывать любое количество выборок одновременно.

Следующей процедурой является проверка в выборках выбросов. Из результатов обработки, представленных на рисунке 4.6, обращают на себя внимание высокие значения эксцесса и асимметрии для выборок в столбцах *a* и *c*. Поэтому можно предположить, что крайние значения (минимальные или максимальные) являются выбросами. Для выборки столбца *a* – это значение 4,1, для выборки столбца *c* – 14,2.

По формуле 4.9 рассчитываем *T* – критерий выброса.

$$T = \frac{4,1 - 13,47}{3,73} = 2,51 \text{ – для выборки в столбце } a;$$

$$T = \frac{14,2 - 5,67}{3,23} = 2,64 \text{ – для выборки в столбце } c.$$

Из таблицы 4.1 при $n = 9$ стандартное значение критерия выброса $T_{st} = 2,2$. Так как *T* – критерий выброса в выборках больше, чем стандартное значение критерия выброса T_{st} , то значения 4,1 в выборке *a* и 14,2 в выборке *c* имеем полное право удалить.

Таким образом, окончательный вариант трех анализируемых выборок представлен в таблице 4.4.

Таблица 4.4 – Окончательный вариант

a	b	c
12,6	15,3	3,9
13,2	15,6	4,2
14	16,3	4,4
14,6	18,5	4,5
14,9	19,3	4,7

15,1	20,2	4,7
16,1	21,1	5,1
16,6	21,2	5,3
	24	
	25,3	

Теперь необходимо выполнить процедуру обработки данных таблицы 4.4, как это было сделано выше (шаги 1 – 3). Результат обработки представлен на рисунке 4.7.

	A	B	C	D	E	F	G	H
1	a	b	c			a	b	c
2	14	15,3	3,9		Среднее	14,6375	19,68	4,6
3	15,1	18,5	4,2		Стандартная ошибка	0,479932	1,073085	0,161466
4	14,9	21,1	4,4		Медиана	14,75	19,75	4,6
5	13,2	20,2	4,5		Мода	#N/D	#N/D	4,7
6	14,6	16,3	4,7		Стандартное отклонение	1,357453	3,393392	0,456696
7	12,6	19,3	4,7		Дисперсия выборки	1,842679	11,51511	0,208571
8	16,1	15,6	5,1		Экссесс	-0,74709	-0,82847	-0,40602
9	16,6	21,2	5,3		Асимметричность	-0,06736	0,257282	0,107982
10		24			Интервал	4	10	1,4
11		25,3			Минимум	12,6	15,3	3,9
12					Максимум	16,6	25,3	5,3
13					Сумма	117,1	196,8	36,8
14					Счет	8	10	8

Рисунок 4.7 – Описательная статистика (окончательный результат обработки)

Проверка на условие нормальности распределения

Для выполнения этой операции, прежде всего, необходимо по формулам 4.7, 4.8 вычислить ошибки показателей эксцесса и асимметрии.

$$s_A = \sqrt{\frac{6}{8}} = 0,31 \text{ – для выборок } a \text{ и } c; \quad s_A = \sqrt{\frac{6}{10}} = 0,24 \text{ – для выборки } b;$$

$$s_E = 2 \cdot 0,31 = 0,61 \text{ – для выборок } a \text{ и } c; \quad s_E = 0,49 \text{ – для выборки } b.$$

Затем найти отношения значений эксцесса и асимметрии по модулю к их ошибкам.

$$t_A = \left| \frac{-0,067}{0,31} \right| = 0,22; \quad t_E = \left| \frac{-0,747}{0,61} \right| = 1,22 \text{ – для выборки } a;$$

$$t_A = \left| \frac{0,257}{0,24} \right| = 1,05; \quad t_E = \left| \frac{-0,828}{0,49} \right| = 1,69 \text{ – для выборки } b;$$

$$t_A = \left| \frac{-0,108}{0,31} \right| = 0,35; \quad t_E = \left| \frac{-0,406}{0,61} \right| = 0,66 \text{ – для выборки } c.$$

Так как найденные значения меньше критического равного трем, то все три выборки удовлетворяют условию нормальности распределения.

Гистограмма

Для построения гистограммы необходимо определить величину класса (кармана) по формуле:

$$k = \frac{X_{\max} - X_{\min}}{n}, \quad n = 1 + 3,322 \cdot \lg N,$$

где N – число наблюдений.

Число классов n округляется до ближайшего целого вверх (например: для столбца a : $n = 4,00006$ – до 5, для столбца b : $n = 4,32$ – до 5, для столбца c : $n = 4,00006$ – до 5)

Размеры интервалов для приведенных в таблице примеров представлены в таблице 4.5.

Таблица 4.5 – Размер интервалов

a	14	14,52	15,04	15,56	16,08	16,6
b	15,3	17,3	19,3	21,3	23,3	25,3
c	3,9	4,18	4,46	4,74	5,02	5,3

Если не рассчитывать размеры интервалов, то они будут определены автоматически.

Опция гистограмма проводит обработку только по одной выборке. Поэтому в примере покажем порядок обработки только для выборки a . Для остальных двух выборок процедура обработки аналогична.

Шаг 1. Вернитесь в модуль **Анализ данных** и выберите опцию **Гистограмма**, после чего щелкните мышкой **ОК** (рисунок 4.8).

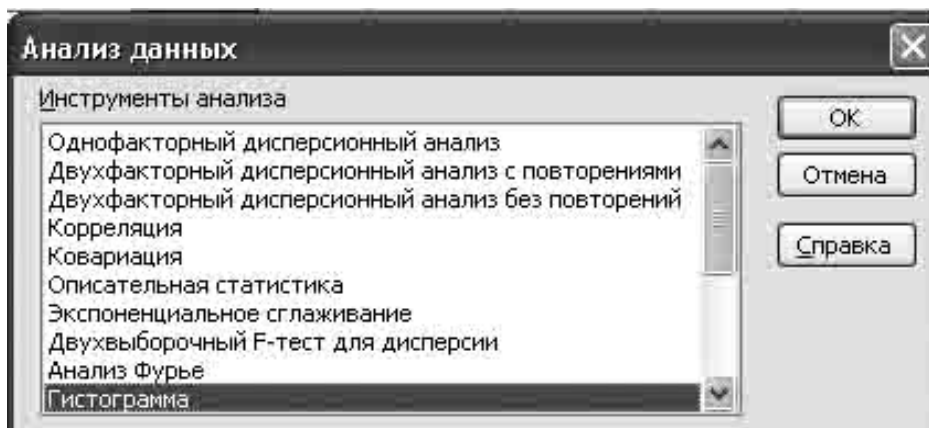


Рисунок 4.8 – Окно **Анализа данных** – гистограмма

Шаг 2. В появившемся окне выполните операции и установки, как показано на рисунке 4.9, после чего щелкните мышкой **ОК**.

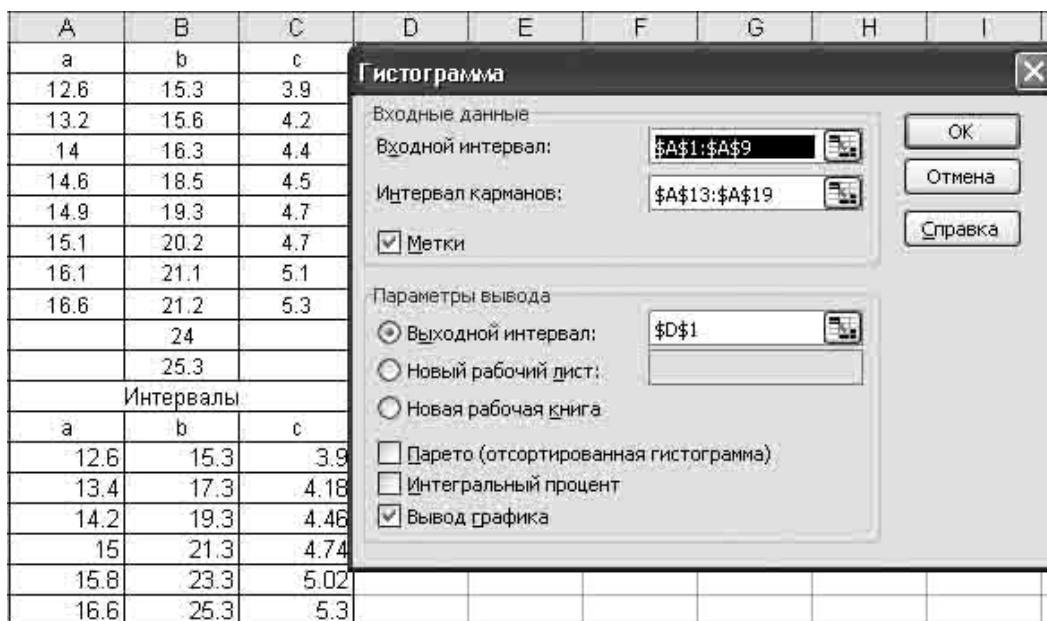


Рисунок 4.9 – Стартовая панель

Шаг 3. Результат обработки появится в указанном поле (выходной интервал \$D\$1 – или любая свободная ячейка, рисунок 4.10).

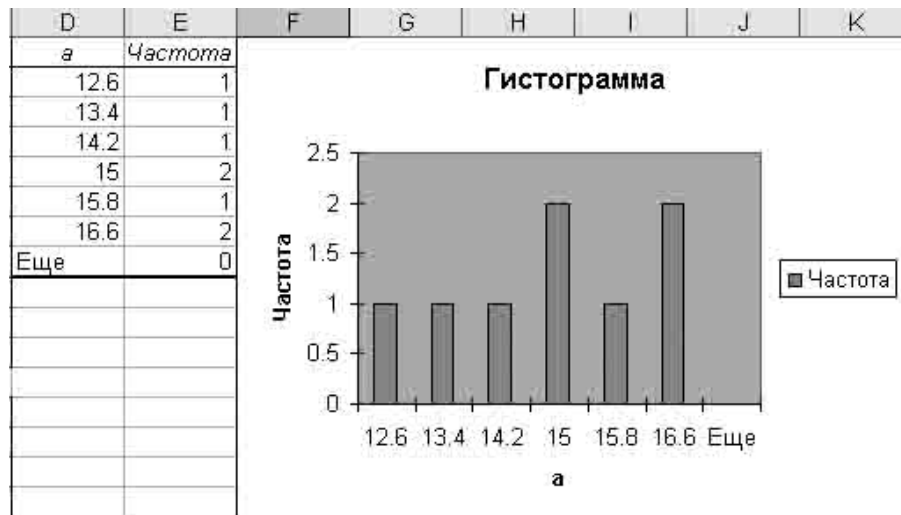


Рисунок 4.10 – Гистограмма

Задания для выполнения

- 1 Введите в таблицу **MS Excel** исходные данные из Приложения А, таблица А1.
- 2 Выполните подготовительные и расчетные процедуры в соответствии с порядком операций, выполненных в настоящем разделе.
- 3 Получите результат и сделайте заключение.

Лабораторная работа 5

Разнообразие значений признака (Statistica 6)

Цель работы: научиться выполнять первичную обработку данных в программном продукте **Statistica 6**.

Краткие теоретические сведения

Краткие теоретические сведения изложены в соответствующем разделе лабораторной работы 4.

Подготовительные процедуры

Процедуры, связанные с сортировкой массивов данных и поисками выбросов выполняются в табличном редакторе Microsoft Excel.

Запустите программный продукт **Statistica 6**.

Сформируйте таблицу исходных данных: в окне **File** выбрать **New** (рисунок 5.1). После этого щелкните мышкой **OK**.



Рисунок 5.1 – Открытие таблицы

В появившемся окне задайте число строк (**Number of cases**) и столбцов (**Number of variables**) (рисунок 5.2). Щелкните мышкой **OK**.

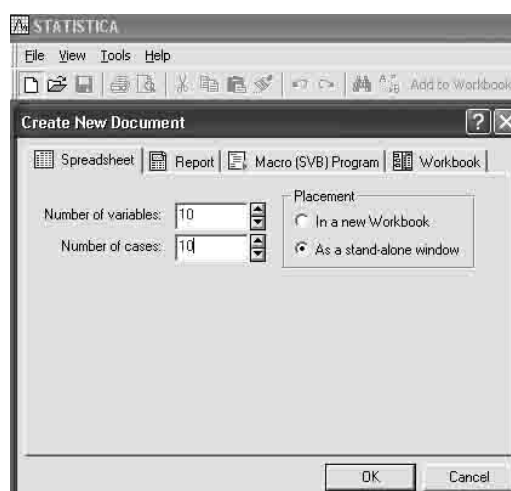


Рисунок 5.2 – Формирование таблицы

Теперь в таблицу необходимо внести исходные данные, осуществив набор непосредственно или вставку копии из файла, например, табличного редактора **MS Excel** (рисунок 5.3).

	1 Var1	2 Var2	3 Var3
1	6	39	20
2	12	49	20
3	12	51	22
4	15	48	25
5	14	102	22
6	6	61	2
7	1	73	1
8	11	63	11
9	12	30	12
10	14	107	19
11	7	37	14
12	1	43	23
13	18	21	10
14	12	44	2
15	12	40	4
16	4	45	13
17	18	44	22
18	8	42	13
19	11	63	11
20	2	33	5
21	12	44	14
22	8	86	36
23	12	49	2
24	16	38	18
25	12	74	1
26	8	17	11
27	16	53	5
28	12	30	9
29	20	73	13
30	11	30	18

Рисунок 5.3 – Исходные данные

Первичный анализ статистических данных

В меню выберите **Statistics** и запустите модуль **Basic Statistics/Tables** (Основные статистики/таблицы). Высветите в стартовой панели модуля **Basic Statistics/Tables** (Основные статистики/таблицы) строку **Descriptive statistics** (Описательная статистика) (рисунок 5.4).

Шаг 1. Нажмите кнопку **OK**. Перед вами откроется окно **Descriptive statistics** (Описательная статистика). Выполните установки, как показано на рисунках 5.5 и 5.6. В окне **Number of intervals** установите необходимое число интервалов для построения гистограмм.

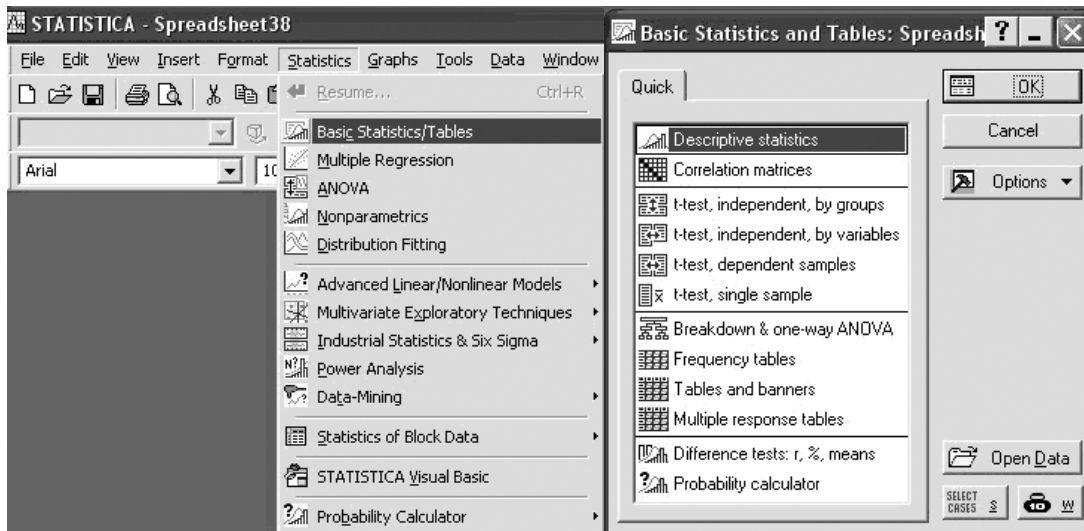


Рисунок 5.4 – Запуск модуля Descriptive statistics (описательная статистика)

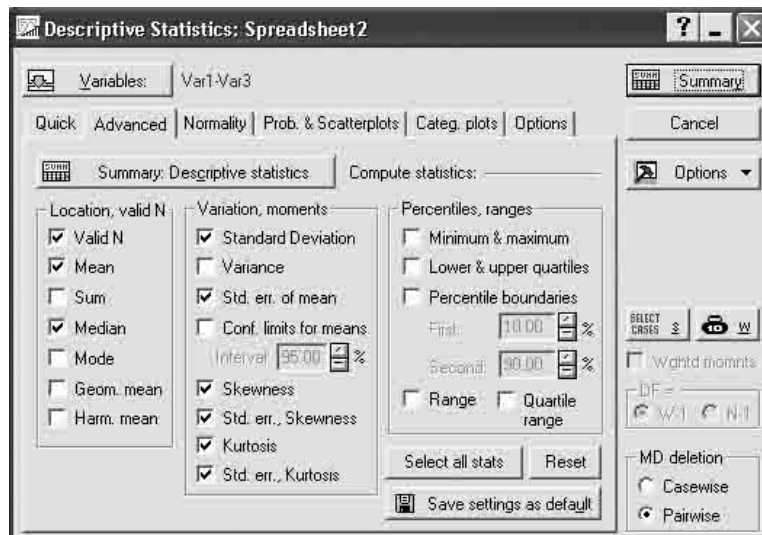


Рисунок 5.5 – Предварительные установки Advanced

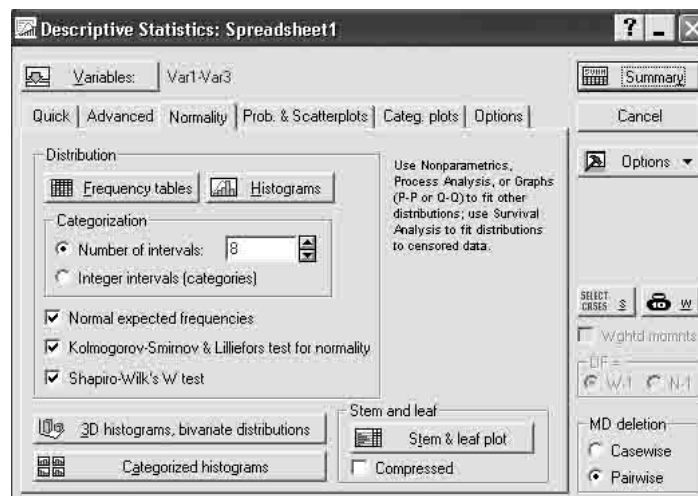


Рисунок 5.6 – Предварительные установки Normality

Шаг 2. Загрузите в систему **Statistica** исходные данные. Для этого левой кнопкой щелкните по клавише **Variables**. В появившемся окне, нажав и не отпуская левую кнопку мышки, пометьте необходимые для последующего анализа столбцы, в данном случае **Var1 – Var3** (рисунок 5.7).

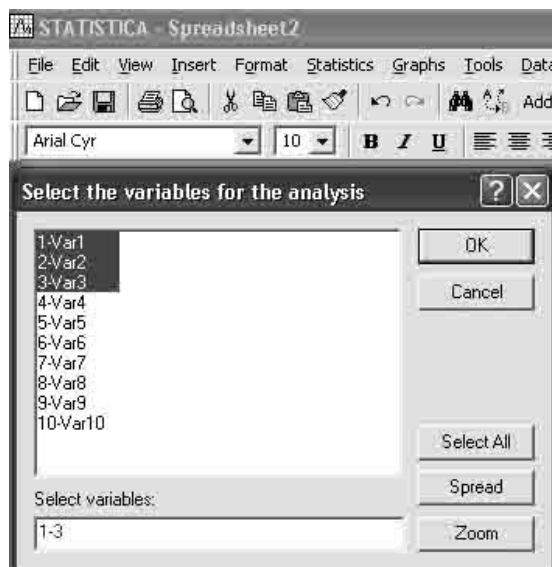


Рисунок 5.7 – Загрузка данных

Шаг 4. В диалоговом окне результатов **Quick** (рисунок 5.8) последовательно нажмите кнопки:

- **Histograms** (Гистограммы);
- **Summary Descriptive statistics** (Описательная статистика);
- **Frequency Tables** (Таблица частот).

