

**Министерство образования и науки Украины**  
**Донбасская государственная машиностроительная академия**

**Л. В. Васильева, В. Н. Черномаз,**  
**С. В. Малыгина, Е. А. Клеваник**

# **ПРИКЛАДНАЯ МАТЕМАТИКА**

## **ЛАБОРАТОРНЫЙ ПРАКТИКУМ**

**для студентов высших учебных заведений**

Утверждено на заседании  
ученого совета ДГМА  
Протокол № 2 от 29.10.2009г.

**Краматорск 2009**

**УДК 519.2**  
**ББК 22.17**  
**В 19**

**Рецензенты:**

**Чигарев В. В.**, д-р техн. наук, профессор, Приазовский государственный технический университет;

**Елисеева О. К.**, канд. техн. наук, доцент, Днепропетровский национальный университет.

Навчальний посібник містить теоретичний і практичний матеріал щодо комп'ютерної обробки даних інженерного експерименту. Запропонований ряд лабораторних робіт, виконуючи які під керівництвом викладача, студенти опановують методологію обробки інженерного експерименту.

Посібник може бути використаний також магістрами й аспірантами.

**Васильева, Л. В.**

**В 19** Прикладная математика : лабораторный практикум для студентов высших учебных заведений / Л. В. Васильева, В. Н. Черномаз, С. В. Малыгина, Е. А. Клеваник. – Краматорск : ДГМА, 2009. – 64 с. ISBN 978-966-379-367-2.

Учебное пособие содержит теоретический и практический материал по компьютерной обработке данных инженерного эксперимента. Предложен ряд лабораторных работ, выполняющих которые под руководством преподавателя, студенты осваивают методологию обработки инженерного эксперимента.

Пособие может быть использовано также магистрами и аспирантами.

**УДК 519.2**  
**ББК 22.17**

ISBN 978-966-379-367-2

© Л.В.Васильева, В.Н.Черномаз,  
С.В.Малыгина, Е.А.Клеваник, 2009  
© ДГМА, 2009

## СОДЕРЖАНИЕ

Введение.....	4
Лабораторная работа 1. Геометрическая вероятность. Статистическое определение вероятности .....	5
Лабораторная работа 2. Статистическая обработка одномерного случайного массива .....	11
Лабораторная работа 3. Прогноз на основании линейной регрессии. Точность прогноза. Теснота линейной связи.....	22
Лабораторная работа 4. Критерий согласия Пирсона.....	35
Приложение А. Краткие сведения о системе STATISTICA.....	44
Приложение Б. Пример выполнения лабораторной работы 2 в среде MathCad .....	51
Приложение В. Пример выполнения лабораторной работы 3 в среде MathCad .....	54
Приложение Г. Тематический план изучения дисциплины, распределение времени на ее изучение, тематика контрольных работ .....	59
Список рекомендуемой литературы .....	63

## ВВЕДЕНИЕ

Данная методическая разработка содержит три лабораторные работы, на выполнение каждой из которых отводится 5 часов: 1 час на подготовку группы и 4 часа машинного времени.

Выбор тем, вынесенных на проработку, определяется следующим:

1) должно быть усвоено и апробировано на практике статистическое понятие вероятности случайного события, поскольку при практической работе это, как правило, единственный способ приписать случайному событию конкретное значение вероятности;

2) должно быть получено представление о скорости статистической сходимости, чтобы в дальнейшей практической деятельности инженер мог предвидеть влияние объема выборки на точность получаемых оценок;

3) должно быть приобретено умение обрабатывать одномерный статистический массив: построение гистограммы, нахождение числовых характеристик и получение практических выводов, исходя из вида гистограммы и значений характеристик;

4) должно быть приобретено умение находить статистическую зависимость между неслучайным фактором  $x$  и случайным откликом на него  $y$ , умение делать по найденной зависимости прогноз для отклика  $y$  и оценивать точность этого прогноза. Оценка точности прогноза особенно важна, так как имеющиеся в распоряжении инженера статистические данные, как правило, сильно зашумлены.

Темы из пунктов 1 и 2 вынесены в первую лабораторную работу, темы из пункта 3 – во вторую, из пункта 4 – в третью работу.

Поскольку обработка статистических данных сопряжена с громоздкими и трудоёмкими вычислениями, необходимо использовать какой-либо математический пакет, специально предназначенный для их обработки. В настоящее время одной из лучших статистических систем считается STATISTICA. Последние версии системы STATISTICA полностью удовлетворяют всем стандартам Windows. Данная система реализует *графически ориентированный* подход к анализу данных, то есть большинство расчётов сопровождается графической иллюстрацией результатов расчёта. Пакет не русифицирован. Поэтому в дальнейшем рядом с английским термином в скобках будет даваться его русский перевод.

Тематический план изучения дисциплины, распределение времени на ее изучение, тематика контрольных работ даны в приложении Г.

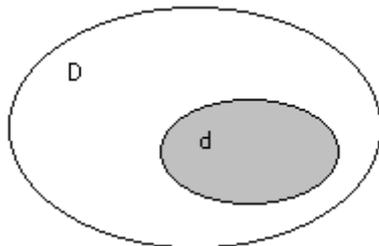
**Лабораторная работа 1**  
**ГЕОМЕТРИЧЕСКАЯ ВЕРОЯТНОСТЬ.**  
**СТАТИСТИЧЕСКОЕ ОПРЕДЕЛЕНИЕ ВЕРОЯТНОСТИ**

**1.1 Краткие теоретические сведения**

В данной лабораторной работе используются одновременно два определения вероятности случайного события – геометрическое и статистическое.

***1.1.1 Геометрическое определение вероятности***

Рассмотрим геометрическое определение для двумерного пространства. Пусть на плоскости имеется некоторая область  $D$ , площадь которой равна  $S^D$  и в ней содержится другая область  $d$  с площадью  $S^d$ . В область  $D$  наудачу бросается точка (рис. 1.1).



*Рисунок 1.1*

Какова вероятность того, что точка попадет в область  $d$ ? Здесь предполагается, что вероятность попадания в какую-либо часть области  $D$  пропорциональна площади этой части и не зависит от ее расположения и формы. В таком случае вероятность попадания в область  $d$  равна

$$P = \frac{S^d}{S^D}. \quad (1.1)$$

В случае одномерной и трехмерной области  $D$  вместо площади нужно говорить, соответственно, о длине и объеме.

***1.1.2 Статистическое определение вероятности***

Производится  $N$  испытаний, в каждом из которых может произойти случайное событие  $A$ . Подсчитывается число испытаний  $N_A$ , в которых событие  $A$  действительно произошло. Вероятность события  $A$  приближенно равна отношению:

$$P(A) \approx \frac{N_A}{N}. \quad (1.2)$$

Интуитивно ясно, что чем больше будет  $N$ , тем больше будет точность этой приближенной оценки.

### 1.1.3 Метод Монте-Карло вычисления геометрической вероятности

Предположим, что нужно вычислить площадь плоской фигуры  $A$  (рис. 1.2), расположенной в прямоугольнике

$$x_l \leq x \leq x_k,$$

$$y_l \leq y \leq y_k,$$

площадь которого равна  $S_I$ .

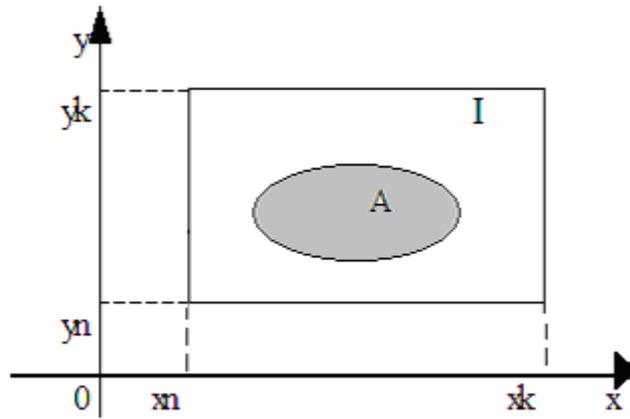


Рисунок 1.2

Внутри прямоугольника генерируются  $N$  случайных точек. Число точек, попавших в область  $A$ , обозначим  $N_A$ . Тогда

$$S(A) \approx S(I) \cdot \frac{N_A}{N}. \quad (1.3)$$

Метод Монте-Карло называется также **методом статистических испытаний**.

С увеличением количества точек  $N$  точность приближенного равенства растет, но медленно. Доказано, что погрешность приближенного равенства (1.2) имеет порядок, близкий к  $\frac{1}{\sqrt{N}}$ . При  $N \approx 1000$  это дает точность порядка 5...10%. Поэтому метод Монте-Карло целесообразно использовать при решении тех задач, в которых результат нужен с небольшой точностью.

## 1.2 Цель лабораторной работы

Должны быть приобретены следующие *умения*: нахождение вероятности суммы и произведения событий, оценка скорости статистической сходимости по результатам статистического эксперимента.

Должны быть усвоены следующие *понятия*: геометрическое определение вероятности, сумма событий, произведение событий, скорость статистической сходимости.

Работа рассчитана на 4 часа.

## 1.3 Задание к лабораторной работе

Используя данные из своего индивидуального задания, выполнить следующее:

- 1) изобразить на рисунке заданные области;
- 2) на отдельном рисунке изобразить области  $AB$ ,  $A+B$ , заштриховав их;
- 3) с помощью рисунка найти точные значения вероятностей  $P(A)$ ,  $P(B)$ ,  $P(AB)$ ,  $P(A+B)$ ;
- 4) найти значения этих вероятностей по методу Монте-Карло (файл MonteCarlo.exe). Оценить относительную погрешность вычислений вероятности по методу Монте-Карло по формуле:

$$\text{Отн.погр} = (P_{\text{точн}} - P_{\text{прибл}}) / P_{\text{точн}} * 100 \% ; \quad (1.4)$$

- 5) оценить скорость сходимости вероятности, вычисленной по методу Монте-Карло, к истинному значению, увеличив число случайных точек в 100 раз. Для этого усреднить относительные погрешности  $O(P_A)$  для  $N = 1000$  и  $N = 100\,000$  и по усредненным результатам сделать заключение о скорости сходимости.

## 1.4 Пример выполнения лабораторной работы с помощью программы MonteCarlo

Исходные данные приведены в таблице 1.1.

Таблица 1.1

Вариант	I	A	B
0	$0 \leq x \leq 10;$ $0 \leq y \leq 10$	$1 \leq x \leq 8;$ $1 \leq y \leq 8$	$(x - 5)^2 + (y - 5)^2 \leq 4$

Здесь область I – прямоугольник, A – прямоугольник, B – окружность.

Запускаем файл MonteCarlo.exe В появившемся окне (рис. 1.3) слева указываем типы заданных областей (в данном примере – прямоугольник и окружность), границы области I и предварительно рассчитанные по формуле (1.1) значения вероятностей  $P(A)$ ,  $P(B)$ ,  $P(AB)$ ,  $P(A+B)$ ;  $N = 1000$ .

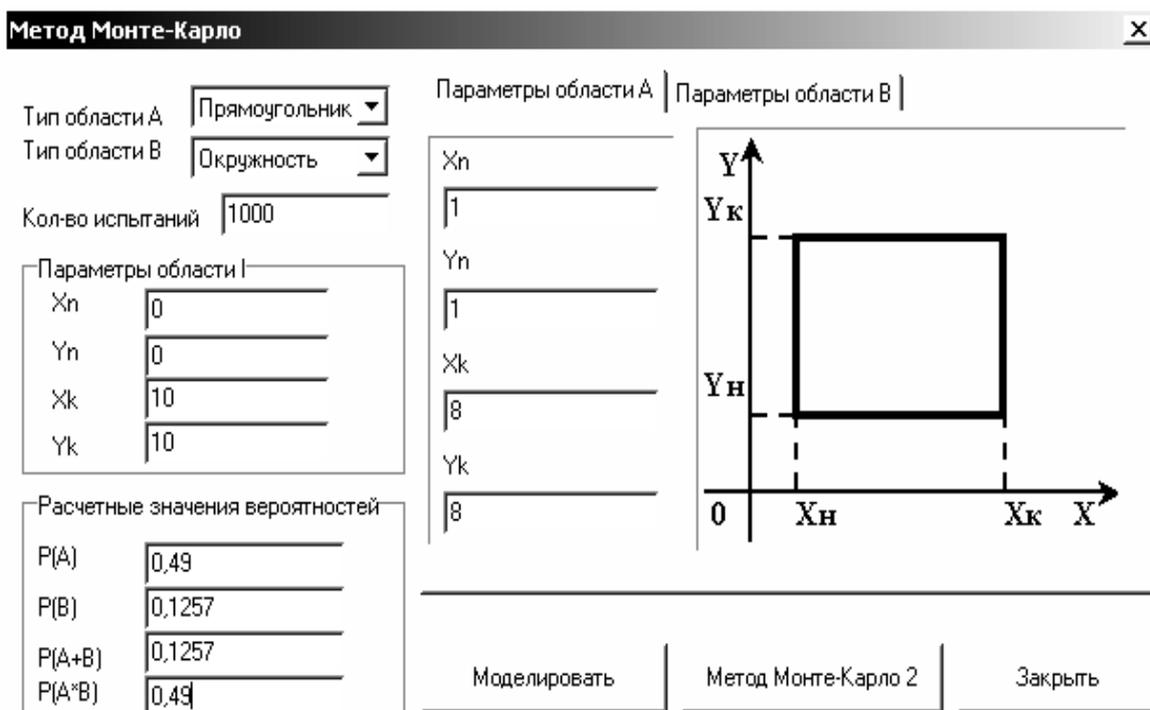


Рисунок 1.3

В правой части окна задаем параметры области А (рис. 1.3 ) и области В (рис. 1.4). Щелкаем кнопку *Моделировать*.

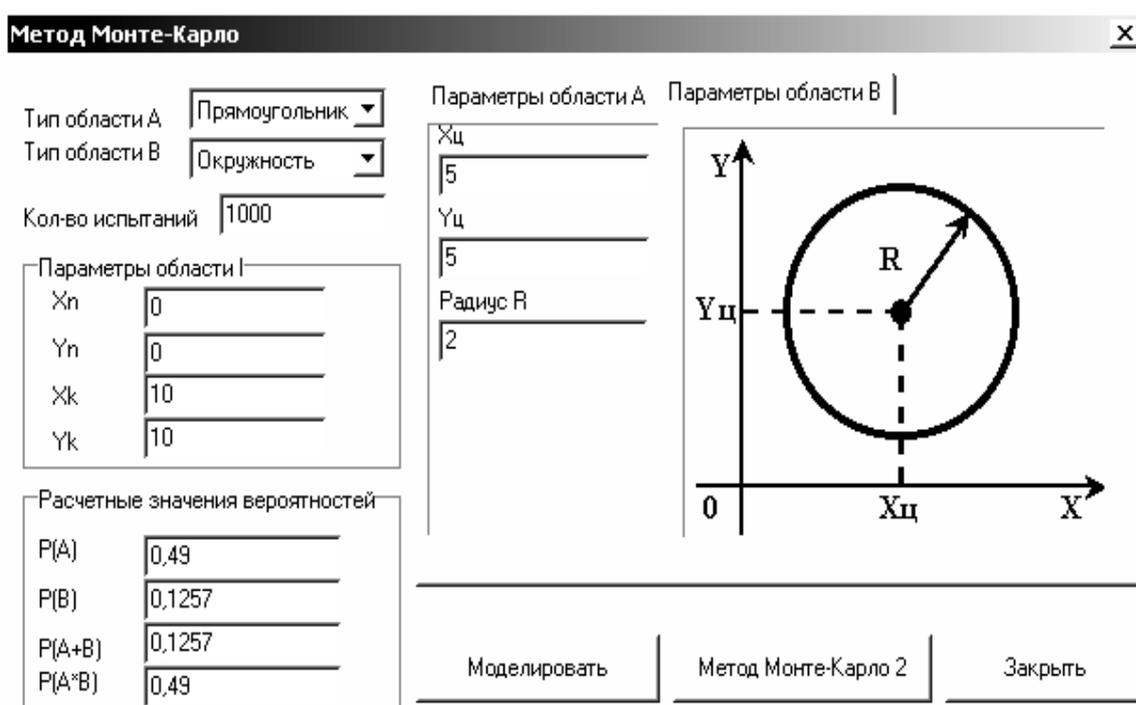


Рисунок 1.4

Сверяем вид ранее нарисованных областей с теми, что будут изображены в появившемся окне (рис. 1.5).

Щелкаем кнопку *Начать моделирование*. В результате расчета получим значения, как на рисунке 1.5.

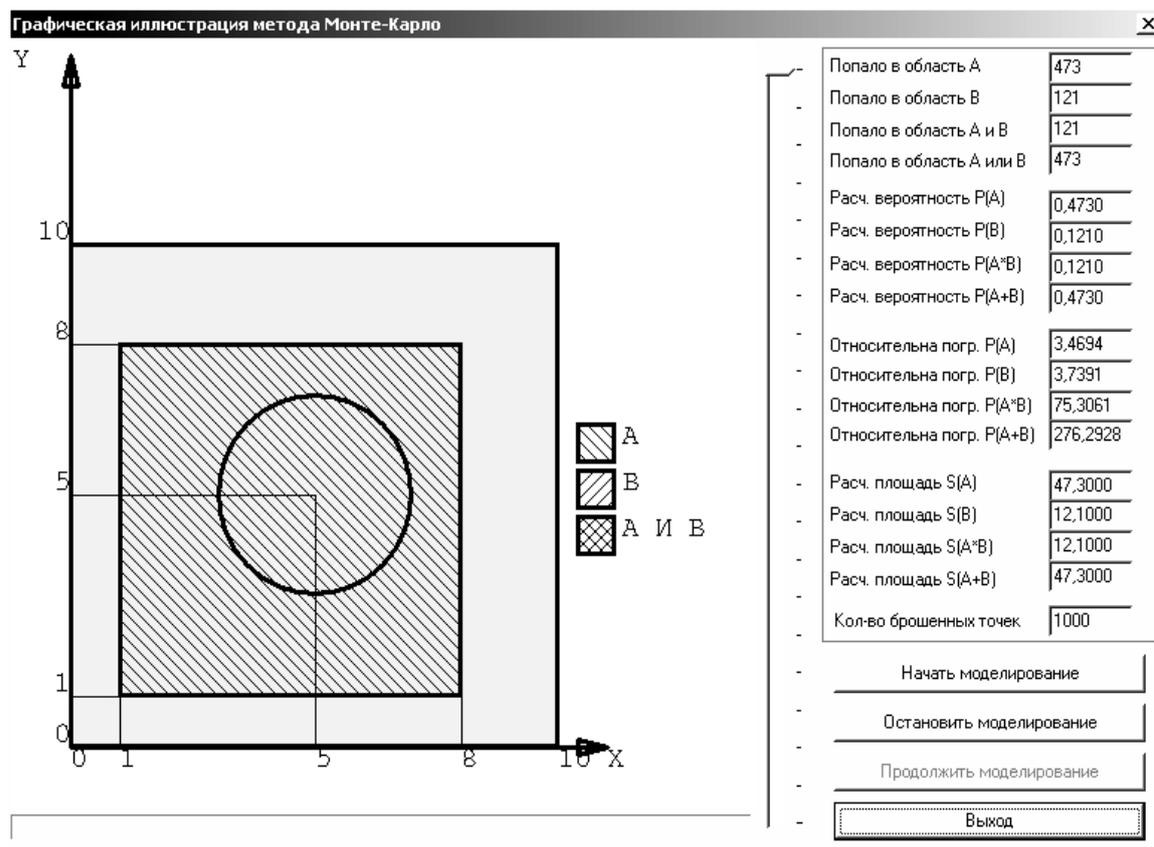


Рисунок 1.5

Расчетная вероятность  $P(A)$  вычисляется по формуле:

$$P(A) \approx \frac{N_A}{N} = \frac{473}{1000}. \quad (1.5)$$

Относительная погрешность (в процентах) для  $P(A)$  вычисляется по формуле:

$$\text{Отн. погр.} = \left| \frac{P(A) - P(A)_{\text{расчетная}}}{P(A)} \right| \cdot 100\%. \quad (1.6)$$

Остальные расчетные вероятности и относительные погрешности вычисляются аналогично.

Если увеличить количество точек в 100 раз ( $N = 100\,000$ ), то результаты расчетов уточнятся (табл. 1.2).

Таблица 1.2

N = 1 000		N = 100 000	
P(A)	O(PA)	P(A)	O(PA)
0,510	4,082	0,491	0,1327
0,498	1,633	0,488	0,3531
0,521	6,327	0,491	0,1306
0,487	0,612	0,492	0,4878
0,474	3,265	0,494	0,8041

Усредняем относительную погрешность для события А при  $N = 1\ 000$  и при  $N = 100\ 000$ . По полученным результатам определяем, на сколько порядков уменьшится относительная погрешность  $O(PA)$ , если количество точек  $N$  возрастет на два порядка с  $N = 1\ 000$  до  $N = 100\ 000$ .