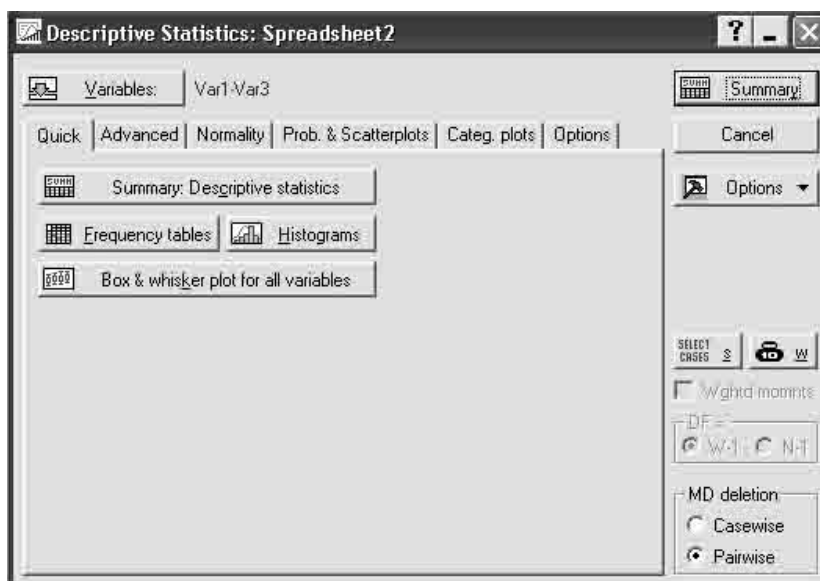


Рисунок 5.8 – Диалоговое окно результатов Quick

Результаты статистической обработки представлены на рисунках 5.9 – 5.11.

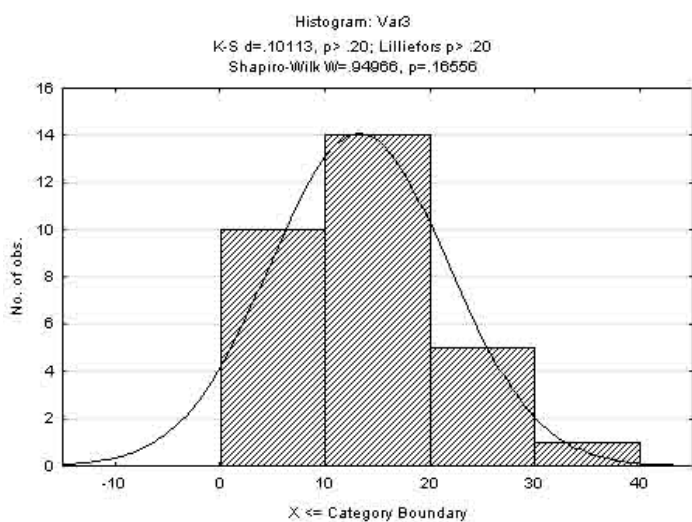
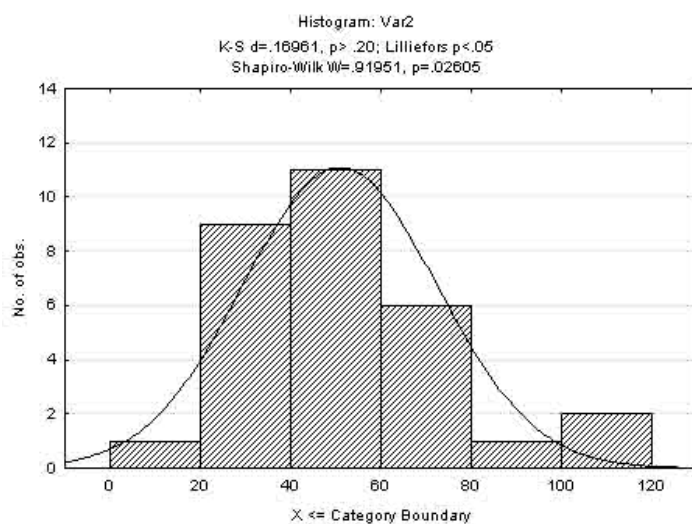
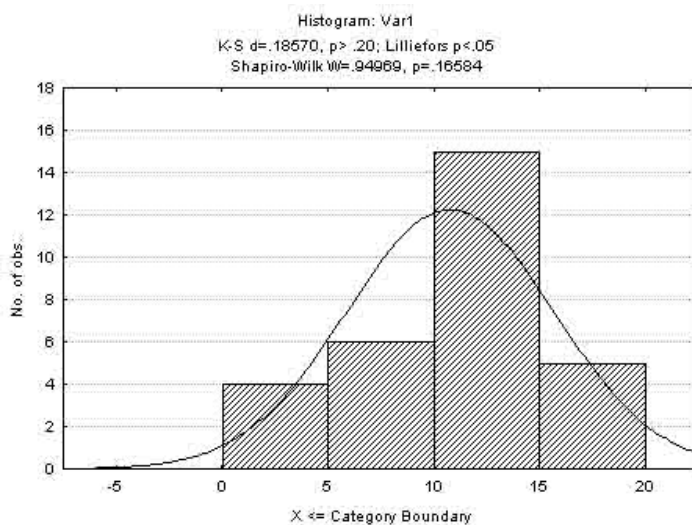


Рисунок 5.9 – Гистограммы

Descriptive Statistics (Spreadsheet8)									
Variable	Descriptive Statistics (Spreadsheet8)								
	Valid N	Mean	Median	Std.Dev.	Standard Error	Skewness	Std.Err. Skewness	Kurtosis	Std.Err. Kurtosis
Var1	30	10.76667	12.00000	4.88970	0.892734	-0.375833	0.426892	-0.189938	0.832746
Var2	30	50.96667	44.50000	21.59898	3.943417	1.026556	0.426892	0.929974	0.832746
Var3	30	13.26667	13.00000	8.50531	1.552849	0.434693	0.426892	0.144419	0.832746

Рисунок 5.10 – Описательная статистика

Frequency table: Var1 (Spreadsheet2)
K-S d= 13123, p> .20; Lilliefors p> .20

Category	Count	Cumulative Count	Percent of Valid	Cumul % of Valid	% of all Cases	Cumulative % of All	Expected Count	Cumulative Expected	Percent Expected	Cumulative % Expected
12.00000<x<=13.00000	1	1	11.11111	11.1111	10.00000	10.0000	0.907640	0.907640	10.08489	10.08489
13.00000<x<=14.00000	2	3	22.22222	33.3333	20.00000	30.0000	1.901904	2.809544	21.13226	31.21716
14.00000<x<=15.00000	3	6	33.33333	66.6667	30.00000	60.0000	2.742458	5.552002	30.47175	61.68891
15.00000<x<=16.00000	1	7	11.11111	77.7778	10.00000	70.0000	2.196046	7.748047	24.40051	86.08941
16.00000<x<=17.00000	2	9	22.22222	100.0000	20.00000	90.0000	0.976131	8.724178	10.84590	96.93531
Missing	1	10	11.11111		10.00000	100.0000				

Frequency table: Var2 (Spreadsheet2)
K-S d= 14039, p> .20; Lilliefors p> .20

Category	Count	Cumulative Count	Percent of Valid	Cumul % of Valid	% of all Cases	Cumulative % of All	Expected Count	Cumulative Expected	Percent Expected	Cumulative % Expected
14.00000<x<=16.00000	2	2	20.00000	20.0000	20.00000	20.0000	1.390803	1.390803	13.90803	13.90803
16.00000<x<=18.00000	1	3	10.00000	30.0000	10.00000	30.0000	1.711916	3.102719	17.11916	31.02719
18.00000<x<=20.00000	2	5	20.00000	50.0000	20.00000	50.0000	2.272930	5.375649	22.72930	53.75649
20.00000<x<=22.00000	3	8	30.00000	80.0000	30.00000	80.0000	2.153470	7.529119	21.53470	75.29119
22.00000<x<=24.00000	1	9	10.00000	90.0000	10.00000	90.0000	1.455902	8.985020	14.55902	89.85020
24.00000<x<=26.00000	1	10	10.00000	100.0000	10.00000	100.0000	0.702276	9.687297	7.02276	96.87297

Frequency table: Var3 (Spreadsheet2)
K-S d= 19481, p> .20; Lilliefors p> .20

Category	Count	Cumulative Count	Percent of Valid	Cumul % of Valid	% of all Cases	Cumulative % of All	Expected Count	Cumulative Expected	Percent Expected	Cumulative % Expected
3.000000<x<=3.500000	1	1	14.28571	14.2857	10.00000	10.0000	0.345756	0.345756	4.93937	4.93937
3.500000<x<=4.000000	0	1	0.00000	14.2857	0.00000	10.0000	1.160817	1.506573	16.58309	21.52247
4.000000<x<=4.500000	3	4	42.85714	57.1429	30.00000	40.0000	2.199653	3.706226	31.42362	52.94608
4.500000<x<=5.000000	2	6	28.57143	85.7143	20.00000	60.0000	2.071772	5.777998	29.59675	82.54283
5.000000<x<=5.500000	1	7	14.28571	100.0000	10.00000	70.0000	0.969709	6.747707	13.85299	96.39582
Missing	3	10	42.85714		30.00000	100.0000				

Рисунок 5.11 – Таблицы частот

Из представленной обработки следует, что все массивы данных описываются функцией нормального распределения, так как отношения эксцессов и асимметрий к их ошибкам меньше трех (рисунок 5.10), по критерию Колмогорова-Смирнова $p > 0,2$ (рисунок 5.9) и по критерию Шапиро-Уилка $p > 0,01$. Дополнительно информацию можно извлечь из анализа результатов обработки, представленных в таблицах (рисунки 5.10, 5.11). Нормальность

распределения также можно проверить по критерию χ^2 -квadrat.

Шаг 1. Для этого в меню выберите **Statistics** и запустите модуль **Distribution Fitting** (Подгонка распределения) (рисунок 5.12).

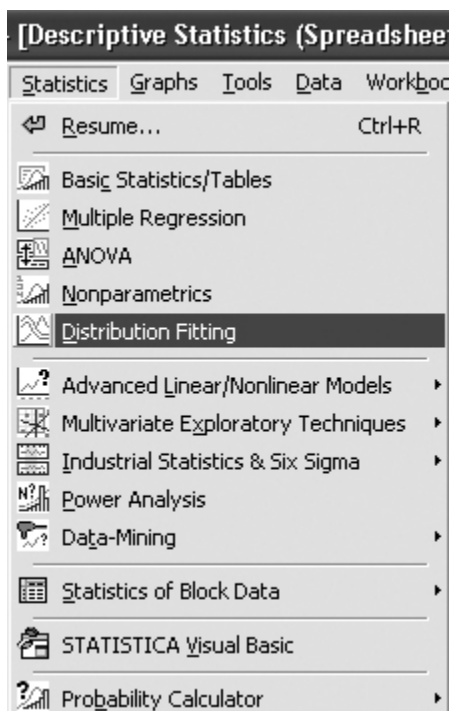


Рисунок 5.12 – Запуск модуля Distribution Fitting (Подгонка распределения)

Шаг 2. В появившемся окне выполнить установки, как показано на рисунке 5.13. Нажмите кнопку **ОК**.

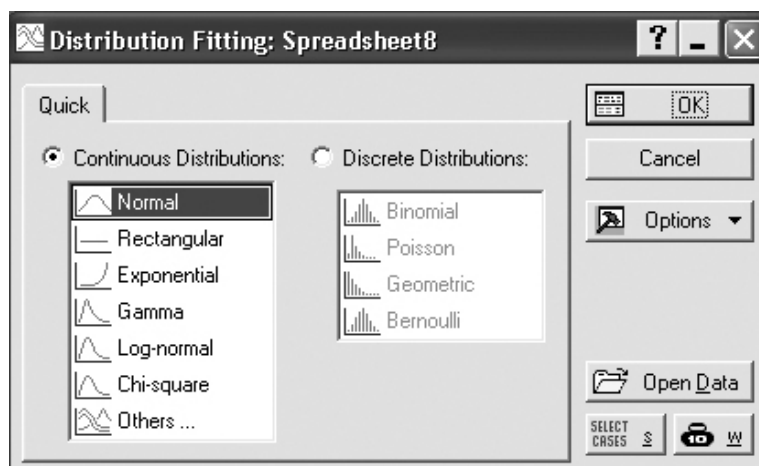


Рисунок 5.13 – Окно модуля Distribution Fitting (подгонка распределения)

Шаг 3. Перед вами откроется окно **Fitting Continuous Distribution** (Подгонка непрерывного распределения). Выполните установки, как

показано на рисунке 5.14.

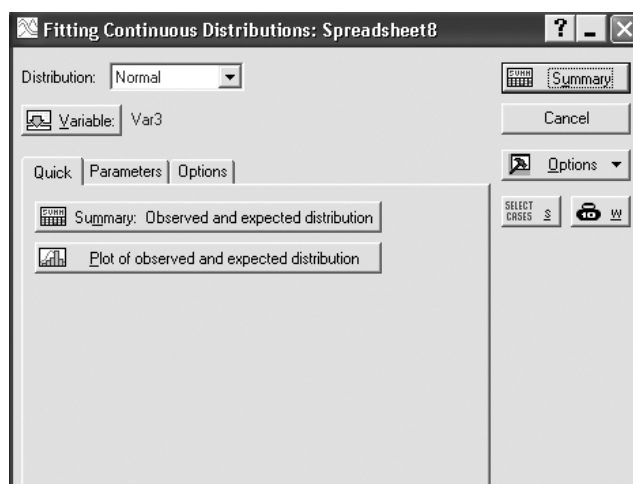


Рисунок 5.14 – Диалоговое окно результатов **Quick**

Шаг 4. Нажмите клавишу **Plot of observed and expected distribution** (График наблюдаемого и ожидаемого распределения).

Результат представлен на рисунке 5.15. Аналогично выполните обработку остальных массивов. Если $p > 0,01$, то массив данных удовлетворяет функции нормального распределения.

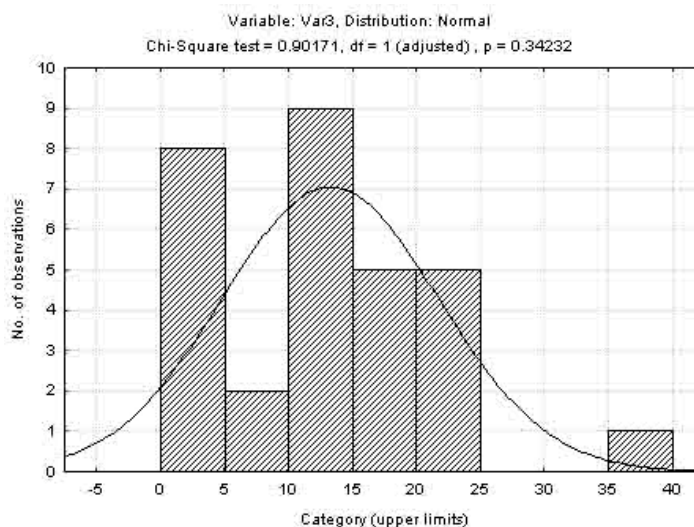


Рисунок 5.15 – Результат обработки

Задания для выполнения

1 Введите в таблицу **Statistica** исходные данные из Приложения А (таблица А1), предварительно отредактировав и проверив их на наличие артефактов в табличном редакторе **MS Excel**.

2 Выполните подготовительные и расчетные процедуры в соответствии с порядком операций, выполненных в настоящем разделе.

3 Получите результат и сделайте заключение.

Лабораторная работа 6

Парная корреляция (MS Excel)

Цель работы: научиться выполнять корреляционный и регрессионный анализ, используя табличный редактор MS Excel.

Краткие теоретические сведения

Корреляция представляет собой меру зависимости переменных. Наиболее известна корреляция Пирсона. При вычислении корреляции Пирсона предполагается, что переменные измерены, как минимум, в интервальной шкале. Некоторые другие коэффициенты корреляции могут быть вычислены для менее информативных шкал. Коэффициенты корреляции изменяются в пределах от $-1,00$ до $+1,00$. Значение $-1,00$ означает, что переменные имеют строгую отрицательную корреляцию. Значение $+1,00$ означает, что переменные имеют строгую положительную корреляцию. Отметим, что значение $0,00$ означает отсутствие корреляции.

Корреляция Пирсона предполагает, что две рассматриваемые переменные измерены, по крайней мере, в интервальной шкале. Она определяет степень, с которой значения двух переменных «пропорциональны» друг другу. Важно, что значение коэффициента корреляции не зависит от масштаба измерения. Например, корреляция между ростом и массой будет одной и той же, независимо от того, проводились измерения в дюймах и фунтах или в сантиметрах и килограммах. Пропорциональность означает просто линейную зависимость. Корреляция высокая, если на графике зависимость можно представить прямой линией (с положительным или отрицательным углом наклона).

Проведенная прямая называется прямой регрессии или прямой, построенной методом наименьших квадратов. Последний термин связан с тем, что сумма квадратов расстояний (вычисленных по оси Y) от наблюдаемых точек до прямой является минимальной. Заметим, что использование квадратов расстояний приводит к тому, что оценки параметров прямой сильно реагируют на выбросы.

Коэффициент корреляции Пирсона (r) представляет собой меру линейной зависимости двух переменных. Если возвести его в квадрат, то полученное значение коэффициента детерминации (r^2) представляет долю вариации, общую для двух переменных (иными словами, степень

зависимости или связанности двух переменных). Чтобы оценить зависимость между переменными, нужно знать как величину корреляции, так и ее значимость.

Уровень значимости, вычисленный для каждой корреляции, представляет собой главный источник информации о надежности корреляции. Значимость определенного коэффициента корреляции зависит от объема выборок. Критерий значимости основывается на предположении, что распределение остатков (т. е. отклонений наблюдений от регрессионной прямой) для зависимой переменной Y является нормальным (с постоянной дисперсией для всех значений независимой переменной X).

Несколько слов о выбросах. По определению, выбросы являются нетипичными, резко выделяющимися наблюдениями. Так как при построении прямой регрессии используется сумма квадратов расстояний наблюдаемых точек до прямой, то выбросы могут существенно повлиять на наклон прямой и, следовательно, на значение коэффициента корреляции. Поэтому единичный выброс (значение которого возводится в квадрат) способен существенно изменить наклон прямой и, следовательно, значение корреляции.

Обычно считается, что выбросы представляют собой случайную ошибку, которую следует контролировать. К сожалению, не существует общепринятого метода автоматического удаления выбросов. Чтобы не быть введенными в заблуждение полученными значениями, необходимо проверить на диаграмме рассеяния каждый важный случай значимой корреляции. Очевидно, выбросы могут не только искусственно увеличить значение коэффициента корреляции, но также реально уменьшить существующую корреляцию.