### 1.5 Индивидуальные задания к лабораторной работе

Вар.	I	A	B
1	$2 \leq x \leq 10; 1 \leq y \leq 7$	$(x - 4)^2 + (y - 4)^2 \leq 4$	$2 \leq x \leq 8; 4 \leq y \leq 6$
2	$0 \leq x \leq 8; 0 \leq y \leq 6$	$3 \leq x \leq 5, 1 \leq y \leq 5$	$0 \leq x \leq 6; 3 \leq y \leq 4$
3	$4 \leq x \leq 11; 2 \leq y \leq 11$	$4 \leq x \leq 10, 6 \leq y \leq 10$	$(x - 6)^2/1 + (y - 6)^2/4 \leq 1$
4	$0 \leq x \leq 8, 0 \leq y \leq 6$	$(x - 4)^2/9 + (y - 3)^2/4 \leq 1$	$(x - 5)^2 + (y - 3)^2 \leq 1$
5	$2 \leq x \leq 9; 0 \leq y \leq 6$	$4 \leq x \leq 5; 0 \leq y \leq 3$	$(x - 7)^2 + (y - 4)^2 \leq 1$
6	$0 \leq x \leq 7, 0 \leq y \leq 5$	$(x - 4)^2 + (y - 2)^2 \leq 4$	$(x - 4)^2/4 + (y - 2)^2/1 \leq 1$
7	$0 \leq x \leq 8; 0 \leq y \leq 7$	$(x - 4)^2 + (y - 3)^2 \leq 4$	$(x - 4)^2/9 + (y - 3)^2/4 \leq 1$
8	$2 \leq x \leq 7, 0 \leq y \leq 4$	$3 \leq x \leq 6, 1 \leq y \leq 4$	$(x - 3)^2 + (y - 3)^2 \leq 1$
9	$0 \leq x \leq 8; 0 \leq y \leq 8$	$(x - 3)^2/9 + (y - 4)^2/4 \leq 1$	$(x - 7)^2 + (y - 4)^2 \leq 1$
10	$0 \leq x \leq 5, 0 \leq y \leq 10$	$(x - 2)^2 + (y - 7)^2 \leq 1$	$1 \leq x \leq 3; 0 \leq y \leq 7$
11	$1 \leq x \leq 10; 0 \leq y \leq 8$	$(x - 3)^2 + (y - 5)^2 \leq 4$	$(x - 3)^2 + (y - 4)^2 \leq 1$
12	$1 \leq x \leq 11, 2 \leq y \leq 12$	$(x - 6)^2/9 + (y - 6)^2/16 \leq 1$	$5 \leq x \leq 8, 5 \leq y \leq 6$
13	$0 \leq x \leq 12 2 \leq y \leq 10$	$1 \leq x \leq 8; 2 \leq y \leq 8$	$(x - 8)^2 + (y - 5)^2 \leq 4$
14	$1 \leq x \leq 9, 0 \leq y \leq 6$	$(x - 4)^2/9 + (y - 3)^2/4 \leq 1$	$(x - 4)^2 + (y - 2)^2 \leq 1$
15	$0 \leq x \leq 9; 0 \leq y \leq 8$	$1 \leq x \leq 7; 1 \leq y \leq 5$	$(x - 7)^2/1 + (y - 5)^2/4 \leq 1$
16	$0 \leq x \leq 8, 1 \leq y \leq 8$	$3 \leq x \leq 5, 1 \leq y \leq 5$	$(x - 4)^2 + (y - 4)^2 \leq 1$
17	$1 \leq x \leq 9; 1 \leq y \leq 10$	$(x - 5)^2 + (y - 5)^2 \leq 1$	$1 \leq x \leq 4; 1 \leq y \leq 3$
18	$0 \leq x \leq 8, 0 \leq y \leq 8$	$1 \leq x \leq 4, 1 \leq y \leq 5$	$(x - 4)^2 + (y - 3)^2 \leq 4$
19	$0 \leq x \leq 10; 0 \leq y \leq 10$	$(x - 5)^2/4 + (y - 3)^2/9 \leq 1$	$(x - 5)^2 + (y - 3)^2 \leq 4$
20	$1 \leq x \leq 10, 0 \leq y \leq 8$	$2 \leq x \leq 8, 2 \leq y \leq 6$	$4 \leq x \leq 6, 2 \leq y \leq 4$
21	$2 \leq x \leq 10; 0 \leq y \leq 8$	$(x - 7)^2 + (y - 3)^2 \leq 4$	$(x - 4)^2 + (y - 3)^2 \leq 1$
22	$0 \leq x \leq 10, 1 \leq y \leq 10$	$6 \leq x \leq 8, 3 \leq y \leq 4$	$(x - 7)^2 + (y - 3)^2 \leq 4$
23	$1 \leq x \leq 9; 1 \leq y \leq 11$	$(x - 4)^2 + (y - 3)^2 \leq 1$	$(x - 4)^2/4 + (y - 3)^2/1 \leq 1$
24	$2 \leq x \leq 6, 0 \leq y \leq 4$	$2 \leq x \leq 5, 1 \leq y \leq 4$	$(x - 3)^2 + (y - 2)^2 \leq 1$

## Лабораторная работа 2

### СТАТИСТИЧЕСКАЯ ОБРАБОТКА ОДНОМЕРНОГО СЛУЧАЙНОГО МАССИВА

#### 2.1 Краткие теоретические сведения

Большинство контролируемых параметров изделий относятся к нормально распределенным случайным величинам: размеры деталей, вес отливок, процентное содержание химических элементов в сплавах, емкость и сопротивление электротехнических изделий и т. д.

При исправном оборудовании и правильно отрегулированном технологическом процессе распределение контролируемого параметра должно быть нормальным, а его среднее значение должно совпадать со значением, заданным в технической документации. Возможные отклонения от этого требования и предполагаемые причины этих отклонений перечислены в таблице 2.1.

Таблица 2.1

№	Нарушение	Причина
1	Распределение контролируемого параметра близко к нормальному, но выборочное среднее не совпадает со значением, заданным технической документацией	Неправильно отрегулирован технологический процесс. Требуется регулировка
2	Распределение контролируемого параметра одномодально, но сильно отличается от нормального	Серьезные неисправности в оборудовании. Требуется ремонт
3	Распределение многомодально	Некачественная выборка, данные взяты из разных генеральных совокупностей. Повторить выборку

Чтобы установить, имеет ли место одно из перечисленных нарушений, производят статистический контроль интересующего параметра. Для этого производят  $n$  случайных его измерений  $x_1, x_2, \dots, x_n$  (выборка объема  $n$ ). По выборке находят следующие числовые характеристики:

- математическое ожидание –

$$\bar{x}^* = \frac{1}{n} \sum_i x_i, \quad (2.1)$$

- дисперсию –

$$D^* = \frac{1}{n} \sum_i (x_i - \bar{x}^*)^2, \quad (2.2)$$

- среднеквадратическое отклонение –

$$\sigma^* = \sqrt{D^*}, \quad (2.3)$$

- асимметрию –

$$As = \frac{1}{(\sigma^*)^3} \cdot \frac{1}{n} \sum_i (x_i - \bar{x}^*)^3, \quad (2.4)$$

- эксцесс –

$$Ek = \frac{1}{(\sigma^*)^4} \cdot \frac{1}{n} \sum_i (x_i - \bar{x}^*)^4 - 3. \quad (2.5)$$

Одномодальность или многомодальность выборочного распределения определяют по виду гистограммы. Близость закона распределения к нормальному определяют по значениям асимметрии и эксцесса.

Сравнение выборочного среднего  $\bar{x}^*$  со значением контролируемого параметра, заданного в технической документации, позволяет установить, правильно ли отрегулирован технологический процесс.

## 2.2 Цель лабораторной работы

Должны быть приобретены следующие *умения*: строить гистограммы; вычислять вероятности попадания случайной величины в заданный промежуток; вычислять вероятность отклонения случайной величины от математического ожидания не более, чем на заданную величину.

Должны быть усвоены следующие *понятия*: объем выборки, математическое ожидание, размах выборки, среднеквадратическое отклонение, асимметрия, эксцесс, теоретическая и эмпирическая плотность нормального распределения.

Работа рассчитана на 4 часа.

## 2.3 Задание к лабораторной работе

Используя данные из своего индивидуального задания, выполнить следующее:

- 1) создать в STATISTICA файл исходных данных под именем своего варианта (например, fio2.sta);
- 2) создать файл автоотчета new.rtf;
- 3) на основании файла lab2\_tv.rtf создать файл отчета по лабораторной работе – fio\_2.rtf;

4) на основании файла *fiо2.sta* создать статистическую таблицу со следующими данными: объем выборки (Valid N), математическое ожидание (Mean), размах выборки (Minimum – Maximum), среднее квадратическое отклонение (Std.Dev.), асимметрия (Skewness), эксцесс (Kurtosis);

5) построить гистограмму, выбрав число частичных интервалов равным 10. Отметить, удовлетворяет ли гистограмма предъявляемым к ней требованиям. Если не удовлетворяет, уменьшить, насколько это допустимо, число частичных интервалов;

6) по значениям асимметрии и эксцесса и виду гистограммы сделать заключение, значительно ли отличается распределение случайной величины от нормального;

7) с помощью вероятностного калькулятора найти вероятность попадания случайной величины в промежуток, заданный в варианте вашего задания;

8) считая, что технологический процесс отрегулирован правильно, а допуск составляет 10 % от значения контролируемого параметра, найти выпуск годной продукции в процентах.

## 2.4 Пример выполнения лабораторной работы в пакете Statistica

Работаем в модуле Basic Statistics and Tables. Порядок работы в системе STATISTICA описан в приложении А.

1 Создать в STATISTICA файл исходных данных под именем своего варианта (например, *fiо2.sta*).

2 Создать файл автоотчета *new.rtf*.

3 Загрузить файл *lab2\_tv.rtf* через пункт меню File – Open Other – Text/Output File. Сохранить этот файл под именем *fiо\_2.rtf*.

4 На основании файла *fiо2.sta* создать статистическую таблицу со следующими данными: объем выборки (Valid N), математическое ожидание (Mean), размах выборки (Minimum – Maximum), среднее квадратическое отклонение (Std.Dev.), асимметрия (Skewness), эксцесс (Kurtosis). Для этого выделить таблицу с данными – Basic Statistics – Analysis – Descriptive statistics (базовые статистики – анализ – описательные статистики) – More statistics (больше статистик) – активировать опции *Valid N* (объем выборки), *Mean* (среднее значение), *Standard Deviation* (среднее квадратическое отклонение), *Minimum* и *Maximum*, *Skewness* (асимметрия), *Kurtosis* (эксцесс) – ОК – Variables (Переменные) – выделить нужную переменную (в данном случае X) – ОК – ОК.

Нужная таблица (табл. 2.2) вставляется в отчет по лабораторной работе из автоотчета:

Таблица 2.2

```

+-----+
| STAT.  | Descriptive Statistics (var00.sta) |
| BASIC  |                                     |
| STATS  |                                     |
+-----+
| Variable | Valid N | Mean | Minimum | Maximum | Std.Dev. | Skeness |
+-----+
|      x  |      80 | 2,457750 | ,310000 | 4,170000 | ,751823 | -229292 |
+-----+
| STAT.  | Descriptive Statistics (var00.sta) |
| BASIC  |                                     |
| STATS  |                                     |
+-----+
| Variable | Kurtosis |
+-----+
|      x  | ,118629 |
+-----+

```

5 Построить гистограмму. Для этого нужно активировать таблицу fio2.sta – пункт верхнего меню Graphs – Stats2D Graphs – Histogramm. Выбрать опции Regular, Off и число частичных интервалов (Categories) равным 10. Получим гистограмму (рис. 2.1).

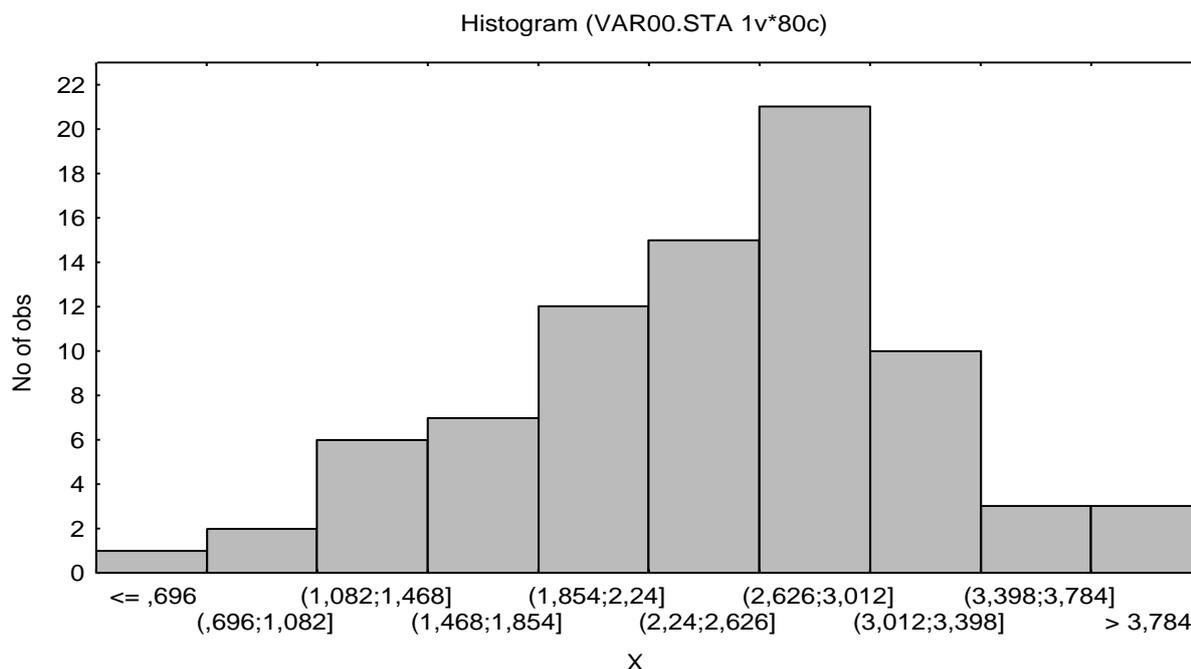


Рисунок 2.1

Под каждым столбцом гистограммы записан интервал изменения переменной, соответствующий данному столбцу. Высота столбцов равна количеству наблюдений, попавших в соответствующий частичный интервал. Данная гистограмма требованиям, предъявляемым к гистограммам, не удовлетворяет (объяснить почему). Уменьшить число частичных интервалов до 5 и построить новую гистограмму (рис. 2.2).