Если после выполнения процедуры поиска и удаления выбросов коэффициент корреляции остался низким и уровень значимости меньше критического 0,05, то следует увеличить размер выборки,

воспользовавшись простым соотношением: $n = \frac{t_{st}^2}{r^2} \approx \frac{4}{r^2}$.

Корреляционный анализ

Выполнение корреляционного анализа покажем на примере результатов измерения массы (Y) и длины туловища (X) подопытных животных. Исходные данные представлены в таблице 6.1 (в табличном редакторе **MS Excel** данные представлены двумя

столбцами).

Шаг 1. Откройте модуль **Анализ данных**, выберите опцию **Корреляция**, после чего щелкните мышкой **ОК**.

Шаг 2. В появившемся окне выполните операции и установки, как показано на рисунке 6.1.

Шаг 3. Щелкните мышкой **ОК**. Результат обработки появится в указанном поле (выходной интервал \$E\$1, таблица 6.2).

Таблица 6.1 – Исходные данные

X	Y	X	Y	X	Y
3,4	14,3	8,4	19,8	10,7	21,3
3,6	14,9	8,5	19,9	11,6	21,3
4,5	17,3	8,8	19,9	12	21,8
4,8	17,3	8,9	20,1	12,3	22
4,9	17,4	8,9	20,1	12,6	22,1
5,2	17,5	8,9	20,1	12,7	22,4
5,4	17,6	8,9	20,1	13,3	22,7
5,7	17,6	9	20,2	13,6	23,5
6,2	17,6	9	20,3	13,8	24,2
6,7	17,8	9,1	20,3	14	24,4
7,1	18	9,3	20,5	15	25,2
7,5	18	9,4	20,6	15,2	25,2
7,7	18,1	9,7	20,9	15,8	25,3
7,8	18,1	9,7	21	15,9	25,7
7,9	18,6	9,9	21,1	16,6	26,8
8	19,7	10,1	21,1	17,1	27,5

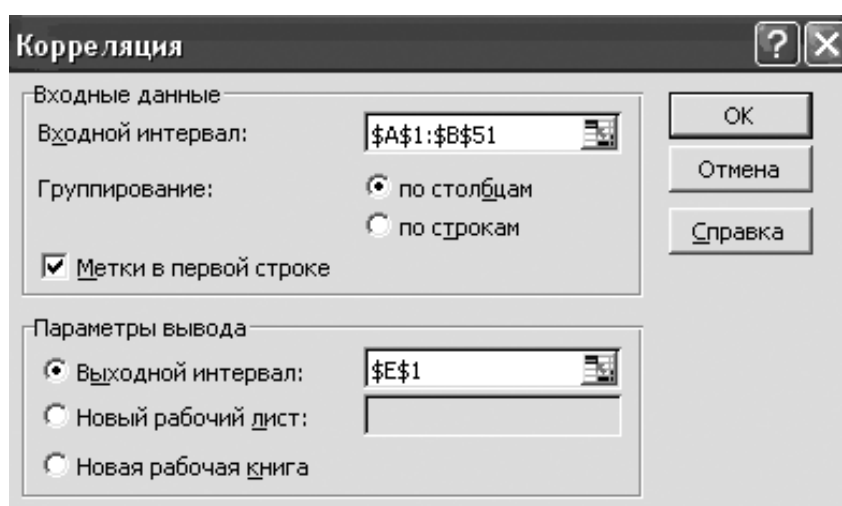


Рисунок 6.1 – Стартовая панель

Таблица 6.2 – Результат обработки

	X	Y
X	1	0,98

Y	0,98	1
---	------	---

Регрессионный анализ

Для выполнения регрессионного анализа использовались исходные данные таблицы 6.1.

Шаг 1. Откройте модуль **Анализ данных** и выберите опцию **Регрессия**, после чего щелкните мышкой **ОК**.

Шаг 2. В появившемся окне выполните операции и установки, как показано на рисунке 6.2.

Шаг 3. Щелкните мышкой **ОК**. Результат обработки появится в указанном поле (выходной интервал \$L\$1, рисунок 6.2, таблицы 6.3–6.5).

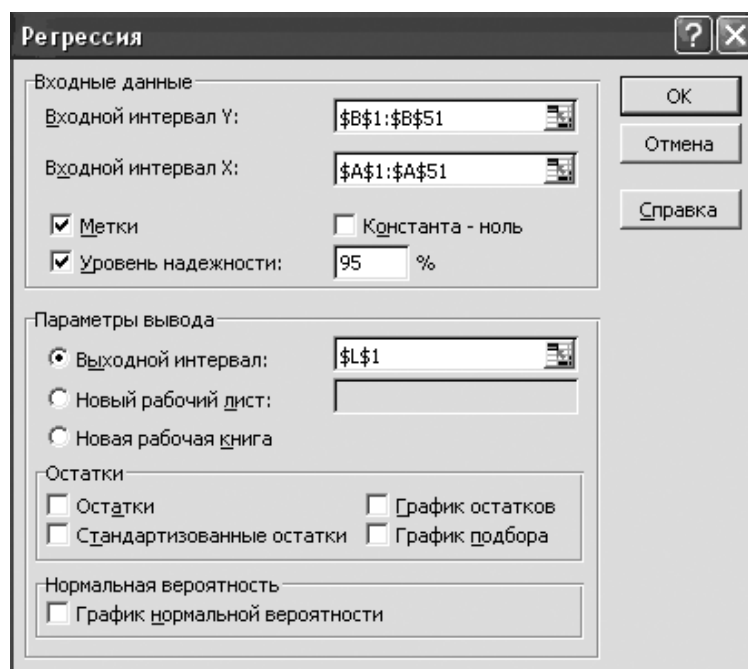


Рисунок 6.2 – Стартовая панель

Таблица 6.3 – Результат обработки

Параметры	Значения
Множественный R	0,98
R-квадрат	0,97
Нормированный R-квадрат	0,97
Стандартная ошибка	0,53
Наблюдения	50

Таблица 6.4 – Дисперсионный анализ

Параметры	df	SS	MS	F	Значимость F
Регрессия	1	397,27	397,27	1403,61	0,000..
Остаток	48	13,586	0,28		

Итого	49	410,86			
-------	----	--------	--	--	--

Таблица 6.5 – Регрессионный анализ

Параметры	Коэффициенты	Стандартная ошибка	t-статистика	P-значение
Y-пересечение	12,67	0,224	56,628	0,000..
X	0,82	0,0218	37,465	0,000..

Таким образом, корреляционная связь между массой и длиной туловища подопытных животных характеризуется высоким ($r = 0,98$) и достоверным коэффициентом корреляции (из таблицы 6.4: критерий Фишера $F = 1403,61$ при уровне значимости существенно меньше $0,05$). Получена очень надежная регрессия, о чем свидетельствует t-статистика из таблицы 6.5 (уровень значимости существенно меньше $0,05$).

Уравнение линейной регрессии Пирсона и коэффициент корреляции:
 $Y = 12,67 + 0,816 \cdot X$; $r = 0,98$.

С менее подробной информацией операции регрессии и корреляции можно выполнить в системе MS Excel, используя модуль Мастер диаграмм.

Шаг 1. В системе MS Excel откройте модуль Мастер диаграмм (рисунок 6.3).

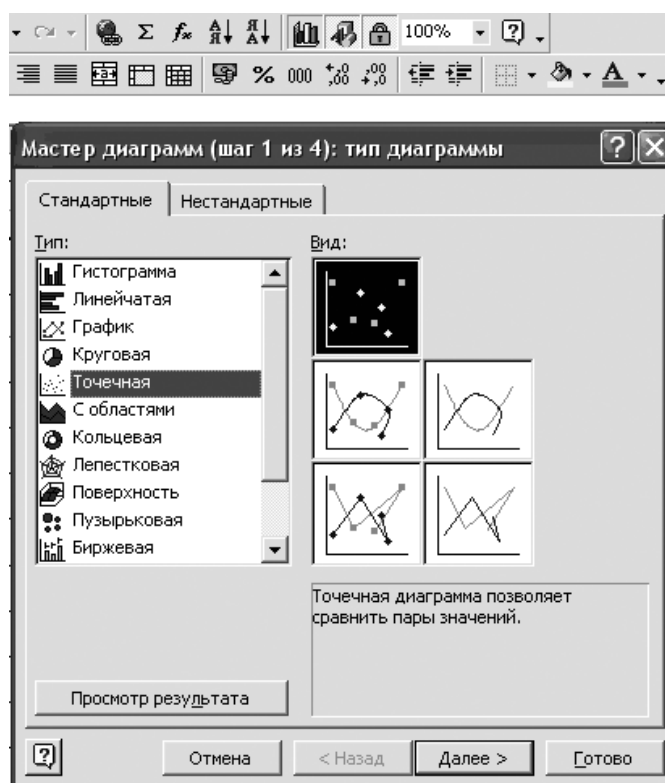


Рисунок 6.3 – Стартовая панель

Шаг 2. Выберите Тип диаграммы **Точечная**.

Шаг 3. Щелкните по кнопке **Далее**. Выберите диапазон данных (рисунок 6.4), оформите график и нажмите **Готово**.

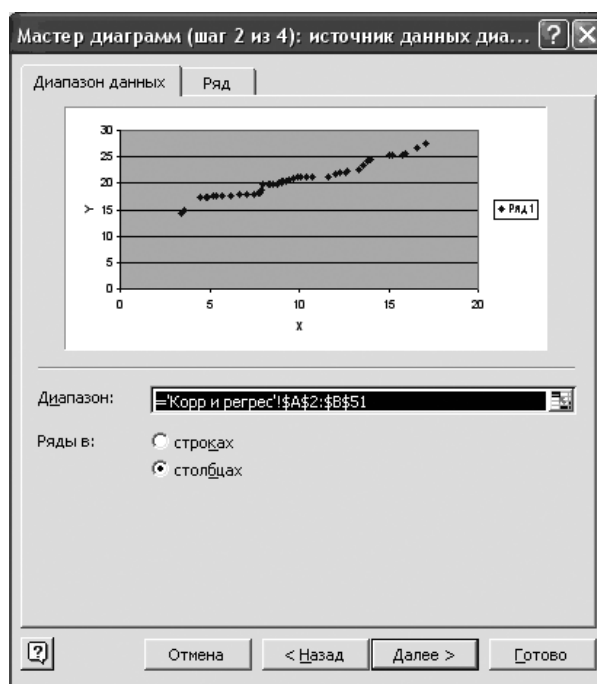


Рисунок 6.4 – Диапазон данных

Шаг 4. На графике щелкните правой кнопкой по любой точке диаграммы.

Шаг 5. Выберите опцию **Добавить линию тренда** и **Тип – линейная** (рисунок 6.5). Щелкните мышкой **ОК**.

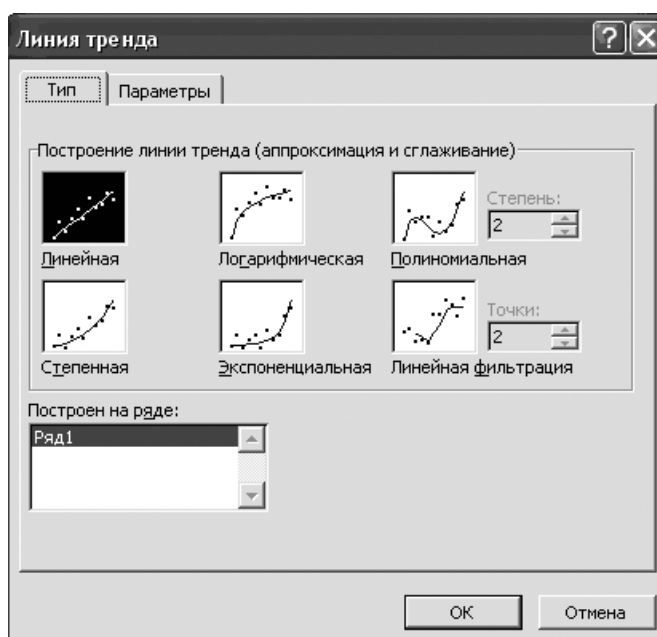


Рисунок 6.5 – Тип аппроксимации кривой

Шаг 6. В опции **Параметры** выберите установки, как показано на рисунке 6.6. Щелкните мышкой **ОК**.

Отредактированная диаграмма представлена на рисунке 6.7.

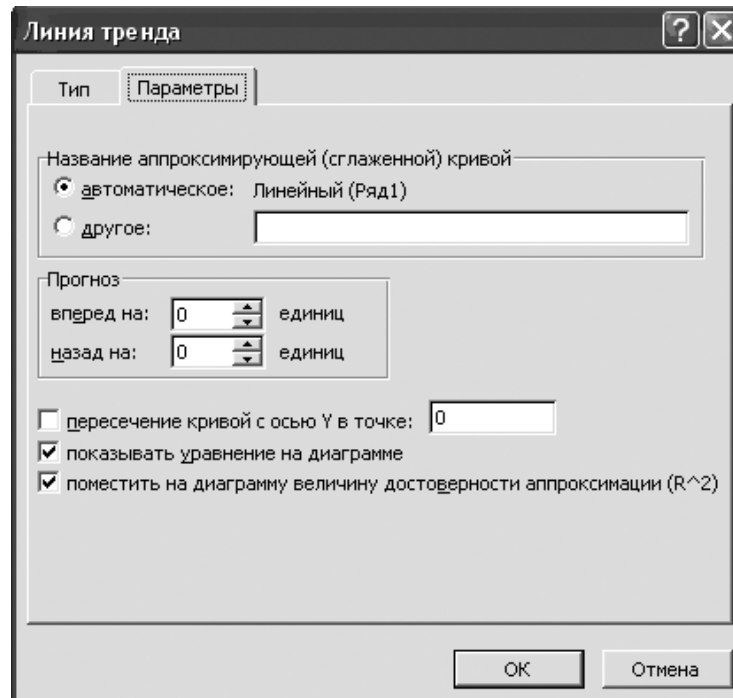


Рисунок 6.6 – Параметры кривой

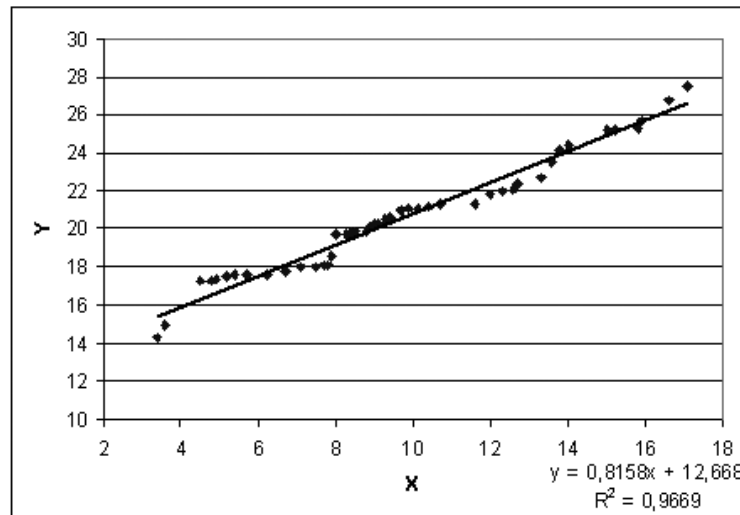


Рисунок 6.7 – Отредактированная диаграмма

Уравнение регрессии и коэффициент детерминации R^2 находятся в правом нижнем углу диаграммы. Как видно, что они такие же, как и при выполнении регрессионного анализа в пакете **Анализ данных Регрессия**.

Задания для выполнения

1 Введите в таблицу **MS Excel** исходные данные из Приложения Б (таблица Б1).

2 Отредактируйте и проверьте их на наличие артефактов в табличном редакторе **MS Excel**.

3 Выполните подготовительные и расчетные процедуры в соответствии с порядком операций, выполненных в настоящем разделе.

4 Получите результат и сделайте заключение.

Лабораторная работа 7

Парная корреляция (Statistica 6)

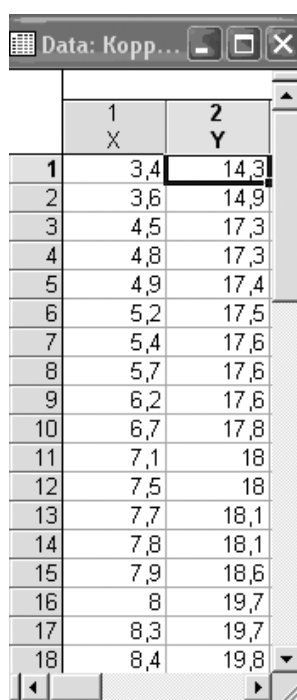
Цель работы: научиться выполнять корреляционный и регрессионный анализ, используя программный продукт **Statistica 6**.

Краткие теоретические сведения

Краткие теоретические сведения изложены в соответствующем разделе лабораторной работы 6.

Корреляционный и регрессионный анализ

Массив исходных данных приведен в таблице 6.1, который из табличного редактора **MS Excel** перемещен в таблицу модуля **Statistica 6** (рисунок 7.1).



	1 X	2 Y
1	3,4	14,3
2	3,6	14,9
3	4,5	17,3
4	4,8	17,3
5	4,9	17,4
6	5,2	17,5
7	5,4	17,6
8	5,7	17,6
9	6,2	17,6
10	6,7	17,8
11	7,1	18
12	7,5	18
13	7,7	18,1
14	7,8	18,1
15	7,9	18,6
16	8	19,7
17	8,3	19,7
18	8,4	19,8

Рисунок 7.1 – Исходные данные

X – независимая переменная

Y – зависимая переменная

Проведем анализ в модуле **Basic statistics/Tables** (Основная статистика). Рассмотрим и установим связь между X и Y .

Шаг 1. Из Переключателя модулей **Statistica** откройте модуль **Basic statistics/Tables** (Основная статистика). Высветите название модуля и далее щелкните мышью по названию модуля: **Basic statistics/Tables**

(рисунок 7.2).

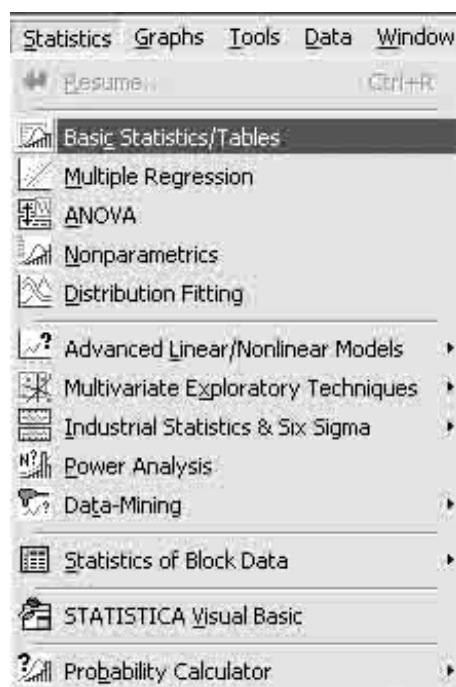


Рисунок 7.2 – Стартовая панель модуля Basic statistics/Tables

Шаг 2. На экране появится (рисунок 7.3). Щелкните мышью по названию **Correlation matrices** (Корреляционная матрица).

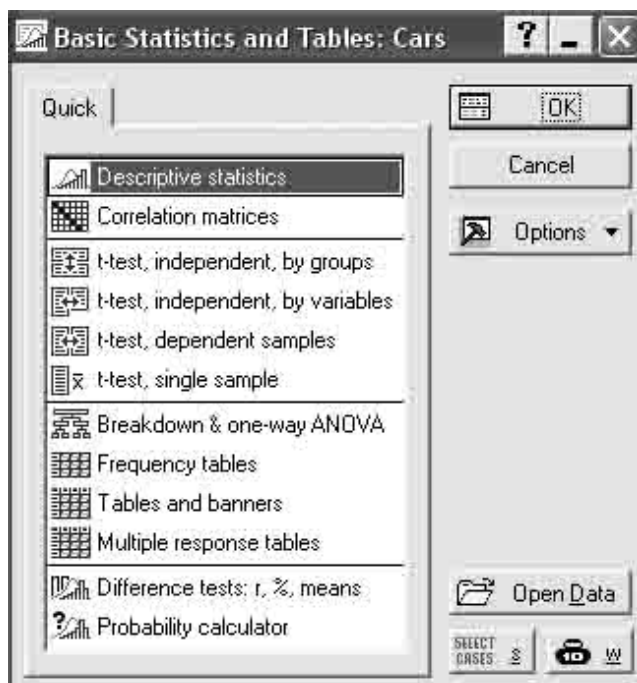


Рисунок 7.3 – Панель запуска

Шаг 3. Выберите переменные для анализа. Выбор переменных осуществляется с помощью кнопки **Two list**, находящейся в центре

верхней части панели (рисунок 7.4).

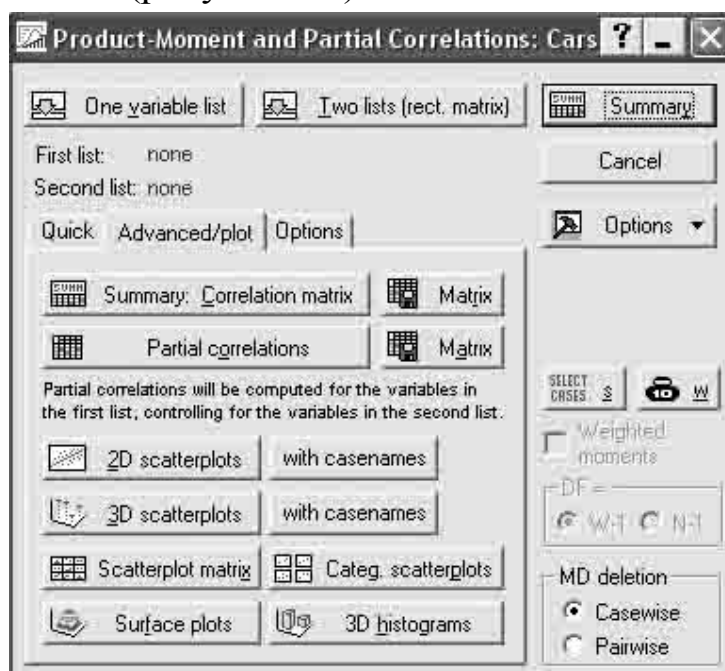


Рисунок 7.4 – Стартовая панель модуля Correlation matrices

После того как кнопка будет нажата, диалоговое окно **Select one or two variable list (выбрать списки зависимых и независимых переменных)** появится на вашем экране (рисунок 7.5).

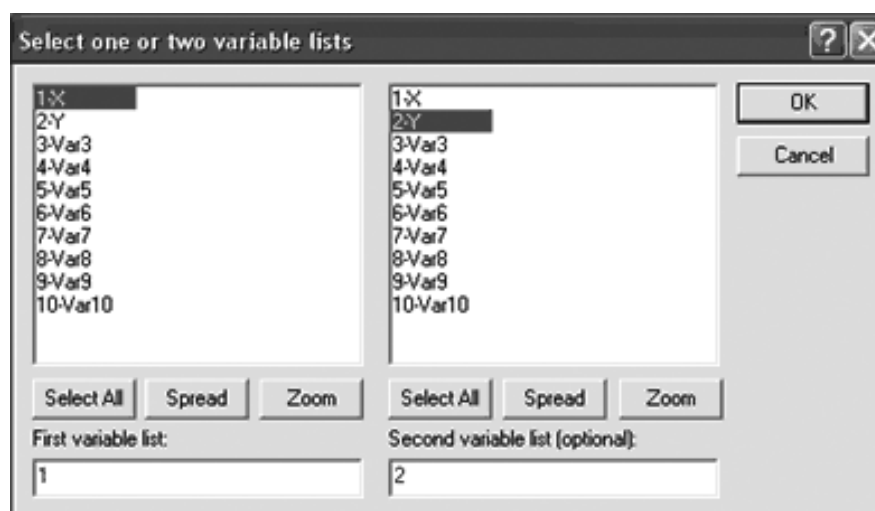


Рисунок 7.5 – Окно выбора переменных для анализа

Шаг 4. Высветив имя переменной в правой части окна, выберите переменную в левой части окна.

После нажатия кнопки **OK** выполните установки, показанные на рисунке 7.6, подсветив **Display detailed table of results**.

Шаг 5. После нажатия кнопки **Summary** программа произведет

расчет корреляции между X и Y , и на экране появится окно результатов (рисунок 7.7).

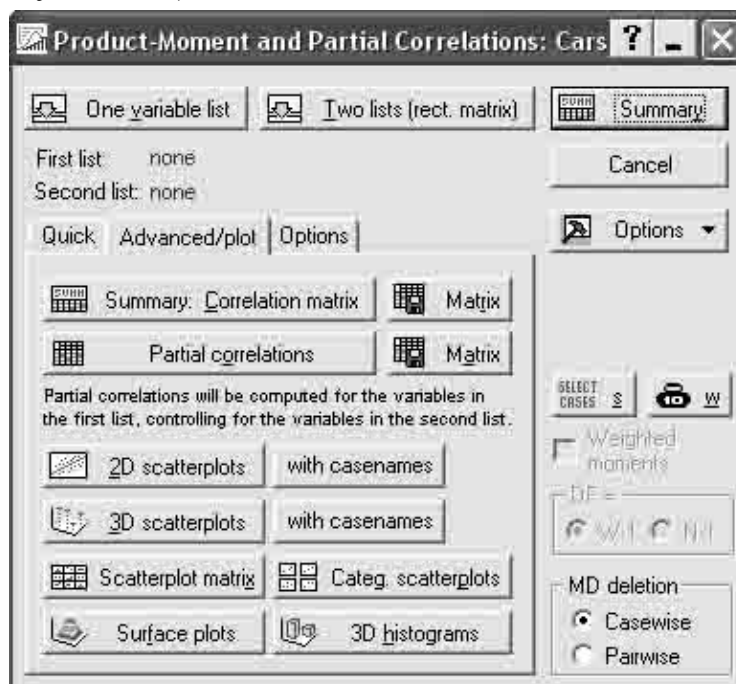


Рисунок 7.6 – Окно предварительных установок

Var. X & Var. Y	Mean	Std. Dev.	r(X,Y)	r ²	t	p	N	Constant dep. Y	Slope dep. Y	Constant dep. X	Slope dep. X
X	9,67600	3,490311									
Y	20,56200	2,895661	0,983328	0,966933	37,46471	0,00	50	12,66835	0,815797	-14,6954	1,185263

Рисунок 7.7 – Результат расчета корреляции

На рисунке 7.7 представлена следующая информация:

- среднее;
- стандартное отклонение;
- значение коэффициента корреляции r ;
- значение коэффициента детерминации r^2 ;
- t – критерий;
- p – уровень значимости;
- число коррелируемых пар;
- 12,67 – свободный член уравнения регрессии.
- 0,82 – коэффициент при независимой переменной уравнения регрессии.

В этом примере $r = 0,98$. Это очень высокое значение (подсвечено красным цветом), показывающее, что построенная регрессия объясняет более 90% разброса значений переменной X относительно среднего.

Из таблицы видно, что оцененная модель имеет вид:

$$Y = 12,67 + 0,816 \cdot X$$

Шаг 6. После нажатия кнопки **2D scatterplots** появится график, на котором данные с подогнутой прямой имеют вид (рисунок 7.8).

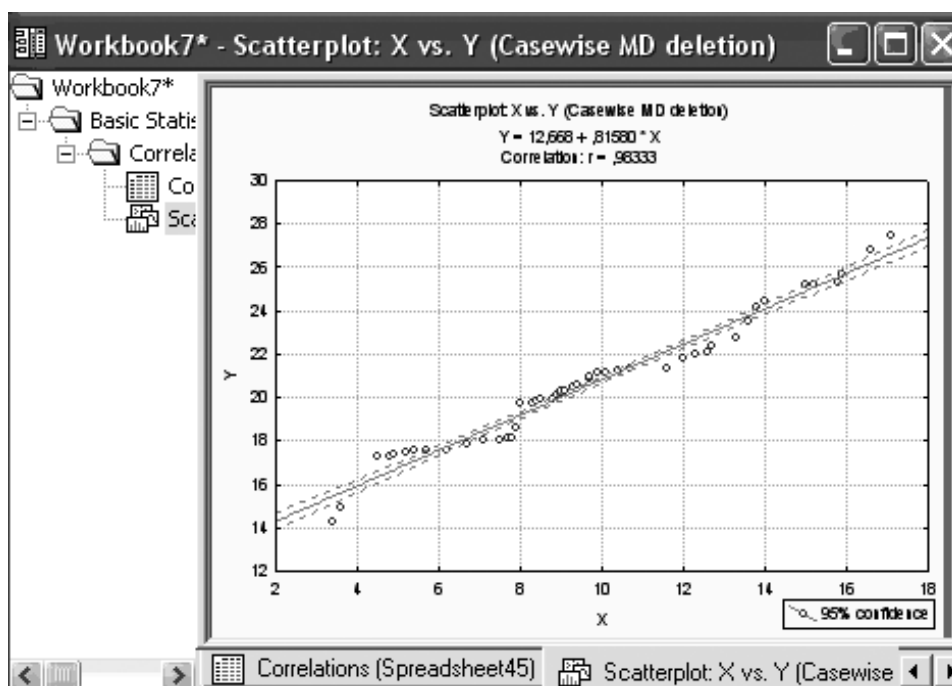


Рисунок 7.8 – Линейная регрессия для данных X и Y

Задания для выполнения

1 Введите в таблицу **Statistica 6** исходные данные из Приложения Б (таблица Б1), предварительно отредактировав и проверив их на наличие артефактов в табличном редакторе **MS Excel**.

2 Выполните подготовительные и расчетные процедуры в соответствии с порядком операций, выполненных в настоящем разделе.

3 Получите результат и сделайте заключение.

Лабораторная работа 8

Частная и множественная линейные корреляции и регрессия (Statistica 6)

Цель работы: научиться выполнять множественный корреляционный и регрессионный анализ в программном продукте Statistica 6.

Краткие теоретические сведения

Типичной практической задачей является задача определения зависимостей в системе данных.

Предположим, вы наблюдаете значения пары переменных X и Y и хотите найти зависимость между ними.

Переменная X носит название **независимой переменной**, или предиктора, переменная Y называется **зависимой переменной**, или **откликом**.

Данная терминология связана с тем, что необходимо определить именно зависимость Y от X или предсказать, какими будут значения Y при данных значениях X .

Значение переменной X в i -м опыте будем обозначать через $X(i)$, соответствующее значение величины Y обозначим через $Y(i)$, $0 < i \leq n$.

Вы наблюдаете значения независимой $X(i)$ и соответствующие им значения зависимой $Y(i)$, и хотите оценить зависимость Y от X . В статистике подобные задачи решаются в рамках регрессионной модели. Будем рассматривать самую простую регрессионную модель – линейную. Однако и в рамках этой модели могут быть решены самые разнообразные практические задачи.

Регрессионный анализ в системе **Statistica 6** проводится в модуле **Multiple regression** (Множественная регрессия).

Описание модели

Дадим точное описание линейной регрессионной модели, в рамках которой будем исследовать зависимость Y от X .

При этом постулируется, что наблюдаемые величины связаны между собой регрессионной зависимостью вида:

$$Y(i) = B_1 \cdot X(i) + B_0 + e(i), \text{ при } 0 < i \leq n$$

где B_1 , B_0 – неизвестные константы, $e(i)$ – ненаблюдаемые случайные величины (наблюдаются только $X(i)$, $Y(i)$, $0 < i \leq n$) со

средним равным нулю (как говорят, являются несмещенными) и неизвестной дисперсией, не меняющейся от опыта к опыту.

Иногда случайные величины $e(i)$, $0 < i \leq n$ называют ошибками наблюдения. Относительно $e(i)$ предполагается, что они не коррелированы в разных опытах. Кроме того, часто предполагается, что ошибки имеют нормальное распределение. В этом случае некоррелированность влечет независимость.

Можно рассматривать и более общие линейные модели, например, с несколькими независимыми переменными:

$$Y(i) = B_1 \cdot X_1(i) + B_2 \cdot X_2(i) + \dots + B_k \cdot X_k(i) + B_0 + e(i), \text{ при } 0 < i \leq n,$$

где $B_0, B_1, B_2, \dots, B_k$ – неизвестные коэффициенты.

Эта модель также может исследоваться в модуле **Multiple regression** (Множественная регрессия).

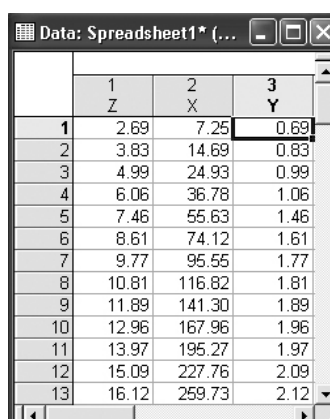
Постановка задачи

Общая задача состоит в том, чтобы по наблюдениям $(X_1, Y_2), \dots, (X_n, Y_n)$:

- оценить параметры модели B_1, B_2, B_0 наилучшим образом;
- построить доверительные интервалы для B_1, B_2, B_0 ;
- проверить гипотезу о значимости регрессии;
- оценить степень адекватности модели.

Ниже на примере выполненных замеров растений определенного вида, где Z – масса растения, X – высота и Y – диаметр ствола, показано, как решается данная задача в системе **Statistica 6**.

Введите исходные данные, как показано на рисунке 8.1.



	1 Z	2 X	3 Y
1	2.69	7.25	0.69
2	3.83	14.69	0.83
3	4.99	24.93	0.99
4	6.06	36.78	1.06
5	7.46	55.63	1.46
6	8.61	74.12	1.61
7	9.77	95.55	1.77
8	10.81	116.82	1.81
9	11.89	141.30	1.89
10	12.96	167.96	1.96
11	13.97	195.27	1.97
12	15.09	227.76	2.09
13	16.12	259.73	2.12

Рисунок 8.1 – Исходные данные

X и Y – независимая переменная; Z – зависимая переменная

Проведем анализ в модуле **Multiple regression** (Множественная

регрессия).

Рассмотрим и установим связь между X , Y и Z .

Шаг 1. Из Переключателя модулей **Statistica** откройте модуль **Multiple regression** (Множественная регрессия). Высветите название модуля и далее щелкните мышью по названию модуля **Multiple regression** (рисунок 8.2).

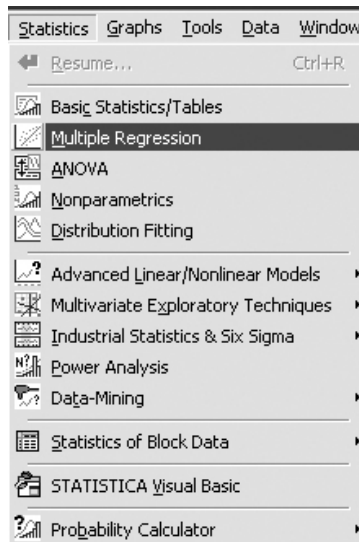


Рисунок 8.2 – Пусковая панель

Шаг 2. На экране появится стартовая панель модуля (рисунок 8.3).

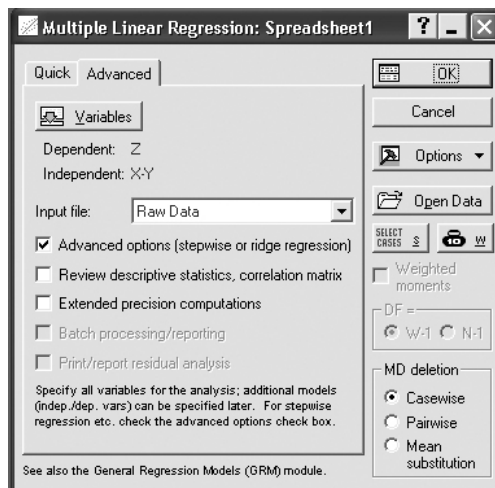


Рисунок 8.3 – Стартовая панель модуля Multiple regression

Выполните установки, как показано на рисунке 8.3. Далее выберите переменные для анализа. Выбор переменных осуществляется с помощью кнопки **Variables** (Переменные), находящейся в левом верхнем углу панели.

После того как кнопка будет нажата, диалоговое окно **Select dependent and independent variable list** (Выбрать списки зависимых и

независимых переменных) появится на вашем экране (рисунок 8.4).

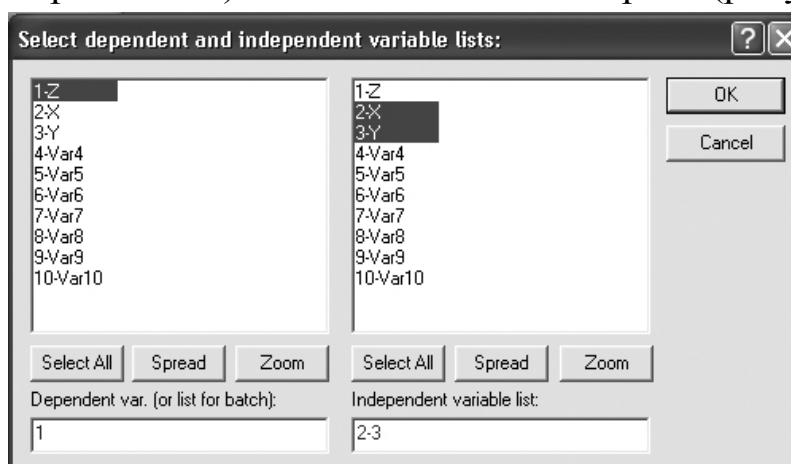


Рисунок 8.4 – Окно выбора переменных для анализа

Высветив имя переменной в левой части окна, выберите зависимые переменные. Высветив имя переменной в правой части окна, выберите независимые переменные.

В данном примере независимыми переменными являются X и Y , зависимой – Z . Высветив имена этих переменных, как показано на рисунке, нажмите кнопку **OK** в правом верхнем углу окна **Select dependent and independent variable list** (Выберите зависимые и независимые переменные). Вы вновь окажетесь в стартовой панели модуля. Переменные для анализа выбраны.

Нажмите кнопку **OK** в правом углу стартовой панели.

Шаг 3. На экране перед вами появится диалоговое окно **Model Definition** (Построение модели) (рисунок 8.5).