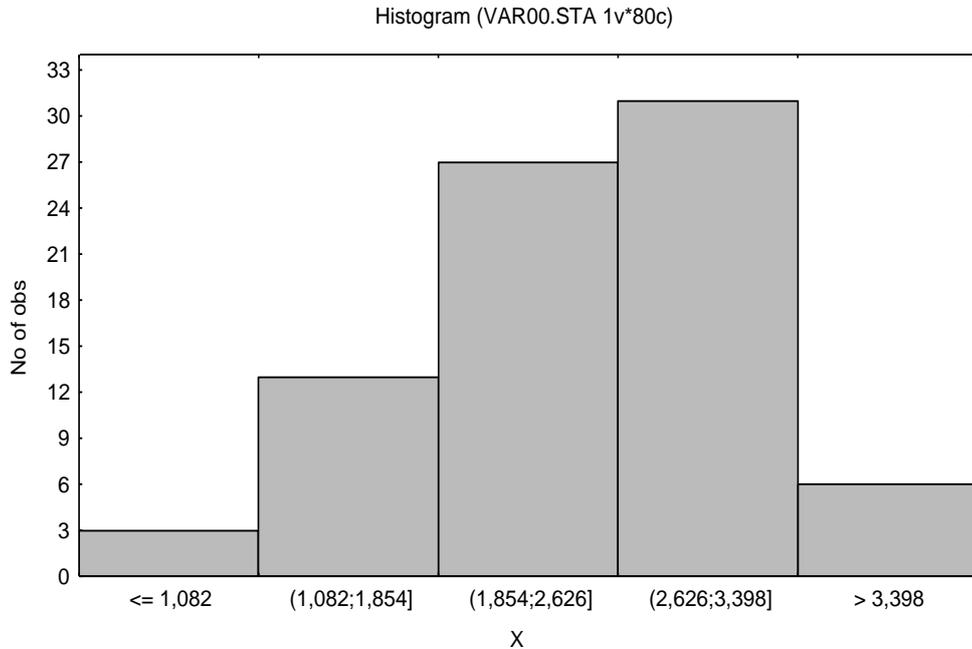


Рисунок 2.2

Ответьте на вопрос: удовлетворяет ли данная гистограмма требованиям, предъявляемым к гистограммам? Можно ли еще уменьшить число частичных интервалов?

6 По значениям асимметрии и эксцесса и виду гистограммы сделать заключение, значительно ли отличается распределение случайной величины от нормального.

7 С помощью вероятностного калькулятора найти вероятность попадания случайной величины в промежуток, заданный в варианте вашего задания. Вероятность попадания случайной величины в промежуток (a, b) рассчитывается по формуле  $P(a < X < b) = F(b) - F(a)$ , где  $F(x)$  – функция распределения случайной величины. Значения функции распределения находим с помощью вероятностного калькулятора: Basic Statistic / Tables – Probability Calculator – Distribution Z(Normal) (рис. 2.3). В поля **mean**, **st.dev.** и **Z** вписываем соответствующие значения, нажимаем на кнопку **Compute** и получаем нужную вероятность:  $F(2,1) = 0,3170$ . Аналогично находим  $F(3,2)$ .

$$P(2,1 < X < 3,2) = F(3,2) - F(2,1) = 0,8381 - 0,3170 = 0,5211.$$

Так как технологический процесс отрегулирован правильно, то выборочное среднее  $\bar{x}$  можно принять за значение параметра, заданного в технической документации. Десятипроцентное отклонение находим по формуле  $\delta = 0,1 \cdot \bar{x} = 0,1 \cdot 2,485$ . Затем вычисляется вероятность  $P(|x - \bar{x}| < \delta) = P(|x - 2,485| < 0,245) = 0,2554$  с помощью вероятностного калькулятора, как показано на рисунке 2.4. Получаем ответ: выход годной продукции при заданном допуске  $0,1\bar{x}$  составляет 25,5 % от всей продукции.

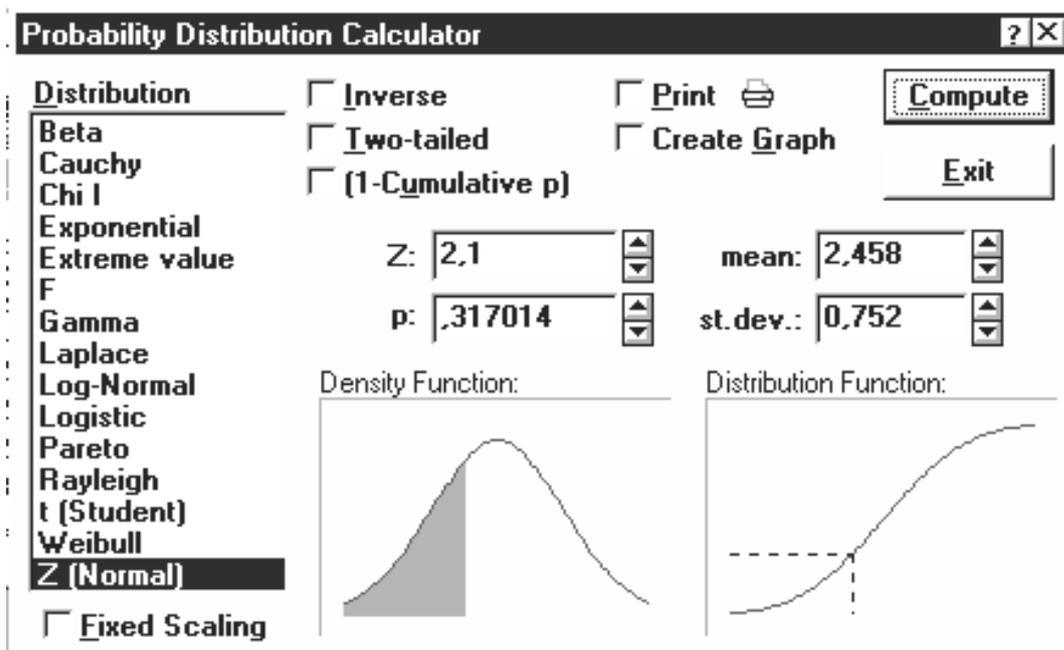


Рисунок 2.3

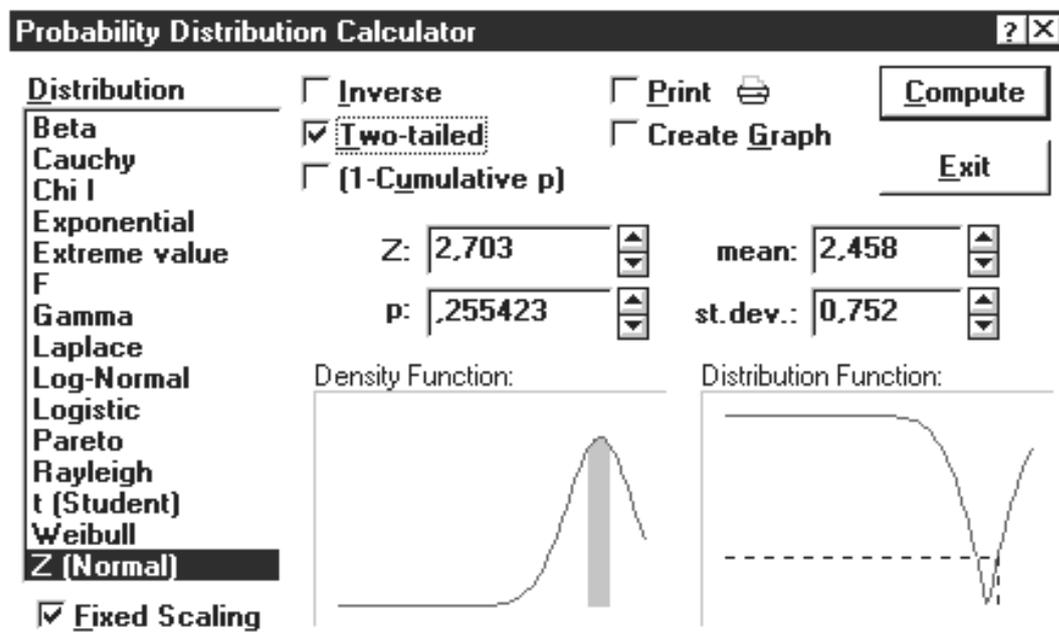


Рисунок 2.4

Примечание. Перед распечаткой отчета рисунки 2.3, 2.4 из текста отчета можно удалить.