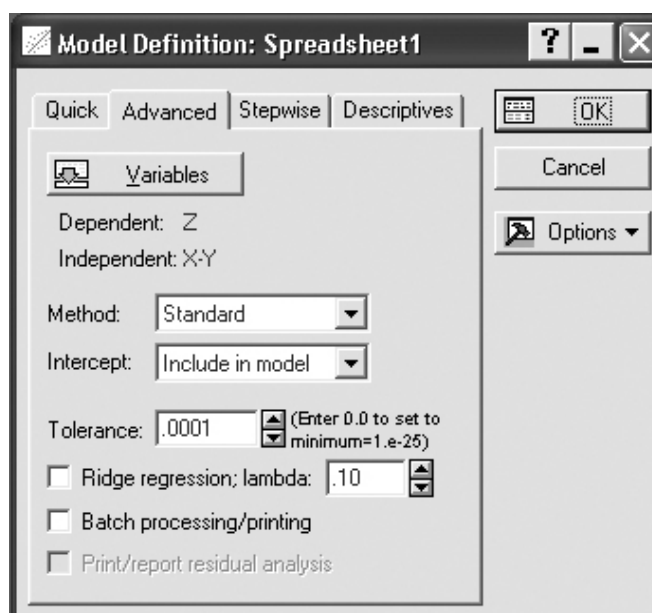


Рисунок 8.5 – Окно построения модели в модуле Multiple Regression

В данном окне выберите стандартный метод оценивания, в опции **Method (Метод): Standard (Стандартный)**.

Далее нажмите кнопку **ОК**.

Программа произведет оценивание параметров модели стандартным методом, и через секунду на экране появится следующее диалоговое окно результатов.

Шаг 4. В диалоговом окне **Multiple Regression Results (Результаты множественной регрессии)** просмотрите результаты оценивания. Результаты можно просмотреть в численном и графическом виде.

Окно результатов анализа имеет следующую структуру: верх окна – информационный. Он состоит из двух частей: в первой части содержится основная информация о результатах оценивания, во второй высвечиваются значимые регрессионные коэффициенты. Внизу окна **Multiple Regression Results (Результаты множественной регрессии)** находятся функциональные кнопки, позволяющие всесторонне просмотреть результаты анализа (рисунок 8.6).

Рассмотрим вначале информационную часть окна.

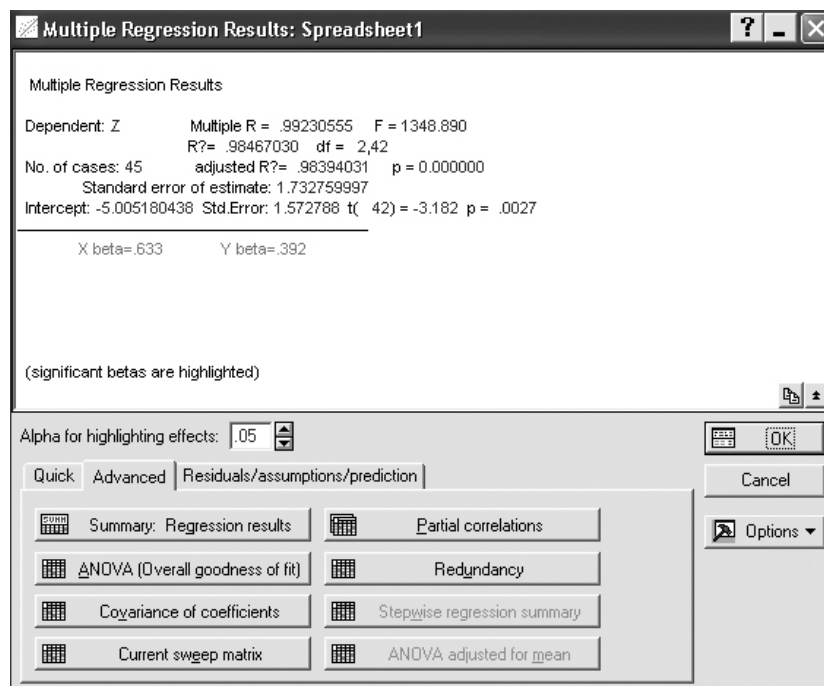


Рисунок 8.6 – Окно оценивания параметров в примере

В ней содержатся краткие сведения о результатах анализа. А именно:

– **Dep. Var.** (Имя зависимой переменной). В данном случае: *Z*;

– **No. of Cases** (Число случаев, по которым построена регрессия). В примере это число равно 45;

– **Multiple R = 0,99** (Коэффициент множественной корреляции);

– **R²** (Квадрат коэффициента множественной корреляции), обычно называемый коэффициентом детерминации. Коэффициент детерминации является чрезвычайно важной характеристикой, поэтому его следует подробно обсудить. Коэффициент детерминации является одной из основных статистик в данном окне, он показывает долю общего разброса (относительно выборочного среднего зависимой переменной), которая объясняется построенной регрессией;

– **Adjusted R²** (Скорректированный коэффициент детерминации), определяемый как: $Adjusted R^2 = R^2 \cdot n / (n - p)$,

где

– n – число наблюдений в модели, p – число параметров модели (число независимых переменных плюс 1, так как в модель включен свободный член);

– **Std. Error of estimate** (Стандартная ошибка оценки). Эта статистика является мерой рассеяния наблюдаемых значений относительно регрессионной прямой;

– **Intercept** (Оценка свободного члена регрессии). Значение коэффициента B_0 в уравнении регрессии;

– **Std. Error** (Стандартная ошибка оценки свободного члена). Стандартная ошибка коэффициента B_0 в уравнении регрессии;

– **t(df) and p-value** (Значение t-критерия и уровень p). **t-критерий** используется для проверки гипотезы о равенстве нулю свободного члена регрессии.

– **F** – значения **F-критерия**;

– **df** – число степеней свободы **F-критерия**;

– **p** – уровень значимости.

В информационной части посмотрим, прежде всего, на значения коэффициента детерминации. Значения коэффициента детерминации лежат в пределах от 0 до 1. В этом примере $R^2 = 0,98$. Это очень хорошее значение, показывающее, что построенная регрессия объясняет 98% разброса значений переменной Z относительно среднего.

Далее посмотрите на значение **F-критерия** и уровень значимости p . **F-критерий** используется для проверки гипотезы о значимости регрессии. В данном случае для проверки гипотезы, утверждающей, что между зависимой переменной Z и независимыми

переменными X и Y нет линейной зависимости, то есть $B_1 = 0$ и $B_2 = 0$, против альтернативы B_1 и B_2 не равны 0. В данном примере наблюдается большое значение **F-критерия** – 1348,89 и даваемый в окне уровень значимости $p = 0,0000$, показывающие, что построенная регрессия высоко значима.

Перейдем в функциональную часть окна результатов.

Прежде всего, нажмите кнопку **Regression summary** (Итоговый результат регрессии). На экране появится электронная таблица вывода – **spreadsheet**, в которой представлены итоговые результаты оценивания регрессионной модели (рисунок 8.7).

Regression Summary for Dependent Variable: Z (Spreadsheet1)						
R= .99230555 R ² = .98467030 Adjusted R ² = .98394031						
F(2,42)=1348.9 p<0.0000 Std.Error of estimate: 1.7328						
	Beta	Std. Err. of Beta	B	Std. Err. of B	t(42)	p-level
N=45						
Intercept			-5.00518	1.572789	-3.18236	0.002748
X	0.632874	0.038588	0.01181	0.000720	16.40092	0.000000
Y	0.391673	0.038588	9.32573	0.918773	10.15020	0.000000

Рисунок 8.7 – Итоговая таблица регрессии

Это стандартная таблица вывода регрессионного анализа. В первом столбце таблицы даны значения коэффициентов **beta** – стандартизованные коэффициенты регрессионного уравнения, во втором – стандартные ошибки **beta**, в третьем – точечные оценки параметров модели.

Свободный член $B_0 = -5,01$.

Коэффициент $B_1 = 0,012$ (при независимой переменной X).

Коэффициент $B_2 = 9,33$ (при независимой переменной Y).

Далее, стандартные ошибки для B_0 , B_1 , B_2 , значения статистик **t-критерия** и т. д.

Из таблицы видно, что оцененная модель имеет вид:

$$Z = 0,012 \cdot X + 9,33 \cdot Y - 5,01$$

Шаг 5. В окне оценивания параметров (рисунок 8.6) нажмите кнопку **Partial Correlation** (Частная корреляция). На рисунке 8.8 высветится результат частных корреляций: $r_{zx-y} = 0,93$ (частная корреляция между Z и X при постоянном значении Y) и $r_{zy-x} = 0,84$ (частная корреляция между Z и Y при постоянном значении X). Достоверность этих корреляций обусловлена высоким значением **t-критерия** и уровнем значимости $p = 0,00...$ Здесь следует обратить

внимание на то, что частные коэффициенты корреляции всегда меньше множественного $R = 0,99$ и соответствующих парных $0,97$ и $0,94$ (рисунки 8.11 и 8.12).

Variable	Beta in	Partial Cor.	Semipart Cor.	Tolerance	R-square	t(42)	p-level
X	0.632874	0.930026	0.313336	0.245124	0.754876	16.40092	0.000000
Y	0.391673	0.842851	0.193917	0.245124	0.754876	10.15020	0.000000

Рисунок 8.8 – Результат расчета частных корреляций

Шаг 6. В окне оценивания параметров (рисунок 8.6) последовательно нажмите кнопки **OK** (рисунок 8.9), **Scatterplot**, **Bivariate Correlation**. В появившемся окне выберите зависимую (**Z variable**) и независимую (**X variable**) переменные (рисунок 8.10 а). Нажмите кнопку **OK**.

Затем выберите зависимую (**Z variable**) и независимую (**Y variable**) переменные (рисунок 8.10 б). Нажмите кнопку **OK**.

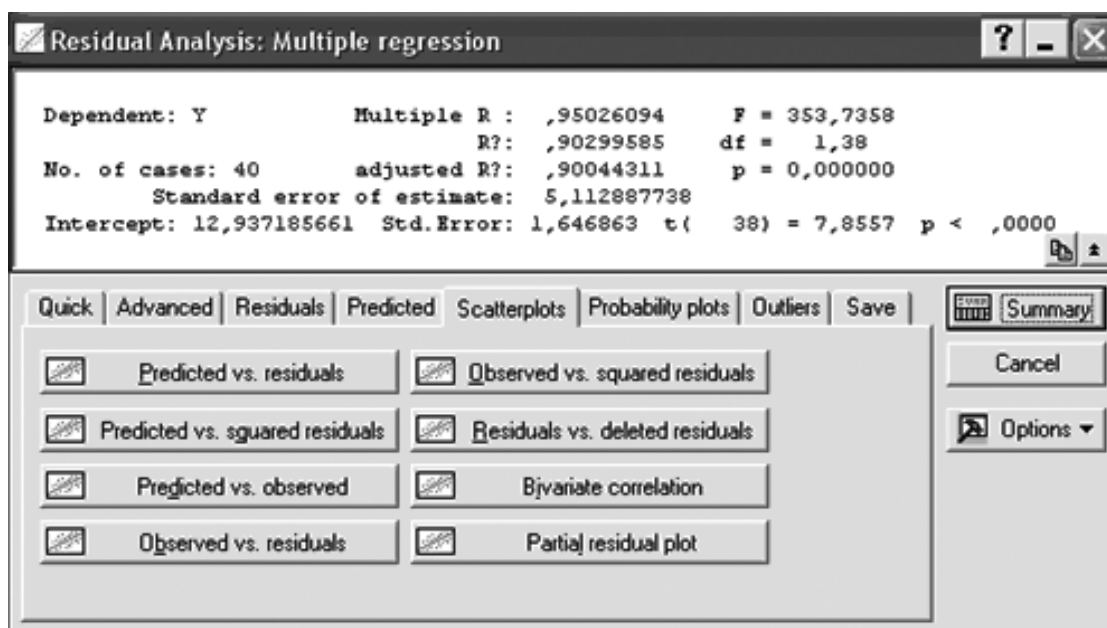
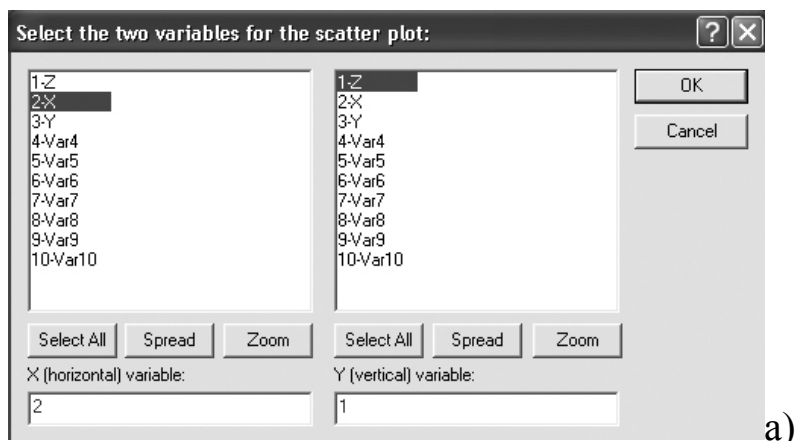
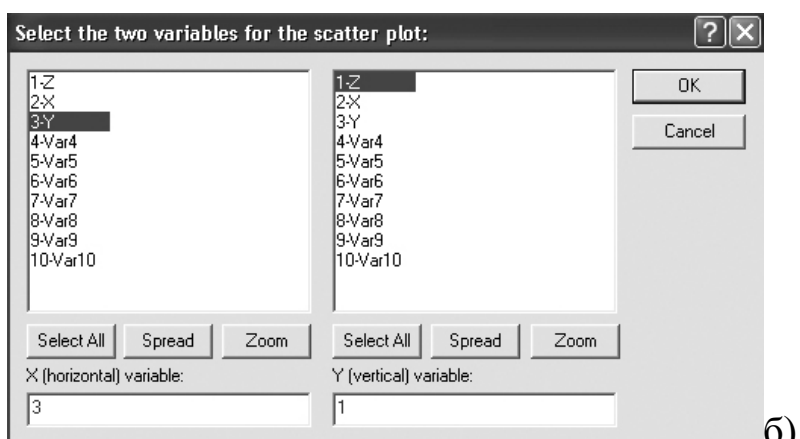


Рисунок 8.9 – Диалоговое окно Анализ остатков в модуле

На полученных графиках данные с подогнанной прямой имеют вид (рисунки 8.11 и 8.12).



a)



б)

Рисунок 8.10 – Выбор переменных а) – Z и X, б) – Z и Y

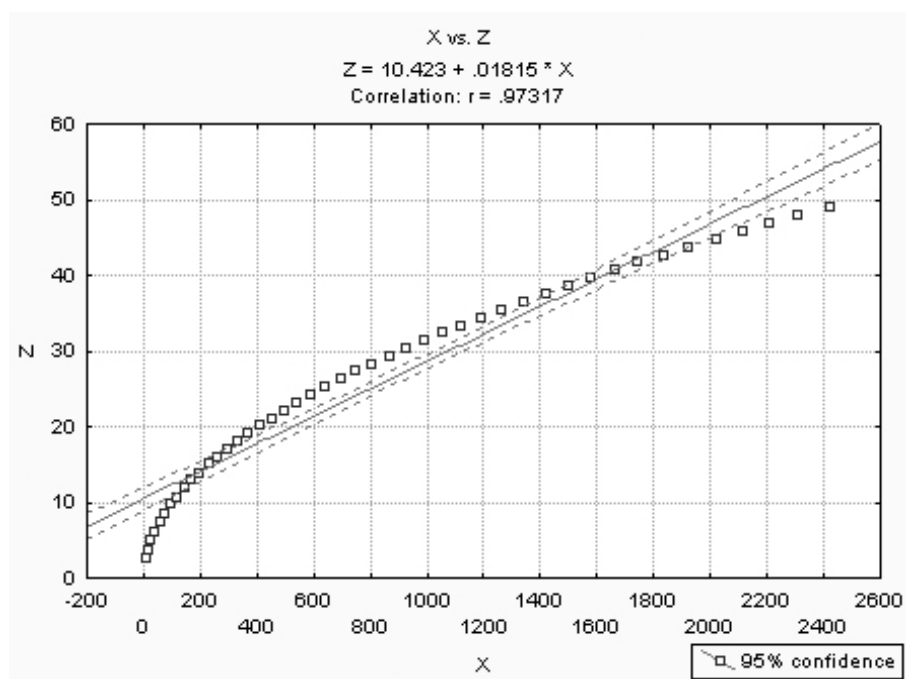


Рисунок 8.11 – Парная корреляция и линейная регрессия между Z и X

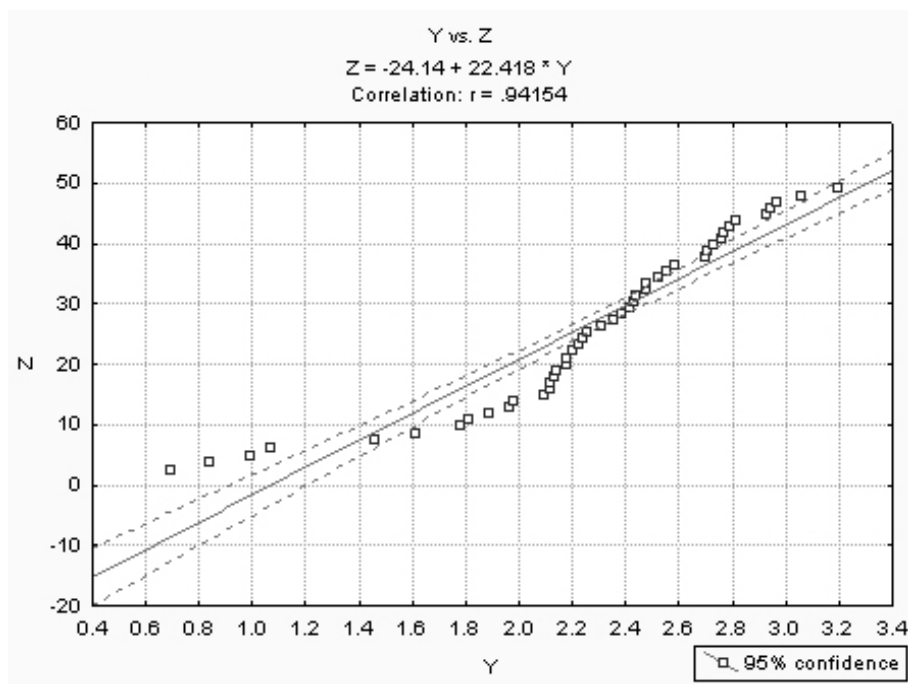


Рисунок 8.12 – Парная корреляция и линейная регрессия между Z и Y
 Шаг 7. Оценка адекватности модели.

Важным элементом анализа является оценка адекватности модели. Для этого в окне **Диалоговое окно Анализ остатков в модуле** (рисунок 8.9) нажмите клавишу **Predicted vs. Observed** (Модельное и наблюдаемое). Результат обработки представлен на рисунке 8.13. В данном случае адекватность модели наблюдаемым значениям не вызывает сомнений.