## 2.5 Индивидуальные задания к лабораторной работе

### Вариант 1

1.67	2.41	0.79	1.41	2.50	2.29	2.58	1.32
3.75	1.94	0.95	3.48	2.39	1.17	1.92	1.04
2.13	1.58	2.18	2.30	3.03	1.50	2.53	1.91
1.31	3.62	1.49	1.98	2.14	3.35	2.89	2.51
2.31	2.34	1.00	2.03	0.64	2.67	0.09	1.78
3.24	1.91	1.20	1.61	2.35	1.73	2.93	2.32
2.84	1.29	2.28	2.54	1.85	2.40	2.22	2.90
2.37	2.68	2.00	2.70	2.33	2.86	0.36	1.98
2.53	0.80	2.89	0.73	1.01	1.85	2.05	1.16
1.76	2.78	2.43	1.85	1.21	1.53	1.54	2.43

$$P(0.93 < X < 1.52) = ?$$

### Вариант 2

2.46	1.70	2.44	0.82	1.50	2.53	2.32	2.61
1.35	3.78	1.97	0.98	3.51	2.42	1.20	1.95
1.07	2.16	1.61	2.21	2.33	3.06	1.53	2.56
1.94	1.34	3.63	1.52	2.01	2.17	3.38	2.92
2.54	2.34	2.37	1.03	2.06	0.67	2.70	1.12
1.81	3.27	1.94	1.23	1.64	2.38	1.76	2.96
2.35	2.87	1.32	2.31	2.57	1.88	2.43	1.88
2.93	2.40	2.71	2.03	2.76	2.36	2.89	0.39
2.01	2.56	0.83	2.92	0.76	1.04	1.88	2.08
1.19	1.79	2.81	2.46	1.88	1.24	1.56	1.57

$$P(0.92 < X < 1.54) = ?$$

### Вариант 3

1.60	2.49	1.73	2.47	0.85	1.53	2.56	2.35
2.64	1.38	3.81	2.00	1.01	3.54	2.45	1.23
1.98	1.10	2.19	1.64	2.24	2.36	3.09	1.56
2.59	1.97	1.37	3.68	1.55	2.04	2.20	3.41
2.95	2.57	2.37	2.40	1.06	2.09	0.70	2.73
0.45	1.84	3.30	1.97	1.26	1.67	2.41	1.79
2.99	2.38	2.90	1.35	2.34	2.60	1.91	2.46
2.28	2.96	2.43	2.74	2.06	2.76	2.39	2.92
0.42	2.04	2.59	0.86	2.95	0.79	1.07	1.91
2.11	1.22	1.82	2.84	2.49	1.91	1.27	1.59

$$P(0.91 < X < 1.55) = ?$$

**Вариант 4**

1.62	1.63	2.52	1.76	2.50	0.88	1.56	2.59
2.38	2.67	1.14	3.84	2.03	1.04	3.57	2.48
1.86	2.01	1.13	2.22	1.67	2.27	2.38	3.12
1.59	2.62	2.00	1.40	3.71	1.58	2.07	2.23
3.44	2.98	2.60	2.40	2.43	1.09	2.12	0.73
2.76	0.18	1.87	3.32	2.00	1.29	1.70	2.44
1.82	3.02	2.41	2.93	1.38	2.37	2.63	1.94
2.49	2.31	2.99	2.46	2.77	2.09	2.79	2.42
2.95	0.45	2.07	2.62	0.89	2.98	0.82	1.10
1.94	2.14	1.25	1.83	2.87	2.52	1.94	1.30

$P(0.90 < X < 1.56) = ?$

**Вариант 5**

3.31	1.15	1.43	2.27	2.47	1.58	2.18	3.20
2.85	2.27	1.63	1.95	1.96	2.85	2.09	2.83
1.21	1.89	2.92	2.71	3.00	1.74	4.17	2.36
1.37	3.90	2.81	1.59	2.34	1.46	2.55	2.00
2.60	2.78	3.45	1.92	2.95	2.33	1.73	4.04
1.91	2.40	2.56	3.77	3.31	2.93	2.73	2.76
1.42	2.45	1.06	3.09	0.31	2.20	3.66	2.33
1.62	2.03	2.77	2.15	3.35	2.74	3.26	1.71
2.70	2.96	2.27	2.82	2.64	3.32	2.79	3.10
2.42	3.12	2.75	3.28	0.78	2.40	2.95	1.22

$P(0.89 < X < 1.57) = ?$

**Вариант 6**

2.46	1.70	2.44	0.82	1.50	2.53	2.32	2.61
1.35	3.78	1.97	0.98	3.51	2.42	1.20	1.95
1.07	2.16	1.61	2.21	2.33	3.06	1.53	2.56
1.94	1.34	3.63	1.52	2.01	2.17	3.38	2.92
2.54	2.34	2.37	1.03	2.06	0.67	2.70	1.12
3.24	1.91	1.20	1.61	2.35	1.73	2.93	2.32
2.84	1.29	2.28	2.54	1.85	2.40	2.22	2.90
2.37	2.68	2.00	2.70	2.33	2.86	0.36	1.98
2.53	0.80	2.89	0.73	1.01	1.85	2.05	1.16
1.76	2.78	2.43	1.85	1.21	1.53	1.54	2.43

$P(1.08 < X < 1.68) = ?$

**Вариант 7**

1.67	2.41	0.79	1.41	2.50	2.29	2.58	1.32
3.75	1.94	0.95	3.48	2.39	1.17	1.92	1.04
2.13	1.58	2.18	2.30	3.03	1.50	2.53	1.91
1.31	3.62	1.49	1.98	2.14	3.35	2.89	2.51
2.31	2.34	1.00	2.03	0.64	2.67	0.09	1.78
1.81	3.27	1.94	1.23	1.64	2.38	1.76	2.96
2.35	2.87	1.32	2.31	2.57	1.88	2.43	1.88
2.93	2.40	2.71	2.03	2.76	2.36	2.89	0.39
2.01	2.56	0.83	2.92	0.76	1.04	1.88	2.08
1.19	1.79	2.81	2.46	1.88	1.24	1.56	1.57

$P(1.07 < X < 1.69) = ?$

**Вариант 8**

1.62	1.63	2.52	1.76	2.50	0.88	1.56	2.59
2.38	2.67	1.14	3.84	2.03	1.04	3.57	2.48
1.86	2.01	1.13	2.22	1.67	2.27	2.38	3.12
1.59	2.62	2.00	1.40	3.71	1.58	2.07	2.23
3.44	2.98	2.60	2.40	2.43	1.09	2.12	0.73
0.45	1.84	3.30	1.97	1.26	1.67	2.41	1.79
2.99	2.38	2.90	1.35	2.34	2.60	1.91	2.46
2.28	2.96	2.43	2.74	2.06	2.76	2.39	2.92
0.42	2.04	2.59	0.86	2.95	0.79	1.07	1.91
2.11	1.22	1.82	2.84	2.49	1.91	1.27	1.59

$P(1.06 < X < 1.70) = ?$

**Вариант 9**

1.60	2.49	1.73	2.47	0.85	1.53	2.56	2.35
2.64	1.38	3.81	2.00	1.01	3.54	2.45	1.23
1.98	1.10	2.19	1.64	2.24	2.36	3.09	1.56
2.59	1.97	1.37	3.68	1.55	2.04	2.20	3.41
2.95	2.57	2.37	2.40	1.06	2.09	0.70	2.73
2.76	0.18	1.87	3.32	2.00	1.29	1.70	2.44
1.82	3.02	2.41	2.93	1.38	2.37	2.63	1.94
2.49	2.31	2.99	2.46	2.77	2.09	2.79	2.42
2.95	0.45	2.07	2.62	0.89	2.98	0.82	1.10
1.94	2.14	1.25	1.83	2.87	2.52	1.94	1.30

$P(1.05 < X < 1.71) = ?$

**Вариант 10**

1.43	2.03	3.05	2.70	2.13	1.48	1.80	1.81
2.70	1.94	3.63	1.06	1.74	2.77	2.56	2.85
1.59	3.08	2.21	1.22	3.75	2.66	1.44	3.19
1.81	2.40	1.85	2.45	2.57	3.30	1.77	2.80
3.18	1.58	2.89	1.76	2.25	2.41	3.62	2.13
1.95	0.45	2.07	2.62	0.89	2.98	0.82	1.10
1.94	2.14	1.25	1.83	2.87	2.52	1.94	1.30
1.62	1.03	2.77	2.15	2.35	2.74	3.26	1.71
2.70	2.96	2.27	1.82	2.64	3.32	2.79	3.10
2.42	1.12	2.75	1.28	0.78	2.40	2.95	1.22

$P(1.04 < X < 1.72) = ?$

**Вариант 11**

2.46	1.70	1.43	2.27	2.47	1.58	2.32	2.61
1.35	3.78	1.63	1.95	1.96	2.85	1.20	1.95
1.07	2.16	1.87	3.32	2.00	1.29	1.53	2.56
1.94	1.34	2.41	2.93	1.38	2.37	3.38	2.92
2.54	2.34	2.99	2.46	2.77	2.09	2.70	1.12
3.24	1.91	2.07	2.62	0.89	2.98	2.93	2.32
2.84	1.29	1.25	1.83	2.87	2.52	2.22	2.90
2.37	2.68	2.77	2.15	3.35	2.74	0.36	1.98
2.53	0.80	2.27	2.82	2.64	3.32	2.05	1.16
1.76	2.78	2.75	3.28	0.78	2.40	1.54	2.43

$P(1.23 < X < 1.83) = ?$

**Вариант 12**

3.31	1.15	2.44	0.82	1.50	2.53	2.18	3.20
2.85	2.27	1.97	0.98	3.51	2.42	2.09	2.83
2.76	0.18	1.61	2.21	2.33	3.06	1.70	2.44
1.82	3.02	3.63	1.52	2.01	2.17	2.63	1.94
2.49	2.31	2.37	1.03	2.06	0.67	2.79	2.42
2.95	0.45	1.20	1.61	2.35	1.73	0.82	1.10
1.94	2.14	2.28	2.54	1.85	2.40	1.94	1.30
1.62	2.03	2.00	2.70	2.33	2.86	3.26	1.71
2.70	2.96	2.89	0.73	1.01	1.85	2.79	3.10
2.42	3.12	2.43	1.85	1.21	1.53	2.95	1.22

$P(1.22 < X < 1.84) = ?$

**Вариант 13**

1.60	2.49	0.79	1.41	2.50	2.29	2.56	2.35
2.64	1.38	0.95	3.48	2.39	1.17	2.45	1.23
1.98	1.10	2.18	2.30	3.03	1.50	3.09	1.56
2.59	1.97	1.49	1.98	2.14	3.35	2.20	3.41
2.95	2.57	1.00	2.03	0.64	2.67	0.70	2.73
2.76	0.18	1.94	1.23	1.64	2.38	1.70	2.44
1.82	3.02	1.32	2.31	2.57	1.88	2.63	1.94
2.49	2.31	2.71	2.03	2.76	2.36	2.79	2.42
2.95	0.45	0.83	2.92	0.76	1.04	0.82	1.10
1.94	2.14	2.81	2.46	1.88	1.24	1.94	1.30

$P(1.21 < X < 1.85) = ?$

**Вариант 14**

0.67	2.41	1.73	2.47	0.85	1.53	1.58	1.32
0.75	1.94	3.81	2.00	1.01	3.54	1.92	1.04
2.13	1.58	2.19	1.64	2.24	2.36	2.53	1.91
1.31	3.62	1.37	3.68	1.55	2.04	2.89	2.51
2.31	0.34	2.37	2.40	1.06	2.09	0.09	1.78
1.81	2.27	1.87	3.32	2.00	1.29	1.76	2.96
2.35	2.87	2.41	2.93	1.38	2.37	2.43	1.88
2.93	2.40	2.99	2.46	1.77	2.09	2.89	0.39
2.01	2.56	2.07	2.62	0.89	2.98	1.88	2.08
1.19	1.79	1.25	1.83	2.87	2.52	1.56	1.57

$P(1.20 < X < 1.86) = ?$

**Вариант 15**

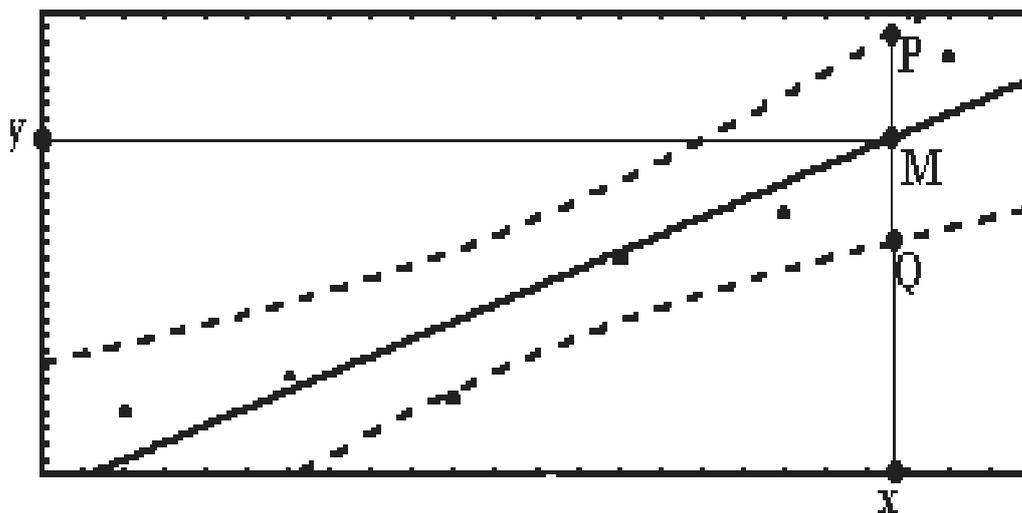
1,62	1,63	2,52	1,76	2,50	0,88	1,56	2,59
2,38	2,67	3,14	3,84	2,03	1,04	3,57	2,48
2,86	2,01	3,13	2,22	1,67	2,27	2,38	3,12
1,59	2,62	2,00	1,40	3,71	1,58	2,07	2,23
0,42	2,04	2,59	0,86	2,95	0,79	1,07	1,91
2,11	3,22	1,82	2,84	2,49	2,91	3,27	1,59
1,76	2,50	1,88	3,08	2,47	2,99	1,44	2,43
2,69	2,00	2,55	2,37	3,05	2,52	1,83	2,15
2,85	3,48	3,01	0,51	2,13	2,68	0,95	3,04
0,88	1,16	2,00	3,20	1,31	1,91	2,93	2,58

$P(1.19 < X < 1.87) = ?$

**Лабораторная работа 3**  
**ПРОГНОЗ НА ОСНОВАНИИ ЛИНЕЙНОЙ РЕГРЕССИИ.**  
**ТОЧНОСТЬ ПРОГНОЗА. ТЕСНОТА ЛИНЕЙНОЙ СВЯЗИ**

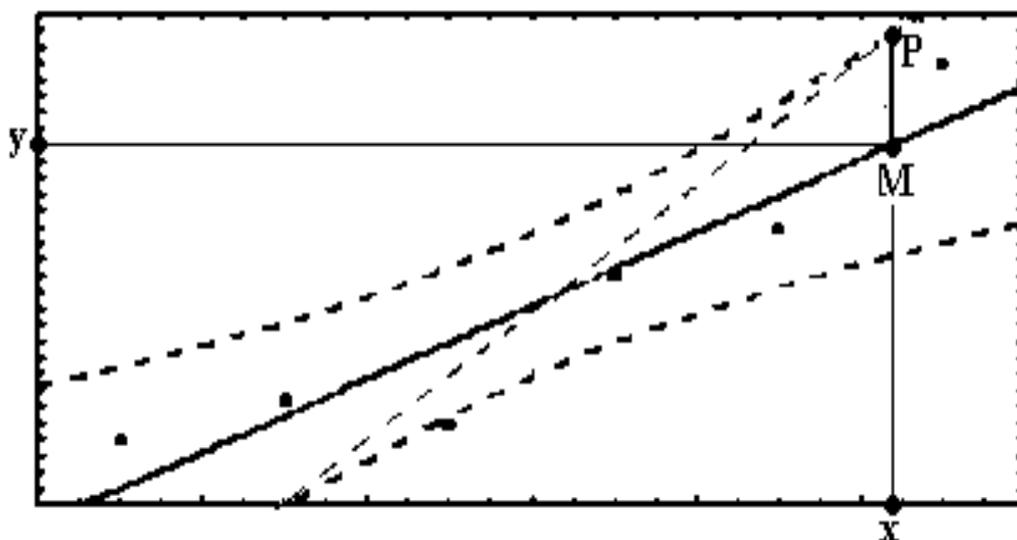
**3.1 Краткие теоретические сведения**

Уравнение линейной регрессии  $y = b_0 + b_1x$  находится по выборке методом наименьших квадратов. На рисунке 3.1 это наклонная прямая, изображенная сплошной линией.



*Рисунок 3.1*

Линейная зависимость между  $x$  и  $y$ :  $y = \beta_0 + \beta_1x + \varepsilon$  ( $\varepsilon$  – случайный член), которая может быть построена по генеральной совокупности, неизвестна. Можно только утверждать, что она с вероятностью  $\gamma$  расположена в доверительной области, ограниченной пунктирными линиями (рис. 3.2).



*Рисунок 3.2*

Вероятность  $\gamma$  называется *уровнем доверия*. Обычно  $\gamma = 0,95$  или  $\gamma = 0,99$  (95 %, 99 %). Точная линия регрессии  $y = \beta_0 + \beta_1x + \varepsilon$  изображена на рис. 3.2 пунктирной прямой. Прогноз  $y$  в точке  $x$  делают по уравнению  $y = b_0 + b_1x$ . Точное значение прогноза может с вероятностью  $\gamma$  соответствовать любой точке доверительного интервала PQ (см. рис. 3.1).

В наиболее неблагоприятном случае прогноз попадает на край доверительной области. В этом случае абсолютная погрешность прогноза равна полуширине доверительного интервала  $\delta = MP = MQ$ . Относительная погрешность прогноза в процентах вычисляется по формуле:

$$\text{Отн. погр.} = \left| \frac{\delta}{y_{\text{прогноза}}} \right| \cdot 100\%. \quad (3.1)$$

### 3.2 Цель лабораторной работы

Должны быть приобретены следующие *умения*:

- 1) нахождение графика и уравнения линейной регрессии и доверительной области при заданном уровне доверия  $\gamma$ ;
- 2) нахождение по графику прогноза отклика  $y$  при заданном значении фактора  $x$ ;
- 3) нахождение по графику максимальной абсолютной погрешности прогноза;
- 4) расчет максимальной относительной погрешности прогноза в процентах;
- 5) оценка тесноты линейной связи по значению коэффициента корреляции.

Должны быть усвоены следующие *понятия*: корреляционное поле, область прогнозов, прогноз, доверительная область, доверительный интервал, уровень доверия, полуширина доверительного интервала, абсолютная и относительная погрешности прогноза, коэффициент корреляции, теснота линейной связи.

Работа рассчитана на 4 часа.

### 3.3 Задание к лабораторной работе

Используя данные из своего индивидуального задания:

- 1 Построить графики линии регрессии с 80 %, 95 % и 99 % доверительными областями.
- 2 Нанести вручную на линию регрессии центр рассеяния.
- 3 Найти по графику интервальный прогноз в точке, соответствующей центру рассеяния для всех трех значений уровня доверия (80 %, 95 %, 99 %), а также прогноз в любой произвольной точке из области прогнозов.

4 Найти по графику полуширину доверительного интервала  $\delta_\gamma$  в точке, соответствующей центру рассеяния для всех трех значений уровня доверия (80 %, 95 %, 99 %):  $\delta_{80}$ ,  $\delta_{95}$ ,  $\delta_{99}$ .

5 Оценить максимальную относительную ошибку прогноза (в процентах) для всех трех значений уровня доверия (80 %, 95 %, 99 %) по формуле

$$\left| \frac{\delta_\gamma}{y_{\text{прогноза}}} \right| \cdot 100\% \quad (\delta_\gamma \text{ и } y_{\text{прогноза}} \text{ находим по чертежу}).$$

6 Сделать вывод о взаимосвязи уровня доверия  $\gamma$  и относительной погрешности прогноза.

7 Найти коэффициент корреляции и оценить по нему тесноту линейной связи.

### 3.4 Пример выполнения лабораторной работы в пакете Statistica

Работаем в модуле Basic Statistics and Tables.

1 *Создать таблицу данных.* Таблица будет иметь две переменные: X – фондоотдача и Y – рентабельность. Затем создать автоотчет и файл отчета в личной папке. Полученные результаты будут переноситься из автоотчета в отчет.

2 *Внести в автоотчет созданную таблицу данных.* Для этого активировать таблицу данных и нажать клавишу F11. На запрос компьютера нажать кнопку ОК. Появится новая таблица From: \*.sta, которую можно закрыть, и в автоотчете появится таблица данных, подобная таблице 3.1.

Таблица 3.1

```

+-----+-----+-----+-----+
| STA|From: lab1.sta (2v * 10c) |
| BAS|
| STA|Variables: 1-2, Cases: 1-10|
+-----+-----+-----+-----+
|   |   X   |   Y   |
+-----+-----+-----+-----+
| 1 | 1,033 | 1,830 |
| 2 |   ,012 |   ,580 |
| 3 |   ,045 | 1,340 |
| 4 |   ,243 | 1,340 |
| 5 |   ,266 | 1,640 |
| 6 |   ,302 | 1,650 |
| 7 |   ,451 | 1,910 |
| 8 | 1,041 | 1,960 |
| 9 | 1,423 | 2,080 |
|10 | 1,914 | 2,180 |
+-----+-----+-----+-----+

```

3 Построить графики линейной регрессии с 80 %, 95 % и 99 % доверительными областями. Graphs – Stats2D Graphs – Scatterplots – Variables (X, Y) – Linear – Confidence Bands (Базовые статистики – графики – плоские графики – точечные графики – доверительные области) – On 0,80 – ОК – Next – 0,95 – ОК – Next – 0,99 – ОК.

Получим три графика: уровень доверия 80 % (рис. 3.3), уровень доверия 95 % (рис. 3.4), уровень доверия 99 % (рис. 3.5).