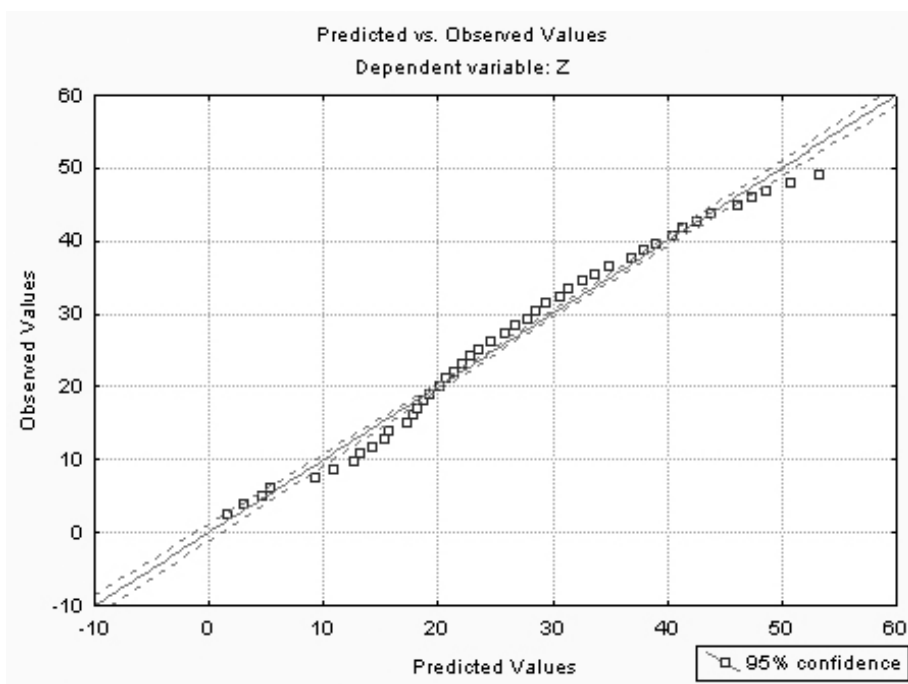


Рисунок 8.13 – Оценка адекватности модели

После того как доказана адекватность модели, полученные результаты можно уверенно использовать для дальнейших действий.

Анализ адекватности основывается на анализе остатков.

Остатки представляют собой разности между наблюдаемыми значениями и модельными, то есть значениями, подсчитанными по модели с оцененными параметрами.

Часто, если остатки не являются нормальными, а также для стабилизации дисперсии применяют преобразования зависимых и независимых переменных, например, логарифмическое преобразование зависимых переменных или извлечение квадратного корня.

Таким образом, применение модуля множественной регрессии показало сильную и достоверную корреляционную связь ($R = 0,99$; $F(2, 42) = 1348$ при $p < 0,00..$) между массой (зависимая переменная) – высотой и диаметром растений (независимые переменные), которая описывается уравнением линейной регрессии вида:

$$Z = 0,012 \cdot X + 9,33 \cdot Y - 5,01$$

В практическом отношении, следовательно, выполнив ряд замеров высоты и диаметра растений и, зная площадь и густоту посева, можно оценить общую биомассу.

Задания для выполнения

1 Введите в таблицу **Statistica 6** исходные данные из Приложения В (таблица В1).

2 Выполните расчетные процедуры в соответствии с порядком операций, выполненных в настоящем разделе.

3 Получите результат и сделайте заключение.

Лабораторная работа 9

Криволинейная корреляция и регрессия (MS Excel)

Цель работы: научиться выполнять корреляционный и регрессионный анализ в случае предполагаемой криволинейной функциональной связи в табличном редакторе MS Excel.

Краткие теоретические сведения

Если связь между изучаемыми явлениями существенно отклоняется от пропорциональной, что легко установить по графику, то коэффициент корреляции непригоден в качестве меры связи. Он может указать на отсутствие сопряженности там, где налицо сильная криволинейная зависимость. Поэтому необходим новый показатель, который правильно измерял бы степень криволинейной зависимости. Таким показателем является корреляционное отношение, обозначаемое греческой буквой η (эта). Оно измеряет степень корреляции при любой ее форме.

Корреляционное отношение измеряет степень криволинейных и прямолинейных связей.

Криволинейная связь между признаками – это такая связь, при которой равномерным изменениям первого признака соответствуют неравномерные изменения второго, причем эта неравномерность имеет определенный закономерный характер.

При графическом изображении криволинейных связей, когда по оси абсцисс откладывают значения первого признака (аргумент – независимая переменная), а по оси ординат – значения второго признака (функция – зависимая переменная) и полученные точки соединяют, получают изогнутые линии. Характер изогнутости зависит от природы коррелируемых признаков.

Корреляционное отношение измеряет степень корреляции при любой ее форме.

В отличие от коэффициента корреляции, который дает одинаковую меру связи признаков (первого со вторым и второго с первым), корреляционное отношение второго признака по первому обычно не бывает равно корреляционному отношению первого признака по второму. Поэтому крайне важно определить какая выборка является аргументом, а какая функцией.

По виду линии на графике можно определить характер связи (прямолинейная или криволинейная), также тип аппроксимации.

Задачей исследователя является подобрать вид функции, которая бы наиболее четко ложилась на поле регрессии, иначе: значение квадрата корреляционного отношения было бы максимально возможным.

Проведение анализа

Исследовалось влияние вводимого перорально с кормом сорбента (X – в г) молочным коровам на снижение содержания ^{137}Cs в молоке (Y – в %).

Исходные данные представлены на рисунке 9.1.

A	B
X	Y
14.5	49.8
5.1	22.5
6.1	28.2
6.5	31.2
6.7	33.5
8.2	41.8
7.4	34.9
10.7	46.9
14.3	48.8
8.1	40.3
8.1	40.3
4.6	21.3
8.4	42.2
9.6	44.9
8.8	43.9
9.1	44.3
9.6	44.8
7.4	36.2
8.5	43.9
11.2	47.1
12.7	47.9
13.4	48.2
13.8	48.7
7.2	34.1
8	40.3

Рисунок 9.1 – Исходные данные

Шаг 1. Из панели инструментов выберите и нажмите команду

Мастер диаграмм. В появившемся окне выберите тип (**Точечная**) и вид диаграммы, как показано на рисунке 9.2.

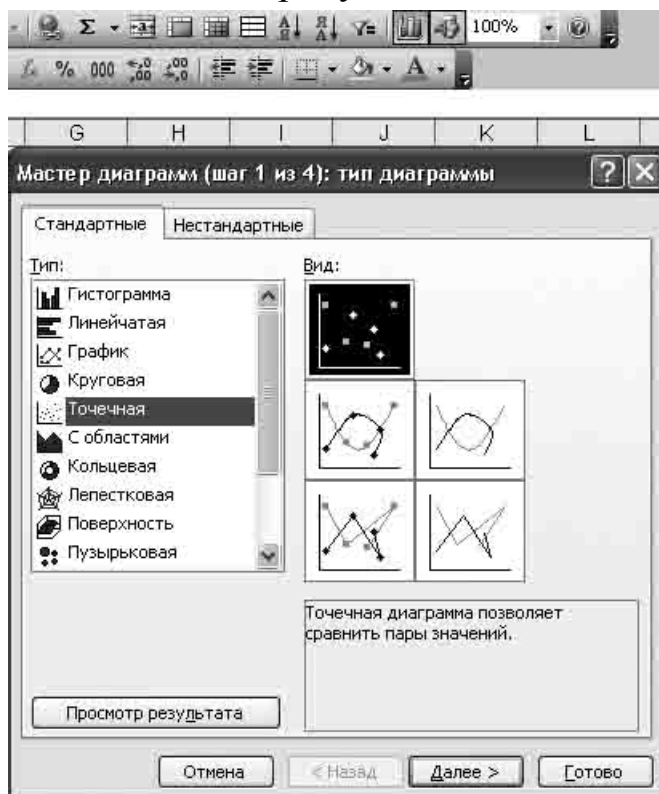


Рисунок 9.2 – Выбор типа и вида диаграммы

Шаг 2. Нажмите кнопку **Далее**. В окне **Диапазон** установите маркер, выберите в таблице анализируемые выборки (рисунок 9.3).

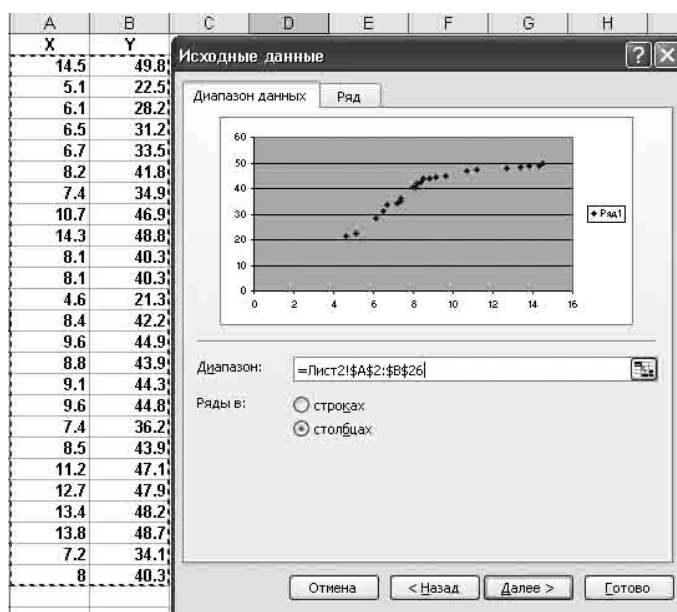


Рисунок 9.3 – Установка диапазона

Шаг 3. Нажмите кнопку **Далее**. Введите названия осей (рисунок 9.4).

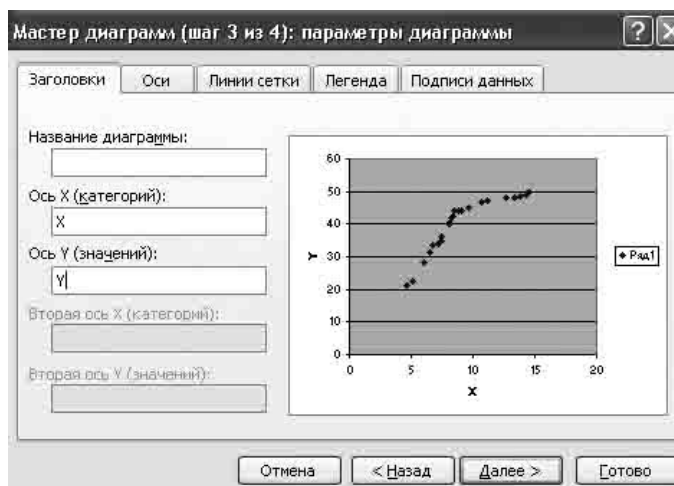


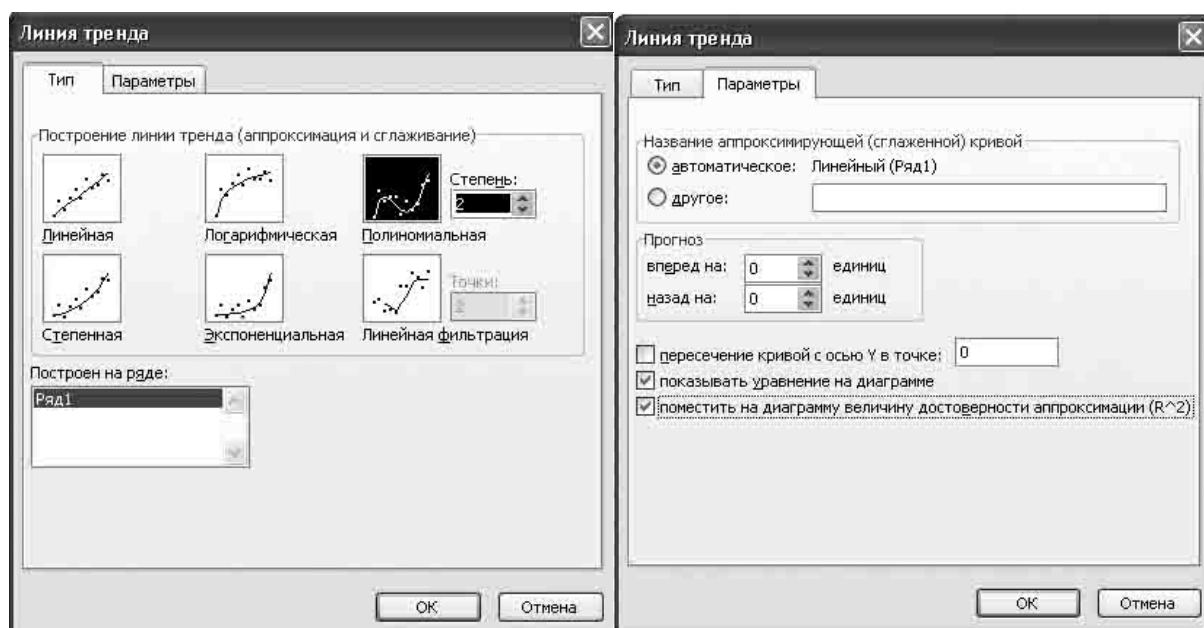
Рисунок 9.4 – Оформление диаграммы

Шаг 4. Щелкните мышкой **Готово**. Результат обработки появится в виде диаграммы.

Шаг 5. На графике щелкните правой кнопкой по любой точке диаграммы.

Шаг 6. Выберите опцию **Добавить линию тренда** и **Тип**. В меню представлены четыре типа аппроксимации: логарифмическая, полиномиальная, степенная и экспоненциальная. В данном случае **полином 2 степени** (рисунок 9.5 а).

Шаг 7. В опции **Параметры** выберите установки, как показано на рисунке 9.5 б. Щелкнете мышкой **ОК**.



а)

б)

Рисунок 9.5 – Тип аппроксимации и параметры кривой

Отредактированная диаграмма представлена на рисунке 9.6.

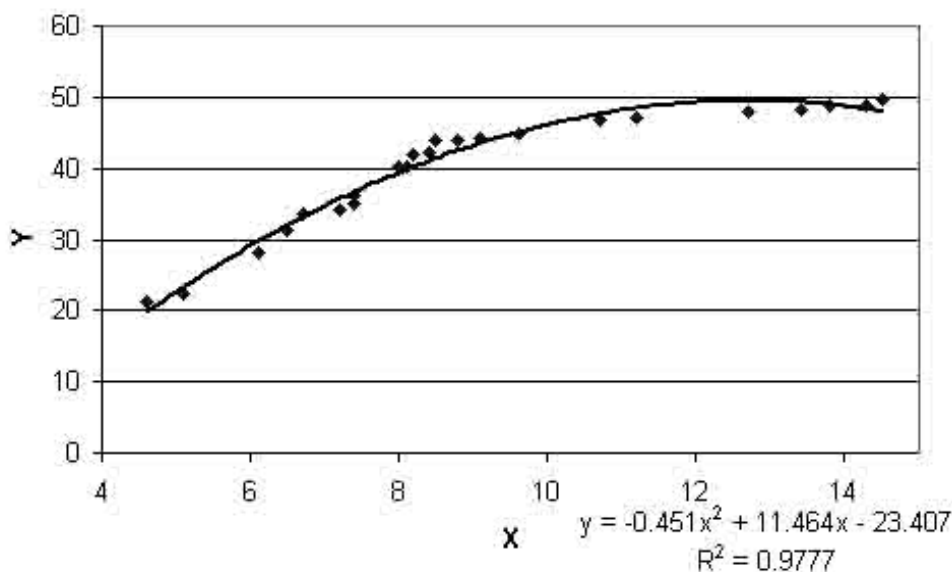


Рисунок 9.6 – Отредактированная диаграмма

Уравнение регрессии и квадрат корреляционного отношения находятся в правом нижнем углу диаграммы.

Результаты анализа показали, что корреляционная связь между массой вводимого сорбента и уровнем снижения содержания ^{137}Cs в молоке велика $r = 0,99$ и описывается полиномом 2 степени:

$$Y = -0,45 \cdot X^2 + 11,46 \cdot X - 23,41.$$

При массе вводимого сорбента больше 12 г наблюдается насыщение. Поэтому оптимальное значение массы сорбента следует рекомендовать в пределах 12 – 14 г.

Задания для выполнения

- 1 Введите в таблицу **MS Excel** исходные данные из Приложения Г (таблица Г1).
- 2 Выполните расчетные процедуры в соответствии с порядком операций, выполненных в настоящем разделе.
- 3 Получите результат и сделайте заключение.

Литература

- 1 Лакин, Г. Ф. Биометрия: учебн. пос. / Г. Ф. Лакин. – М. : «Высшая школа», 1990. – 142 с.
- 2 Зайцев, Г. Н. Математическая статистика в экспериментальной ботанике: учебн. пос. / Г. Н. Зайцев. – М. : «Наука», 1984. – 424 с.
- 3 Плохинский, Н. А. Биометрия: 2-е издание учебн. пос. / Н. А. Плохинский. – М. : «МГУ», 1970. – 368 с.
- 4 Свалов, Н. Н. Вариационная статистика / Н. Н. Свалов. – М. : «Лесная промышленность», 1977. – 177 с.
- 5 Рокитский, П. Ф. Биологическая статистика: изд. 3-е испр. учебн. пос. / П. Ф. Рокитский. – Мн. : «Высшая школа», 1973. – 320 с.
- 6 Боровиков, В. П. Популярное введение в программу Statistica / В. П. Боровиков. – М. : «КомпьютерПресс», 1998. – 69 с.
- 7 Жученко, Ю. М. Статистическая обработка информации с применением персональных компьютеров: практическое руководство для студентов 5 курса / Ю. М. Жученко. – Гомель, УО ГГУ им. Ф. Скорины, 2007. – 101 с.
- 8 Мюллер, П. Таблицы по математической статистике: справочник / П. Мюллер [и др.]. – М. : «Финансы и статистика», 1982. – 64 с.
- 9 Павловский, З. Введение в математическую статистику: перевод с польск. / З. Павловский. – М. : «Статистика», 1967. – 285 с.
- 10 Карасев, А. И. Теория вероятностей и математическая статистика: учебн. пос. / А. И. Карасев. – М. : «Статистика», 1979. – 279 с.
- 11 Бейли, Н. Математика в биологии и медицине: перевод с англ. / Н. Бейли. – М. : «Мир», 1970. – 167 с.
- 12 Урбах, В. Ю. Статистический анализ в биологических и медицинских исследованиях: учебн. пос. / В. Ю. Урбах. – М. : «Медицина», 1975. – 321 с.
- 13 Лапач, С. Н. Статистические методы в медико-биологических исследованиях с использованием Excel / С. Н. Лапач [и др.]. – К. : «МОРИОН», 2000. – 196 с.
- 14 Реброва, О. Ю. Статистический анализ медицинских данных: применение пакета прикладных программ STATISTICA / О. Ю. Реброва. – М.: «МедиаСфера», 2002. – 84 с.