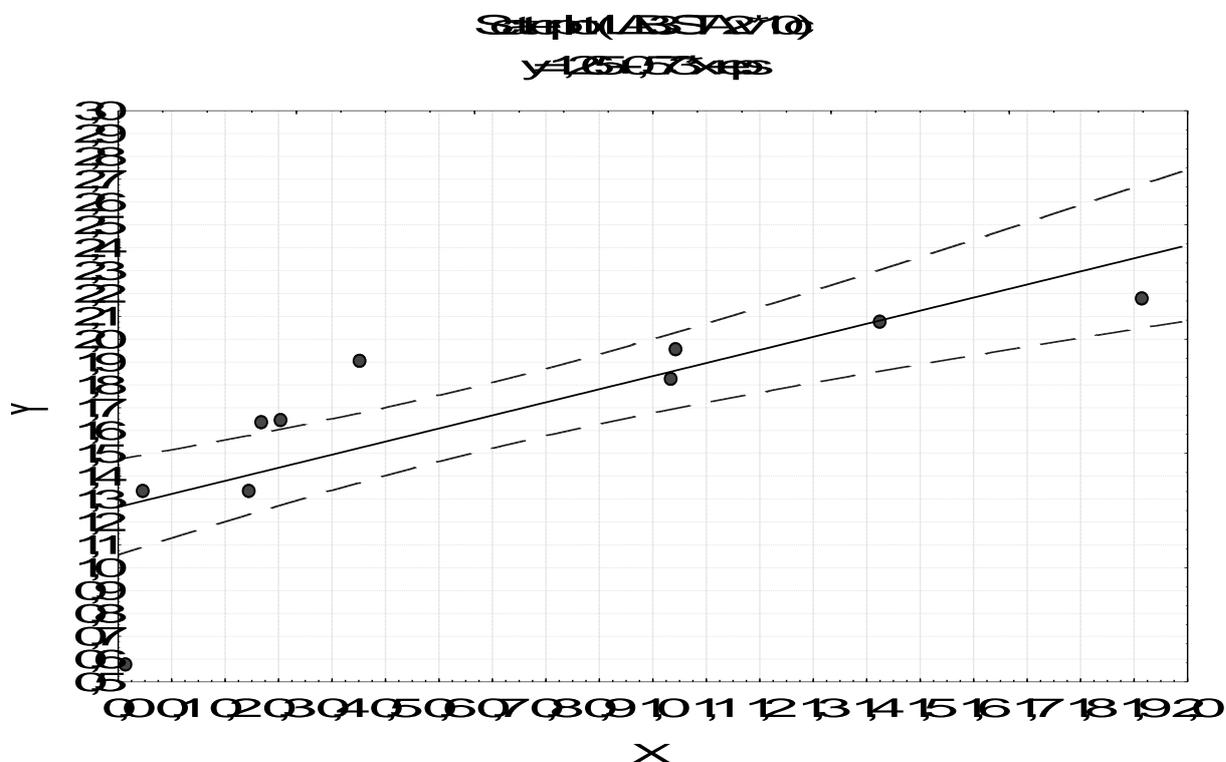


Рисунок 3.3

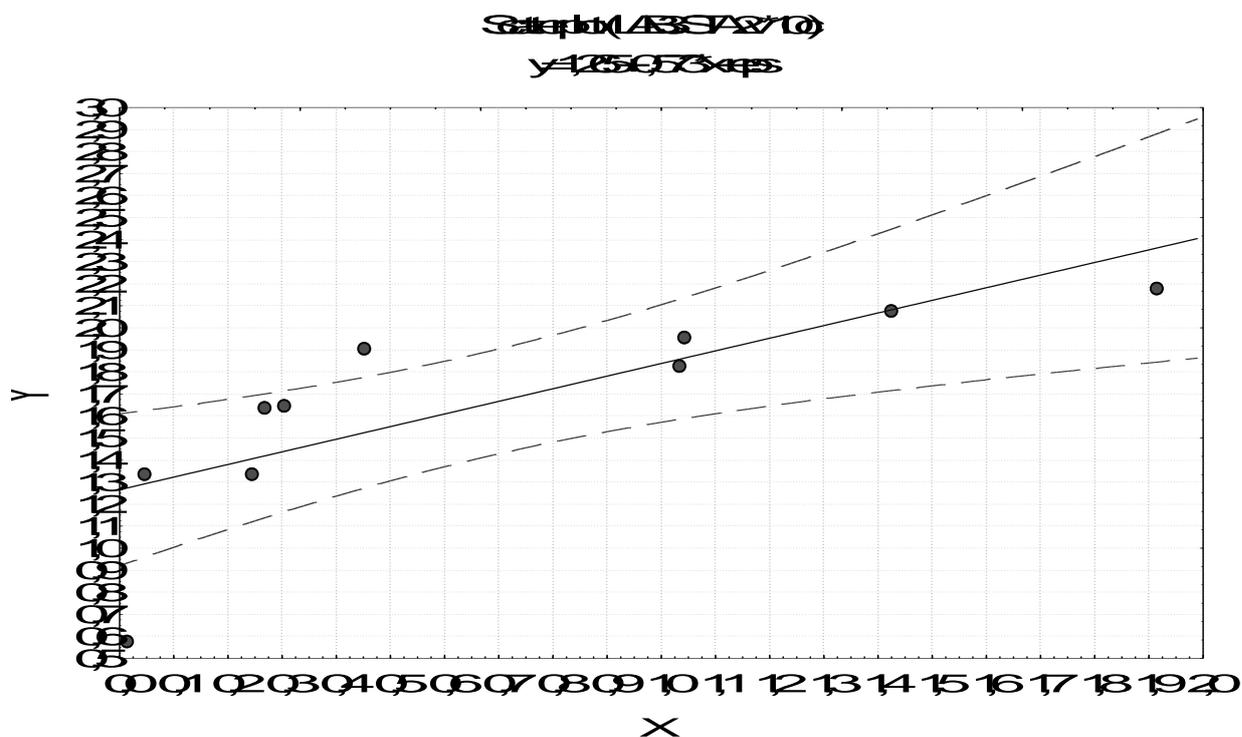


Рисунок 3.4

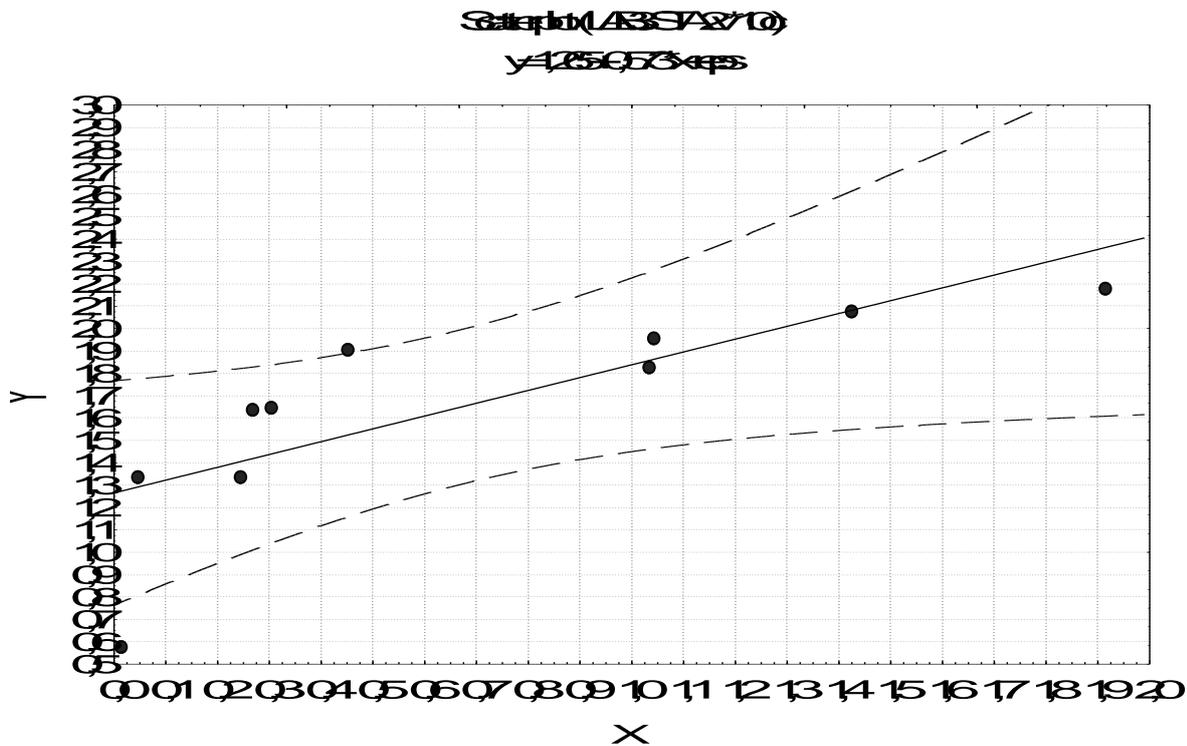


Рисунок 3.5

4 Найти координаты центра рассеяния и область прогнозов. Активировать таблицу с данными – Basic Statistics – Analysis – Descriptive statistics (описательные статистики) – More statistics (Больше статистик) – активировать опции *Mean* (среднее значение), *Minimum* и *Maximum* – ОК – *Variables* (переменные) – выделить обе переменные (в данном случае X и Y) – ОК – ОК. Требуемые статистики выводятся в автоотчет в виде таблицы 3.2:

Таблица 3.2

```

+-----+-----+-----+-----+
| STAT.  | Descriptive Statistics (new.sta) |
| BASIC  |                                     |
| STATS  |                                     |
+-----+-----+-----+-----+
| Variable | Mean | Minimum | Maximum |
+-----+-----+-----+-----+
|      x  | ,673000 | ,012000 | 1,914000 |
|      y  | 1,651000 | ,580000 | 2,180000 |
+-----+-----+-----+-----+

```

Нужная таблица вставляется в отчет по лабораторной работе из автоотчета через буфер.

Средние значения (Mean) дают координаты центра рассеяния  $X = 0,673$  и  $Y = 1,651$ . Выполнение пунктов 2–6 задания производится вручную и с помощью калькулятора.

Область прогнозов задается интервалом (Xmin; Xmax), в данном примере (0,012; 1,914).

5 Найдите коэффициент корреляции. Для этого создаем корреляционную таблицу: активируем таблицу данных – Analysis – Correlation matrices – Variables for analysis X,Y (анализ – корреляционные матрицы – переменные для анализа X, Y). Результат получаем в виде таблицы 3.3.