Приложение А (обязательное)

Таблица А1 Задания к лабораторным работам 4 и 5

Варианты заданий														
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
13,45	81,92	3,00	20,15	5,48	8	12	6	39	27	1	20	3	29	30
7,35	147,02	2,17	8,79	9,35	2	22	12	49	32	19	20	14	55	71
8,53	109,30	2,53	12,56	3,94	13	1	12	51	22	3	22	11	39	34
16,58	91,55	1,73	5,65	9,94	4	9	15	48	13	16	25	6	46	40
12,86	79,85	2,74	15,44	1,52	22	8	14	102	19	14	22	28	54	51
10,11	80,34	0,78	2,18	13,31	19	11	6	61	12	20	2	15	89	80
17,22	95,98	2,01	7,49	1,67	5	1	1	73	16	12	1	20	71	44
8,72	95,22	1,66	5,26	5,67	11	5	11	63	27	20	11	14	40	39
20,76	150,85	2,73	15,38	12,35	12	12	12	30	1	11	12	3	55	23
5,30	108,46	1,48	4,38	10,36	23	5	14	107	28	18	19	12	69	65
91,55	43,99	1,99	7,30	11,69	6	17	7	37	10	32	14	1	29	52
79,85	106,15	2,21	9,11	5,00	14	7	1	43	17	7	23	1	67	47
80,34	135,03	2,95	19,13	1,84	12	6	18	21	35	9	10	47	69	71
95,98	119,52	2,39	10,90	7,19	6	14	12	44	21	18	2	21	41	40
95,22	111,39	2,28	9,80	2,25	12	7	12	40	13	23	4	26	84	51
4,48	87,89	2,84	17,15	5,44	1	20	4	45	3	17	13	24	20	47
4,35	77,79	2,28	9,76	1,90	10	8	18	44	10	22	22	2	63	44
5,00	147,68	2,26	9,54	7,90	5	7	8	42	5	11	13	8	52	74
4,85	127,11	1,77	5,84	1,98	11	18	11	63	9	13	11	1	74	61
4,99	146,63	2,26	9,62	14,39	21	7	2	33	21	36	5	24	75	64
4,86	129,39	1,07	2,92	9,06	11	7	12	44	31	42	14	16	33	51
4,49	89,21	2,85	17,21	14,61	11	6	8	86	19	18	36	11	52	58
	132,30	1,65	5,19	3,67	18	11	12	49	2	16	2	20	59	48
	133,92	1,72	5,56	6,23	1	3	16	38	2	28	18	26	35	70
	111,96	1,20	3,31	18,84	1	15	12	74	46	23	1	20	63	50
	127,69	2,15	8,59	42,65	11	5	8	17	19	22	11	1	33	63
	118,32	1,99	7,32	3,92	10	1	16	53	17	1	5	12	48	39
	55,86	1,75	5,75	11,07	9	6	12	30	9	16	9	18	61	79
	155,82	0,15	1,16	6,22	23	8	20	73	32	6	13	2	13	9
	54,09	1,50	4,49		14	15	11	30	5	22	18	15	29	21
	69,98	2,55	12,75										34	59
	40,91	2,57	13,08											
	117,10	2,87	17,57											
	80,15	2,32	10,13											
	59,97	2,56	12,90											
	104,08	2,00	7,38											
	77,97	2,16	8,66											
	134,53	2,56	12,89											
	47,15	1,02	2,78											

Приложение Б (обязательное)

Таблица Б1 Задания к лабораторным работам 6 и 7

Варианты заданий													
1		2		3		4		5		6		7	
X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y
14,7	171,0	5,2	17,5	5,2	17,5	12,0	3,6	1,4	28,4	6,0	3,9	26,8	5,1
13,7	202,2	5,4	17,6	5,4	17,6	12,0	3,7	1,8	28,4	8,2	2,9	47,3	5,0
15,2	124,6	5,7	17,6	5,7	17,6	12,0	4,0	1,9	28,5	6,3	2,1	49,6	4,7
16,3	119,0	6,2	17,6	6,2	17,6	11,9	4,6	2,1	28,7	18,1	1,9	58,0	4,6
16,2	124,3	6,7	17,8	6,7	17,8	11,9	4,9	2,3	29,0	13,9	3,8	59,3	4,5
14,8	175,2	7,1	18	7,1	18	11,7	5,0	2,4	29,0	4,3	2,8	59,7	4,4
16,1	161,7	7,5	18	7,5	18	11,3	5,0	2,4	29,2	11,2	2,9	63,3	4,3
13,9	169,4	7,7	18,1	7,7	18,1	11,2	5,6	2,5	29,3	11,3	2,4	68,5	4,3
14,3	199,9	7,8	18,1	7,8	18,1	10,3	5,9	2,5	29,3	2,9	3,4	73,0	4,2
13,3	115,6	7,9	18,6	7,9	18,6	10,0	6,2	2,7	29,4	20,5	4,0	73,2	4,2
14,0	155,0	8	19,7	8	19,7	9,6	6,3	2,7	29,5	9,4	2,2	74,3	4,2
14,2	141,8	8,3	19,7	8,3	19,7	9,5	6,3	2,9	29,5	15,5	2,4	76,9	4,1
14,4	186,0	8,4	19,8	8,4	19,8	9,4	6,3	2,9	29,5	7,7	1,9	78,1	3,8
14,6	159,4	8,5	19,9	8,5	19,9	9,3	6,4	2,9	29,6	11,3	2,0	80,3	3,7
15,1	122,0	8,8	19,9	8,8	19,9	9,2	6,4	3,0	29,6	7,1	3,7	83,0	3,5
14,6	72,2	8,9	20,1	8,9	20,1	9,1	6,5	3,0	29,7	-0,2	2,4	84,0	3,4
14,7	123,8	8,9	20,1	8,9	20,1	9,0	6,6	3,1	29,7	17,1	4,3	88,2	3,4
14,6	166,5	8,9	20,1	8,9	20,1	8,9	7,1	3,2	29,9	12,9	2,8	92,7	3,4
16,3	117,0	8,9	20,1	8,9	20,1	8,6	7,1	3,3	29,9	7,7	3,8	93,6	3,4
14,9	174,4	9	20,2	9	20,2	8,2	7,1	3,3	29,9	10,5	3,7	95,6	3,3
14,8	186,8	9	20,3	9	20,3	8,2	7,2	3,4	30,0	16,5	3,6	95,9	3,2
14,5	131,1	9,1	20,3	9,1	20,3	8,2	7,4	3,4	30,1	12,1	4,8	96,4	3,0
15,9	199,6	9,3	20,5	9,3	20,5	8,2	7,5	3,5	30,1	9,3	1,9	97,2	3,0
14,3	93,3	9,4	20,6	9,4	20,6	8,2	7,6	3,5	30,1	8,8	3,0	102,1	2,9
13,4	135,6	9,7	20,9	9,7	20,9	7,9	8,0	3,5	30,1	7,6	2,8	102,2	2,9
15,5	121,8	9,7	21	9,7	21	7,7	8,1	3,5	30,2	8,7	3,1	106,1	2,8
15,9	127,6	9,9	21,1	9,9	21,1	7,5	8,2	3,6	30,3	7,3	3,5	106,4	2,8
14,9	149,5	10,1	21,1	10,1	21,1	7,3	8,3	3,6	30,4	1,9	3,7	109,7	2,8
14,5	131,7	10,4	21,2	10,4	21,2	7,3	8,3	3,6	30,4	4,6	3,5	119,4	2,7
15,7	156,5	10,7	21,3	10,7	21,3	7,2	8,8	3,6	30,5	10,7	3,2	131,0	2,6
14,6	185,5	11,6	21,3	11,6	21,3	7,0	9,1	3,7	30,6	5,8	3,0	133,7	2,6
15,8	129,3	12	21,8	12	21,8	6,9	9,3	3,8	30,6	10,6	3,4	134,8	2,4
13,6	223,6	12,3	22	12,3	22	6,3	9,6	3,9	30,6	13,1	3,2	136,1	2,4
14,2	198,6	12,6	22,1	12,6	22,1	6,1	9,6	3,9	30,7	12,9	3,3	138,8	2,4
13,5	180,8	12,7	22,4	12,7	22,4	0,7	29,5	4,0	30,8	8,7	1,6	139,1	2,4
14,6	132,9	13,3	22,7	13,3	22,7	35,9	1,1	4,0	31,0	11,8	2,3	139,4	2,3
15,0	172,4	13,6	23,5	8,8	44	1,1	43,0	4,0	31,1	9,0	5,0	161,0	2,0

Приложение В (обязательное)

Таблица В1 Задания к лабораторной работе 8

Варианты заданий														
1			2			3			4			5		
Z	X	Y	Z	X	Y	Z	X	Y	Z	X	Y	Z	X	Y
2419	2,69	0,69	5,8	0,057	2	31,4	0,04	2	0,021	38	2	2,69	7,25	0,69
2309	3,83	0,83	7,7	0,059	2,3	31,4	0,045	2,3	0,025	34,7	2,3	3,83	14,69	0,83
2205	4,99	0,99	8,5	0,061	2,7	30,4	0,048	2,7	0,044	31,5	2,7	4,99	24,93	0,99
2110	6,06	1,06	9,9	0,068	2,9	27,1	0,062	2,9	0,045	30,6	2,9	6,06	36,78	1,06
2018	7,46	1,46	10,1	0,072	4,3	26,6	0,063	4,3	0,046	26,9	4,3	7,46	55,63	1,46
1919	8,61	1,61	11,2	0,073	5	26,5	0,066	5	0,05	26,8	5	8,61	74,12	1,61
1831	9,77	1,77	12,8	0,074	5,9	25,2	0,066	5,9	0,05	26,4	5,9	9,77	95,55	1,77
1744	10,81	1,81	12,9	0,078	6,1	24,8	0,068	6,1	0,056	24,7	6,1	10,81	116,82	1,81
1661	11,89	1,89	14	0,078	6,6	24,6	0,073	6,6	0,061	24	6,6	11,89	141,30	1,89
1578	12,96	1,96	14,2	0,078	7,1	24,5	0,079	7,1	0,063	23,7	7,1	12,96	167,96	1,96
1497	13,97	1,97	14,6	0,079	7,2	24,1	0,081	7,2	0,068	23,3	7,2	13,97	195,27	1,97
1420	15,09	2,09	15,4	0,081	8,1	23,7	0,083	8,1	0,069	23,1	8,1	15,09	227,76	2,09
1338	16,12	2,12	16,1	0,082	8,3	23,7	0,084	8,3	0,071	23	8,3	16,12	259,73	2,12
1263	17,12	2,12	16,2	0,083	8,3	23,3	0,084	8,3	0,072	22,9	8,3	17,12	292,97	2,12
1191	18,13	2,13	16,8	0,085	8,4	22,5	0,085	8,4	0,072	22,8	8,4	18,13	328,63	2,13
1120	19,14	2,14	17,1	0,086	8,5	22,3	0,092	8,5	0,074	22,7	8,5	19,14	366,34	2,14
1054	20,17	2,17	17,1	0,088	8,8	22,1	0,096	8,8	0,075	22,5	8,8	20,17	407,02	2,17
988	21,17	2,17	17,3	0,088	8,8	22	0,098	8,8	0,078	22,4	8,8	21,17	448,37	2,17
925	22,20	2,20	17,7	0,09	9	21,8	0,098	9	0,085	21,9	9	22,20	492,72	2,20
865	23,22	2,22	17,7	0,092	9,2	20,8	0,099	9,2	0,088	21,4	9,2	23,22	539,13	2,22
805	24,24	2,24	17,8	0,092	9,4	20	0,1	9,4	0,089	19,9	9,4	24,24	587,61	2,24
748	25,25	2,25	18,7	0,093	9,5	20	0,102	9,5	0,09	19,7	9,5	25,25	637,63	2,25
691	26,30	2,30	19,5	0,096	10	19,9	0,102	10	0,092	19,7	10	26,30	691,83	2,30
637	27,35	2,35	19,7	0,099	10,5	19,8	0,103	10,5	0,092	19,6	10,5	27,35	748,10	2,35
587	28,38	2,38	20,4	0,099	10,8	19,7	0,104	10,8	0,094	19,4	10,8	28,38	805,40	2,38
539	29,42	2,42	20,7	0,106	11,2	19,6	0,109	11,2	0,095	18,4	11,2	29,42	865,30	2,42
492	30,42	2,42	21	0,107	11,3	18,7	0,109	11,3	0,096	18	11,3	30,42	925,67	2,42
448	31,43	2,43	21,7	0,108	11,4	18,3	0,11	11,4	0,097	17,7	11,4	31,43	988,07	2,43
407	32,48	2,48	22,5	0,108	11,9	18,2	0,115	11,9	0,102	17,3	11,9	32,48	1054,73	2,48
366	33,48	2,48	22,5	0,11	11,9	18	0,115	11,9	0,103	16,7	11,9	33,48	1120,68	2,48
328	34,52	2,52	22,6	0,112	12,4	17,2	0,116	12,4	0,106	16,7	12,4	34,52	1191,47	2,52
292	35,55	2,55	22,8	0,112	12,8	16,8	0,117	12,8	0,106	15,9	12,8	35,55	1263,76	2,55
259	36,58	2,58	22,9	0,112	13,2	16,8	0,118	13,2	0,106	15,8	13,2	36,58	1338,11	2,58
227	37,69	2,69	23,5	0,113	14,8	15,8	0,122	14,8	0,107	15,6	14,8	37,69	1420,88	2,69
195	38,70	2,70	23,5	0,113	14,9	15,7	0,123	14,9	0,113	15,4	14,9	38,70	1497,80	2,70
167	39,73	2,73	23,9	0,12	15,3	15,6	0,124	15,3	0,122	14,6	15,3	39,73	1578,30	2,73
141	40,76	2,76	24,5	0,12	15,8	15	0,125	15,8	0,124	14,2	15,8	40,76	1661,38	2,76

Приложение Г (обязательное)

Таблица Г1 Задания к лабораторным работам 9 и 10

Варианты заданий													
1		2		3		4		5		6		7	
X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y
14,5	49,8	33,3	0,03	9,06	2,13	0,9	66,9	0,97	0,4	1,5	0,7	0,87	0,15
5,1	22,5	7,1	0,2	8,95	2,12	0,01	0,01	1,21	2,79	1,47	0,69	0,98	0,11
6,1	28,2	11,2	0,13	5,58	1,61	-0,8	-46,4	0,83	0,09	1,58	0,73	1,32	0,04
6,5	31,2	14,3	0,08	11,26	2,38	1,2	165,1	1,16	1,83	1,57	0,73	1,22	0,05
6,7	33,5	17,2	0,06	9,14	2,14	1,5	314,8	0,95	0,31	1,45	0,68	0,81	0,18
8,2	41,8	22,2	0,05	2,87	0,83	1,1	141,1	1,53	22,9	1,45	0,69	1,47	0,02
7,4	34,9	34,7	0,03	12,01	2,42	-0,1	-0,1	1,09	1,13	1,02	0,51	0,96	0,11
10,7	46,9	11,2	0,1	6,86	1,81	0,4	5,8	0,48	0	1,18	0,58	0,77	0,2
14,3	48,8	3,7	0,37	7	1,89	0,8	60,5	1,1	1,15	0,92	0,46	0,93	0,12
8,1	40,3	33,8	0,03	2,86	0,99	-0,1	-0,3	1,15	1,73	0,96	0,48	0,88	0,14
8,1	40,3	10,9	0,12	7,59	1,97	1,2	175	1,04	0,73	1,78	0,79	0,96	0,11
4,6	21,3	13,1	0,08	10,47	2,25	2,6	1845,6	1,23	3,15	0,55	0,2	0,73	0,23
8,4	42,2	18,1	0,07	6,78	1,77	1,6	421,1	0,75	0,04	1,21	0,6	1,56	0,02
9,6	44,9	20,1	0,06	2,57	0,69	1,2	192,9	1,02	0,62	0,86	0,43	1,25	0,05
8,8	43,9	23,4	0,05	7,59	1,96	0,7	42,4	1,11	1,3	0,61	0,26	1,27	0,04
9,1	44,3	30,1	0,04	9,59	2,17	-0,01	-0,01	1,14	1,63	1,09	0,54	0,68	0,26
9,6	44,8	27	0,04	11,08	2,35	0,02	0,02	1,35	7,22	0,86	0,43	1,12	0,07
7,4	36,2	25,4	0,04	12,33	2,43	0,2	1,1	1,11	1,31	1,12	0,56	1,53	0,02
8,5	43,9	25,6	0,04	11,76	2,42	1,6	443,8	1,78	88,7	0,53	0,19	0,72	0,23
11,2	47,1	30,1	0,04	9,49	2,2	0,4	8,4	0,79	0,06	1,16	0,58	1,13	0,07
12,7	47,9	27,7	0,04	12,2	2,48	-0,3	-2,8	0,84	0,1	1,05	0,52	1,12	0,07
13,4	48,2	26	0,05	9,62	2,22	-0,5	-13,1	0,83	0,1	0,57	0,22	0,94	0,12
13,8	48,7	23,2	0,05	12,53	2,48	0,8	50	1,21	2,76	0,9	0,45	1,32	0,04
7,2	34,1	19	0,06	3,51	1,06	0,3	4,3	1,2	2,61	0,75	0,36	0,95	0,12
8	40,3	32,1	0,03	4,76	1,46	0,6	19,8	1,31	5,75	0,86	0,43	1,2	0,06
		16,8	0,07	10,41	2,3	0,2	1,1	1,67	49,21	0,77	0,37	0,72	0,23
		12,1	0,09	8,71	2,09	0,6	23,2	1,18	2,29	1,61	0,74	0,69	0,25
		9,3	0,17	8,7	2,12	1,4	255,9	0,65	0,01	0,68	0,31	0,96	0,11
		6,7	0,26	9,68	2,17	0,8	44,8	0,89	0,17	1,01	0,51	0,9	0,14
		21,5	0,05	10,04	2,24	1,5	326,4	0,68	0,01	1,36	0,65	1,67	0,01

Приложение Д (обязательное)

Таблица Д1 Задания к лабораторным работам 11 и 12

Варианты заданий											
1				2				3			
a	b	c	d	a	b	c	d	a	b	c	d
9,4	21,1	12,5	21,0	1,95	6,87	1,47	1,94	24,4	21,5	17,3	33,0
8,4	8,2	12,1	5,0	2,87	3,34	2,77	2,05	15,8	23,2	22,8	14,2
4,9	16,6	12,1	24,1	1,89	3,32	3,06	1,69	19,5	16,8	11,0	18,3
13,4	10,5	15,4	14,8	2,97	6,19	2,58	2,34	16,5	20,5	17,8	27,2
11,1	18,4	2,0	29,1	1,70	5,75	2,38	2,41	16,5	8,2	23,5	5,6
3,6	11,3	9,5	20,1	2,05	5,61	3,25	1,21	15,0	19,4	11,9	22,6
15,1	18,3	12,1	11,2	3,02	7,33	3,32	2,83	21,4	28,6	15,0	31,4
13,2	11,2	11,0	19,5	1,57	4,43	3,02	1,56	17,5	21,3	24,6	26,4
12,0	19,4	5,8	20,5	1,49	4,45	2,10	1,81	17,2	19,8	19,9	33,9
11,6	10,0	16,0	22,9	2,28	5,84	3,25	2,61	11,5	12,7	20,8	26,6
18,2	11,2	12,1	17,7	3,29	5,20	2,57	2,43	11,5	23,5	22,8	14,1
12,4	11,1	16,6	13,8	1,91	3,68	2,86	1,36	12,0	20,2	29,2	20,7
15,2	6,3	15,5	9,3	3,32	4,09	3,54	2,25	15,3	19,1	25,4	25,8
9,8	16,3	11,5	20,4	2,21	7,21	2,76	1,60	30,2	29,8	19,8	21,7
8,0	18,8	9,3	16,7	4,32	3,28	3,32	1,03	33,3	25,9	25,6	29,4
12,0	10,1	14,4	18,3	1,57	6,31	2,89	1,53	15,4	22,5	10,8	27,0
13,3	17,6	12,2	29,1	1,07	4,69	2,51	1,84	32,1	31,9	13,1	23,2
12,0	6,6	13,2	13,1	2,19	5,98	2,61	1,34	40,1	11,1	12,3	29,9
16,7	14,7	6,2	19,5	0,02	4,00	3,30	2,35	22,3	21,3	14,3	17,8
7,3	18,2	8,1	20,0	2,06	5,27	1,97	1,88	25,0	14,4	24,2	38,9
8,7	5,5	13,9	12,4								
13,1	17,7	12,2	14,3								
7,6	11,9	5,3	15,5								
11,2	5,5	15,7	10,4								
8,5	10,4	14,2	31,7								
19,4	9,2	10,5	11,5								
10,6	10,7	13,6	21,2								
7,3	10,8	10,1	15,5								
17,9	13,8	13,0	15,5								
20,3	19,1	11,4	26,1								
17,2	15,6	13,3	15,7								
13,3	12,9	12,6	24,0								
10,2	14,7	12,5	20,7								
15,9	15,2	9,8	13,7								
15,0	14,8	17,6	15,0								
10,1	9,0	11,0	9,0								

Окончание таблицы Д1

Варианты заданий												
4				5				6				
a	b	c	d	a	b	c	d	a	b	c	d	e
5,8	9,5	31,3	1,1	9,4	21,1	12,5	21,0	2,30	4,72	0,70	1,00	0,21
7,8	10,6	10,8	0,7	8,4	8,2	12,1	5,0	1,20	4,55	1,37	0,94	0,43
12,4	3,1	27,8	1,6	4,9	16,6	12,1	24,1	1,55	1,64	0,91	1,06	0,31
8,5	24,9	19,8	0,9	13,4	10,5	15,4	14,8	2,17	4,01	1,50	1,20	0,52
5,7	10,0	18,1	1,1	11,1	18,4	2,0	29,1	2,21	5,34	1,31	1,24	0,34
10,7	19,0	28,2	0,7	3,6	11,3	9,5	20,1	1,77	3,47	1,07	1,20	0,02
11,3	16,2	43,7	1,6	15,1	18,3	12,1	11,2	0,76	6,03	1,39	0,97	0,34
12,5	20,9	19,1	0,7	13,2	11,2	11,0	19,5	2,23	5,06	1,15	0,95	0,24
4,7	2,1	28,0	1,0	12,0	19,4	5,8	20,5	2,44	4,25	1,19	0,83	0,46
11,8	24,5	20,1	1,1	11,6	10,0	16,0	22,9	2,04	3,61	0,98	0,73	0,60
5,4	17,2	15,0	1,1	18,2	11,2	12,1	17,7	2,38	3,55	1,08	0,95	0,31
10,3	13,5	20,7	0,8	12,4	11,1	16,6	13,8	2,84	4,54	1,27	0,97	0,28
6,5	14,7	12,6	0,9	15,2	6,3	15,5	9,3	2,46	3,19	1,22	1,14	0,62
10,1	7,2	11,5	1,1	9,8	16,3	11,5	20,4	1,53	4,23	1,44	1,13	0,11
14,2	16,0	16,6	1,0	8,0	18,8	9,3		1,68	4,84	0,80	0,83	0,81
9,8	4,9	22,9	1,1	12,0	10,1	14,4		3,20	3,97	1,04	0,82	
10,0	19,4	12,0	1,3	13,3	17,6	12,2		1,45	5,10	0,68	0,98	
12,4	11,6	21,0	1,4	12,0	6,6	13,2		1,72	4,87	1,18	1,08	
6,7	15,8	15,1	0,7	16,7	14,7	6,2		2,30	2,06	0,99		
5,7	7,8	26,2	0,7	7,3	18,2	8,1		1,74	5,21	0,94		
12,1	16,8	24,0	1,1	8,7	5,5	13,9		2,35	3,50			
8,0	19,3	26,2	0,9		17,7	12,2		2,43	5,31			
10,4	19,1	8,0	1,6		11,9	5,3		1,16	3,81			
9,8	17,1	35,8	0,9		5,5	15,7		2,30	4,61			
					10,4	14,2		1,66	3,55			
						10,5			3,17			
						13,6			3,93			
						10,1			4,82			
						13,0			3,91			
						11,4			5,68			
						13,3						
						12,6						
						12,5						
						9,8						
						17,6						
						11,0						

Приложение Е (обязательное)

Таблица Е1 Задания к лабораторным работам 13 и 14

Варианты заданий														
1					2					3				
	a	b	c	d		a	b	c	d		a	b	c	d
A	0,02	3,28	1,47	1,03	A	14,2	24,9	43,7	1,1	A	3,6	21,1	2,0	5,0
A	1,07	3,32	1,97	1,21	A	12,5	24,5	35,8	0,7	A	4,9	19,4	5,3	9,0
A	1,49	3,34	2,10	1,34	A	12,4	20,9	31,3	1,6	A	7,3	19,1	5,8	9,3
A	1,57	3,68	2,38	1,36	A	12,4	19,4	28,2	0,9	A	7,3	18,8	6,2	10,4
A	1,57	4,00	2,51	1,53	A	12,1	19,3	28,0	1,1	A	7,6	18,4	8,1	11,2
A	1,70	4,09	2,57	1,56	A	11,8	19,1	27,8	0,7	A	8,0	18,3	9,3	11,5
A	1,89	4,43	2,58	1,60	A	11,3	19,0	26,2	1,6	A	8,4	18,2	9,5	12,4
A	1,91	4,45	2,61	1,69	A	10,7	17,2	26,2	0,7	A	8,5	17,7	9,8	13,1
A	1,95	4,69	2,76	1,81	B	10,4	17,1	24,0	1,0	A	8,7	17,6	10,1	13,7
A	2,05	5,20	2,77	1,84	B	10,3	16,8	22,9	1,1	A	9,4	16,6	10,5	13,8
B	2,06	5,27	2,86	1,88	B	10,1	16,2	21,0	1,1	A	9,8	16,3	11,0	14,3
B	2,19	5,61	2,89	1,94	B	10,0	16,0	20,7	0,8	A	10,1	15,6	11,0	14,8
B	2,21	5,75	3,02	2,05	B	9,8	15,8	20,1	0,9	B	10,2	15,2	11,4	15,0
B	2,28	5,84	3,06	2,25	B	9,8	14,7	19,8	1,1	B	10,6	14,8	11,5	15,5
B	2,87	5,98	3,25	2,34	B	8,5	13,5	19,1	1,0	B	11,1	14,7	12,1	15,5
B	2,97	6,19	3,25	2,35	B	8,0	11,6	18,1	1,1	B	11,2	14,7	12,1	15,5
B	3,02	6,31	3,30	2,41	C	7,8	10,6	16,6	1,3	B	11,6	13,8	12,1	15,7
B	3,29	6,87	3,32	2,43	C	6,7	10,0	15,1	1,4	B	12,0	12,9	12,1	16,7
B	3,32	7,21	3,32	2,61	C	6,5	9,5	15,0	0,7	B	12,0	11,9	12,2	17,7
B	4,32	7,33	3,54	2,83	C	5,8	7,8	12,6	0,7	B	12,0	11,3	12,2	18,3
					C	5,7	7,2	12,0	1,1	B	12,4	11,2	12,5	19,5
					C	5,7	4,9	11,5	0,9	B	13,1	11,2	12,5	19,5
					C	5,4	3,1	10,8	1,6	B	13,2	11,1	12,6	20,0
					C	4,7	2,1	8,0	0,9	B	13,3	10,8	13,0	20,1
										C	13,3	10,7	13,2	20,4
										C	13,4	10,5	13,3	20,5
										C	15,0	10,4	13,6	20,7
										C	15,1	10,1	13,9	21,0
										C	15,2	10,0	14,2	21,2
										C	15,9	9,2	14,4	22,9
										C	16,7	9,0	15,4	24,0
										C	17,2	8,2	15,5	24,1
										C	17,9	6,6	15,7	26,1
										C	18,2	6,3	16,0	29,1
										C	19,4	5,5	16,6	29,1
										C	20,3	5,5	17,6	31,7

Окончание таблицы Е1

Варианты заданий																
4				5				6				7				
	a	b	c	d		a	b	c		a	b		a	b	c	d
A	15,5	18,8	59,9	0,8	A	0,55	0,62	2,12	A	9,4	21,1	A	9,4	21,1	12,5	21,0
A	21,2	23,9	62,7	33,9	A	1,32	2,68	4,77	A	8,4	8,2	A	8,4	8,2	12,1	5,0
A	15,5	25,3	56,9	67,4	A	1,12	2,36	2,34	A	4,9	16,6	A	4,9	16,6	12,1	24,1
A	19,0	26,4	46,2	87,5	A	0,93	1,40	2,79	A	13,4	10,5	A	13,4	10,5	15,4	14,8
A	12,4	26,4	86,2	88,2	A	0,95	1,66	2,85	A	11,1	18,4	A	11,1	18,4	2,0	29,1
A	11,6	27,2	34,5	98,0	A	0,87	0,87	1,05	A	3,6	11,3	A	3,6	11,3	9,5	20,1
A	15,8	27,8	44,6	108,1	A	0,51	0,02	1,12	B	20,3	19,1	B	15,1	18,3	12,1	11,2
B	9,0	28,0	52,7	109,0	B	0,99	1,50	4,85	B	11,2	5,5	B	13,2	11,2	11,0	19,5
B	11,9	29,6	46,3	117,6	B	1,44	2,05	4,54	B	15,1	18,3	B	12,0	19,4	5,8	20,5
B	24,1	30,7	65,0	135,7	B	1,44	3,63	4,75	B	13,2	11,2	B	11,6	10,0	16,0	22,9
B	13,5	31,0	51,5	135,8	B	0,14	2,45	4,59	B	12,0	19,4	B	18,2	11,2	12,1	17,7
B	18,4	31,6	52,5	142,5	B	1,55	1,76	1,41	B	11,6	10,0	B	12,4	11,1	16,6	13,8
B	21,2	32,2	71,7	146,0	B	0,72	0,33	4,29	C	18,2	11,2	C	15,2	6,3	15,5	9,3
B	7,1	32,2	68,5	148,3	B	1,02	1,29	3,42	C	12,4	11,1	C	9,8	16,3	11,5	20,4
C	9,9	32,8	56,6	159,0	C	0,92	3,08	4,87	C	15,2	6,3	C	8,0	18,8	9,3	16,7
C	15,1	33,1	65,8	161,1	C	0,92	1,16	5,40	C	7,6	11,9	C	12,0	10,1	14,4	18,3
C	20,0	33,2	34,2	174,4	C	2,09	4,45	5,12	C	9,8	16,3	C	13,3	17,6	12,2	29,1
C	15,0	35,3	70,4	198,5	C	1,71	1,09	4,23	C	8,0	18,8	C	12,0	6,6	13,2	13,1
C	15,9	36,0	71,9	200,6	C	0,85	1,30	3,87	D	12,0	10,1	D	16,7	14,7	6,2	19,5
C	18,7	36,6	45,4	235,4	C	1,27	2,40	4,43	D	13,3	17,6	D	7,3	18,2	8,1	20,0
C	4,1	39,3	45,6	261,8	C	0,46	0,72	4,80	D	12,0	6,6	D	8,7	5,5	13,9	12,4
					D	0,76	0,01	1,48	D	16,7	14,7	D	13,1	17,7	12,2	14,3
					D	0,59	3,39	2,65	D	7,3	18,2	D	7,6	11,9	5,3	15,5
					D	1,01	3,20	2,94	D	8,7	5,5	D	11,2	5,5	15,7	10,4
					D	0,54	1,79	4,30	E	13,1	17,7	E	8,5	10,4	14,2	31,7
					D	1,40	1,36	5,88	E	8,5	10,4	E	19,4	9,2	10,5	11,5
					D	1,12	2,24	4,06	E	19,4	9,2	E	10,6	10,7	13,6	21,2
					D	0,41	3,35	3,67	E	10,6	10,7	E	7,3	10,8	10,1	15,5
									E	7,3	10,8	E	17,9	13,8	13,0	15,5
									E	17,9	13,8	E	20,3	19,1	11,4	26,1
												F	17,2	15,6	13,3	15,7
												F	13,3	12,9	12,6	24,0
												F	10,2	14,7	12,5	20,7
												F	15,9	15,2	9,8	13,7
												F	15,0	14,8	17,6	15,0
												F	10,1	9,0	11,0	9,0

Приложение Ж (обязательное)

Таблица Ж1 Задания к лабораторной работе 15

1 задание					2 задание			3 задание		
Класс	1 призн.	2 призн.	3 призн.	4 призн.	Класс	1 призн.	2 призн.	Класс	1 призн.	2 призн.
a	1,01	0,48	0,87	1,12	d	0,247	0,295	a	9,23	5,26
a	1,14	1,11	1,38	1,17	d	0,491	0,495	c	19,82	14,00
a	1,22	1,44	1,02	0,79	c	0,768	0,240	a	11,59	8,06
a	1,09	0,94	0,91	1,1	c	0,838	0,354	a	11,06	6,99
a	0,65	0,89	1,09	1,12	c	0,921	0,320	a	7,58	2,83
a	1,21	1,22	1,11	1,4	c	0,837	0,409	a	9,35	5,73
a	0,95	1,04	0,76	1,15	b	0,642	0,309	a	10,27	6,24
b	2,01	1,24	0,78	0,8	b	0,754	0,288	d	25,95	20,80
b	1,64	0,93	0,96	1,48	c	0,844	0,340	a	11,04	6,83
b	2,13	1,38	1,15	0,93	d	0,433	0,242	b	16,03	11,53
a	1,41	0,85	1,15	0,73	c	0,923	0,289	e	31,41	26,44
a	1,22	1,03	1,1	0,87	c	0,824	0,243	b	16,92	11,96
b	2,81	1,13	1	0,65	c	0,963	0,335	b	16,02	11,47
b	2,05	1,25	0,91	0,88	c	0,800	0,223	b	17,74	12,96
b	2,31	0,87	1,14	1,37	a	0,828	0,696	b	17,00	11,96
b	1,8	0,84	0,94	0,89	c	1,000	0,366	b	18,36	13,36
a	1,15	1,27	1,25	1,16	b	0,715	0,354	b	15,73	11,25
a	0,74	2	0,86	0,85	a	0,885	0,573	b	17,98	13,18
a	0,89	2,05	0,9	0,87	b	0,658	0,278	b	17,33	12,11
a	0,68	1,39	0,71	0,99	a	0,739	0,632	a	12,62	8,28
a	0,94	2,07	1,09	0,78	b	0,684	0,357	c	21,71	16,34
a	1,06	1,92	1,01	0,82	a	0,684	0,592	c	22,13	16,88
a	0,76	2,27	1,09	1,17	d	0,512	0,428	e	30,27	26,10
a	1,12	1,86	0,9	1,26	a	0,763	0,777	c	20,92	15,50
a	1,11	1,9	1,33	0,93	d	0,452	0,339	e	30,17	25,93
a	1,03	1,25	0,91	1,02	d	0,468	0,428	c	22,83	17,16
a	1,57	2,1	0,94	1,09	c	0,849	0,298	c	21,02	15,63
a	1,05	0,56	1,82	0,95	c	0,761	0,252	c	21,84	16,83
a	0,92	0,93	2,17	1,17	d	0,298	0,201	e	31,82	26,62
a	1,11	0,9	2,36	1,03	c	0,820	0,438	d	25,06	19,97
a	1,44	1,29	2,11	0,95	c	0,886	0,252	d	26,06	21,15
a	0,81	1,03	2,36	1,12	a	0,761	0,534	c	20,60	15,17
a	0,69	1,29	1,79	0,92	c	0,827	0,361	d	24,51	19,51
a	1,04	0,83	2,76	1,26	a	0,894	1,000	e	28,94	23,35
a	0,98	1,17	1,62	0,93	d	0,381	0,459	d	26,12	21,25
a	1,23	0,88	2,69	0,63	b	0,750	0,301	d	23,90	18,94
a	1,25	1,11	2,52	1,38	a	0,613	0,512	d	23,74	18,34
c	1,26	0,95	0,66	2,44	d	0,443	0,369	d	25,30	20,05
c	1,03	1,27	1,39	2,15	c	0,853	0,231	c	20,51	15,11

Приложение И (обязательное)

Таблица И1 Задания к лабораторной работе 16

1 задание (4 класса.)		2 задание (3 класса)				3 задание (5 классов)	
1 признак	2 признак	1 признак	2 признак	3 признак	4 признак	1 признак	2 признак
0,25	0,30	24,4	38,0	2,6	7,0	13,16	8,54
0,49	0,50	19,6	30,2	2,0	6,2	10,40	6,32
0,77	0,24	27,7	40,7	2,7	7,4	10,44	6,47
0,84	0,35	23,8	35,8	2,4	6,7	9,89	5,97
0,92	0,32	19,8	30,6	2,1	6,2	12,81	8,39
0,84	0,41	28,3	41,8	2,9	7,5	7,80	3,41
0,64	0,31	22,2	33,1	2,3	6,3	12,18	8,07
0,75	0,29	21,1	31,5	2,2	6,2	13,27	8,75
0,84	0,34	30,4	44,1	3,0	7,8	8,17	4,48
0,43	0,24	22,1	32,8	2,2	6,3	18,36	13,35
0,92	0,29	30,7	44,6	3,1	8,0	24,73	19,53
0,82	0,24	59,1	83,4	5,5	12,9	15,54	11,24
0,96	0,33	6,2	15,6	0,2	3,6	17,65	12,60
0,80	0,22	39,5	61,2	3,8	9,7	17,41	12,19
0,83	0,70	46,0	67,4	4,3	10,8	17,90	13,17
1,00	0,37	46,3	67,7	4,4	10,8	16,41	11,74
0,72	0,35	36,9	56,7	3,6	9,1	15,74	11,28
0,89	0,57	41,2	61,9	4,0	9,9	16,83	11,82
0,66	0,28	47,7	69,2	4,4	11,1	20,40	14,73
0,74	0,63	47,0	68,8	4,4	11,0	20,60	15,14
0,68	0,36	46,7	68,7	4,4	11,0	20,72	15,45
0,68	0,59	46,6	67,9	4,4	10,9	19,93	14,38
0,51	0,43	45,7	67,3	4,3	10,7	22,09	16,87
0,76	0,78	38,7	59,4	3,8	9,4	20,94	15,53
0,45	0,34	43,7	64,7	4,2	10,5	21,40	16,02
0,47	0,43	43,5	63,7	4,2	10,4	16,33	11,73
0,85	0,30	42,7	62,9	4,1	10,3	21,34	16,00
0,76	0,25	60,8	86,9	5,7	13,5	21,67	16,27
0,30	0,20	60,1	85,1	5,6	13,3	21,00	15,58
0,82	0,44	58,1	81,5	5,4	12,7	30,33	26,10
0,89	0,25	54,0	78,3	5,2	12,0	25,62	20,30
0,76	0,53	65,3	93,2	6,1	14,5	26,03	21,01
0,83	0,36	66,6	93,8	6,2	14,6	24,12	19,06
0,89	1,00	54,0	78,4	5,2	12,0	25,35	20,11
0,38	0,46	64,2	92,4	6,0	14,1	24,47	19,23
0,75	0,30	59,7	84,9	5,6	13,3	24,17	19,19
0,61	0,51	56,5	81,0	5,4	12,6	25,80	20,43
0,44	0,37	65,2	93,2	6,1	14,3	23,85	18,81
0,85	0,23	60,5	86,1	5,6	13,5	29,33	24,94

Учебное издание

ЖУЧЕНКО Юрий Михайлович

**ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ
В БИОЛОГИИ И ХИМИИ
Лабораторный практикум**

Учебно-методическое пособие

Редактор В. И. Шкредова
Корректор В. В. Калугина

Подписано в печать 2010. Формат 60×84 1/6
Бумага офсетная. Ризография. Усл. печ. л.
Уч. – изд. л. Тираж **100** экз. Заказ №

Издатель и полиграфическое исполнение:
учреждение образования
«Гомельский государственный университет
имени Франциска Скорины»
ЛИ № 02330/0549481 от 14. 05. 2009.
ЛП № 02330/0150450 от 03. 02. 2009.
Ул. Советская, 104, 246019, г. Гомель