Таблица 3.3

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+
| STAT.   |Correlations (new.sta)
| BASIC   |Marked correlations are significant at p < ,05000
| STATS   |N=9 (Casewise deletion of missing data)
|+-----+-----+-----+-----+-----+-----+-----+-----+
| Variable |           X           |           Y           |
|+-----+-----+-----+-----+-----+-----+-----+-----+
|           X |           1,00           |           ,78           |
|           Y |           ,78           |           1,00           |
|+-----+-----+-----+-----+-----+-----+-----+-----+

```

Коэффициент корреляции для переменных X, Y равен 0,78. Так как  $0,6 < 0,78 < 0,9$ , то линейная связь между этими переменными достаточная.

### 3.5 Индивидуальные задания к лабораторной работе

#### Вариант 1

Фондоотдача и уровень рентабельности по метизному заводу за год характеризуются следующими данными (табл. 3.4):

Таблица 3.4

Номер завода	Фактор	Показатель
	Фондоотдача, грн	Уровень рентабельности, %
1	1,24	39,4
2	0,63	23,2
3	1,18	37,2
4	1,12	35,1
5	0,44	20,0
6	1,19	37,9
7	0,48	20,1
8	0,65	23,4
9	0,26	13,4
10	0,75	24,8
11	1,03	32,2
12	0,89	30,2
13	0,16	10,3
14	0,67	23,7
15	0,90	31,3

### **Вариант 2**

Фондоотдача и уровень рентабельности по вагоноремонтному заводу за год характеризуются следующими данными (табл. 3.5):

*Таблица 3.5*

Номер завода	Фактор	Показатель
	Фондоотдача, грн	Уровень рентабельности, %
1	38,9	10,7
2	33,3	11,3
3	37,7	12,2
4	31,1	12,4
5	29,4	10,9
6	37,2	11,3
7	35,6	11,1
8	34,1	14,0
9	0,26	6,8
10	22,8	7,1
11	21,7	8,9
12	26,	4,2
13	23,3	7,4
14	24,5	11,4
15	29,9	4,8

### **Вариант 3**

Фондоотдача и уровень рентабельности по металлоремонтным цехам заводов области за год характеризуются следующими данными (табл. 3.6):

*Таблица 3.6*

Номер завода	Фактор	Показатель
	Фондоотдача, грн	Уровень рентабельности, %
1	5,46	27,6
2	5,53	24,9
3	7,05	32,1
4	7,29	37,1
5	7,40	36,9
6	7,10	33,4
7	6,25	31,3
8	8,64	39,3
9	5,18	24,8
10	1,81	20,0
11	2,30	25,5
12	5,53	26,4
13	2,22	20,3
14	3,54	29,1
15	3,23	27,7

#### **Вариант 4**

Фондоотдача и уровень рентабельности по хлебозаводам области за год характеризуются следующими данными (табл. 3.7):

*Таблица 3.7*

Номер завода	Фактор	Показатель
	Фондоотдача, грн	Уровень рентабельности, %
1	20,1	12,2
2	64,2	17,6
3	61,1	17,5
4	13,3	10,3
5	10,8	12,8
6	17,2	13,1
7	34,1	16,9
8	32,3	14,4
9	27,8	16,0
10	24,2	16,4
11	55,5	18,3
12	17,1	10,8
13	11,1	10,0
14	25,5	14,0
15	31,1	16,1

#### **Вариант 5**

Фондоотдача и уровень рентабельности по станциям ТО области за год характеризуются следующими данными (табл. 3.8):

*Таблица 3.8*

Номер станции	Фактор	Показатель
	Фондоотдача, грн	Уровень рентабельности, %
1	1,25	9,2
2	2,32	14,7
3	1,71	10,3
4	1,64	10,0
5	1,38	9,9
6	1,18	9,1
7	1,44	9,8
8	1,17	6,4
9	1,72	13,0
10	2,21	11,8
11	1,64	13,2
12	1,73	11,4
13	1,17	8,1
14	1,39	9,0
15	1,07	11,1

### **Вариант 6**

Фондоотдача и уровень рентабельности по плодоконсервным заводам области за год характеризуются следующими данными (табл. 3.9):

*Таблица 3.9*

Номер завода	Фактор	Показатель
	Фондоотдача, грн	Уровень рентабельности, %
1	1,08	20,1
2	1,05	12,9
3	0,99	18,0
4	1,02	11,7
5	0,98	17,9
6	1,04	16,8
7	1,03	15,6
8	1,10	14,3
9	1,03	18,1
10	0,89	17,8
11	0,78	13,0
12	0,99	14,2
13	1,43	24,2
14	1,03	20,0
15	1,05	19,3

### **Вариант 7**

Фондоотдача и уровень рентабельности по хлебозаводам области за год характеризуются следующими данными (табл. 3.10):

*Таблица 3.10*

Номер завода	Фактор	Показатель
	Фондоотдача, грн	Уровень рентабельности, %
1	33,4	12,3
2	29,1	14,7
3	25,3	10,9
4	27,1	16,1
5	43,3	22,3
6	47,2	21,1
7	49,3	24,3
8	35,7	13,3
9	45,8	27,6
10	43,4	28,3
11	42,1	25,1
12	40,1	20,2
13	33,3	13,7
14	41,2	19,9
15	34,0	14,2

### **Вариант 8**

В табл. 3.11 приведены данные об удельном весе механизированных работ и производительности труда по плодоовощным заводам области за год.

*Таблица 3.11*

Номер завода	Фактор	Показатель
	Удельный вес механизированных работ, %	Производительность труда, грн
1	84	4300
2	83	4150
3	67	3000
4	63	3420
5	69	3300
6	70	4300
7	73	3420
8	81	4100
9	77	3700
10	72	3500
11	80	4000
12	85	4450
13	83	4270
14	70	3300
15	87	4500

### **Вариант 9**

В табл. 3.12 приведены данные об удельном весе простоев оборудования и уровне рентабельности по молокозаводам области за год.

*Таблица 3.12*

Номер завода	Фактор	Показатель
	Удельный вес простоев оборудования, %	Уровень рентабельности, %
1	18,1	9,5
2	7,8	19,4
3	7,4	8,7
4	6,4	18,3
5	7,8	16,4
6	17,1	8,8
7	10,2	17,8
8	14,1	13,7
9	20,0	7,0
10	16,7	10,2
11	16,0	10,4
12	20,4	7,3
13	16,2	10,7
14	16,0	14,0
15	20,1	7,3

### Вариант 10

В таблице 3.13 приведены данные об уровне технической подготовки рабочих и уровне заработной платы по сахарным заводам области за год.

Таблица 3.13

Номер завода	Фактор	Показатель
	Удельный вес рабочих с технической подготовкой, %	Заработная плата за месяц, грн
1	40	142,20
2	33	152,33
3	37	154,20
4	39	149,95
5	37	154,37
6	41	149,80
7	49	170,11
8	38	168,33
9	55	193,30
10	43	172,72
11	56	189,39
12	47	187,01
13	44	173,40
14	55	187,87
15	54	184,20

### Вариант 11

Фондоотдача и уровень рентабельности по метизным заводам за год характеризуются следующими данными (табл. 3.14):

Таблица 3.14

Номер завода	Фактор	Показатель
	Фондоотдача, грн	Уровень рентабельности, %
1	20,0	2,0
2	12,8	1,8
3	9,2	1,1
4	5,3	3,5
5	18,6	10,1
6	10,8	3,3
7	28,7	24,2
8	13,8	1,9
9	28,6	20,8
10	22,9	19,2
11	14,0	3,4
12	13,0	2,7
13	12,8	1,4
14	25,0	20,1
15	13,8	7,8

### **Вариант 12**

Фондоотдача и уровень рентабельности по вагоноремонтному заводу за год характеризуются следующими данными (табл. 3.15):

*Таблица 3.15*

Номер завода	Фактор	Показатель
	Фондоотдача, грн	Уровень рентабельности, %
1	80,0	20,0
2	87,2	37,5
3	90,8	43,4
4	94,7	45,6
5	81,4	23,4
6	89,2	25,0
7	71,3	17,2
8	86,2	33,3
9	71,4	15,0
10	77,7	18,7
11	86,0	24,8
12	87,0	34,5
13	87,2	33,1
14	75,0	19,2
15	86,2	31,8

### **Вариант 13**

В таблице 3.16 приведены данные об удельном весе в товарообороте потребительской кооперации продукции собственного производства и уровне рентабельности предприятий области за год.

*Таблица 3.16*

Номер предприятия	Фактор	Показатель
	Удельный вес продукции собственного производства, %	Уровень рентабельности, %
1	25,2	9,5
2	58,2	9,4
3	42,2	8,7
4	46,8	8,3
5	60,5	6,4
6	66,1	8,8
7	26,5	7,8
8	59,9	13,7
9	43,2	7,0
10	47,8	6,7
11	61,8	10,4
12	68,1	7,3
13	32,0	8,9
14	60,2	9,4
15	44,2	7,3

### **Вариант 14**

В таблице 3.17 приведены данные об относительном уровне издержек обращения и уровне рентабельности по магазинам промышленных товаров за год.

*Таблица 3.17*

Номер магазина	Фактор	Показатель
	Относительный уровень издержек обращения, %	Уровень рентабельности, %
1	7,89	8,9
2	14,41	4,3
3	6,01	10,2
4	9,17	4,9
5	6,78	8,3
6	8,91	7,8
7	6,17	13,1
8	10,11	4,9
9	5,98	13,3
10	6,10	10,7
11	5,90	13,7
12	8,13	5,6
13	9,01	4,7
14	6,00	11,1
15	6,13	10,8

## Лабораторная работа 4 КРИТЕРИЙ СОГЛАСИЯ ПИРСОНА

### 4.1 Краткие теоретические сведения

*Критериями согласия* называют специальные случайные величины, назначение которых – проверять гипотезу о виде распределения.

Критерий согласия Пирсона является универсальным. Для его применения нужны выборки больших объемов ( $n \geq 100$ ).

Пусть дана выборка:  $x_1, x_2, \dots, x_n$ . По ней рассчитывают наблюдаемое значение критерия  $\chi_{\text{набл}}^2$ , по статистическим таблицам находят критическое значение критерия  $\chi_{\text{кр}}^2$ . В зависимости от того, как соотносятся эти две величины, гипотезу о виде распределения либо принимают, либо отвергают.

#### 4.1.1 Последовательность действий

1 Выдвигают гипотезу о законе распределения случайной величины, из которой взята выборка. Размах выборки от  $x_{\min}$  до  $x_{\max}$  разбивают на  $s$  непересекающихся частей (рис. 4.1), которые могут быть различной длины:  $m_1 + m_2 + \dots + m_s = n$ . Подсчитывают, сколько наблюдений  $m_i$  попало в  $i$ -й отрезок разбиения. Величины  $m_1, m_2, \dots, m_s$  называют *выборочными частотами*.

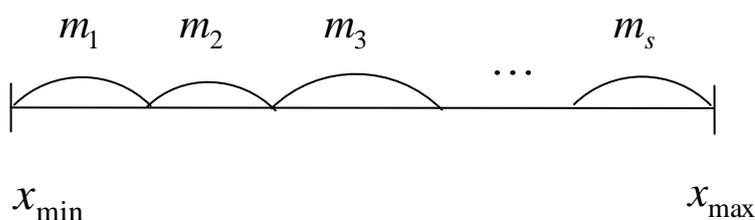


Рисунок 4.1

2 Вычисляют теоретические частоты  $m_i'$ :

$$m_i' = np_i, \quad (4.1)$$

где  $p_i$  – вероятность попадания случайных величин  $x$  в интервал  $i$ . Эта величина зависит от того, какой теоретический закон распределения проверяется.

Очевидно, что если гипотеза о виде закона распределения верна, то  $m_i \approx m_i'$ .

3 Вычисляют статистику

$$\chi_{набл}^2 = \sum_{i=1}^s \frac{(m_i - m_i')^2}{m_i}, \quad (4.2)$$

которая при  $n \rightarrow \infty$  стремится к распределению  $\chi^2$  с  $\nu = s - r - 1$  степенями свободы, где  $r$  – число параметров теоретического распределения  $F(x)$ , оцениваемых на основании выборки, а  $s$  – число интервалов, на которые разбито все множество наблюдаемых данных (табл. 4.1).

Таблица 4.1

Вид распределения	Число степеней свободы
нормальное	$s - 1 - 2 = s - 3$
показательное	$s - 1 - 1 = s - 2$
равномерное	$s - 1 - 0 = s - 1$

4 Для нахождения критического значения  $\chi_{кр}^2$  используют статистические таблицы  $\chi^2(\alpha, n)$  или вероятностный калькулятор в программе Statistica, где  $\alpha$  – уровень значимости гипотезы.

5 Если  $\chi_{набл}^2 < \chi_{кр}^2$ , то гипотеза о виде распределения принимается. Если  $\chi_{набл}^2 > \chi_{кр}^2$  – отвергается.

## 4.2 Цель лабораторной работы

Должны быть приобретены следующие *умения*: построение гистограмм в паре с кривыми распределения; разбиение размаха выборки на частичные интервалы произвольной длины; нахождение критического значения  $\chi_{кр}^2$  для заданного уровня значимости; расчет теоретических частот распределения.

Должны быть усвоены следующие *понятия*: критерий согласия, выборочные (эмпирические) частоты, теоретические частоты, статистическая гипотеза, уровень значимости гипотезы, число степеней свободы.

Работа рассчитана на 4 часа.

### 4.3 Задание к лабораторной работе

Используя данные из своего индивидуального задания, выполнить следующее:

1 Создать в STATISTICA файл исходных данных под именем своего варианта (например, var1.sta).

2 Создать файл автоотчета.

3 На основании файла данных создать статистическую таблицу со следующими данными: объем выборки (Valid N), математическое ожидание (Mean), среднее квадратическое отклонение (Std.Dev).

4 По файлу данных построить 3 гистограммы. Гистограммы построить в паре со следующими кривыми распределения: нормальное, экспоненциальное, логнормальное. По виду рисунков сделать заключение о том, можно ли однозначно утверждать, какой вид распределения соответствует выборке.

5 Проверить соответствие вида распределения выборке с помощью тестов Колмогорова-Смирнова и  $\chi^2$ .

### 4.4 Пример выполнения лабораторной работы в пакете Statistica

Работаем в модуле Basic Statistics and Tables.

1 Создаем в STATISTICA файл исходных данных под именем своего варианта (например, var1.sta).

2 Создаем файл автоотчета.

3 Создаем статистическую таблицу со следующими данными: объем выборки (Valid N), матем. ожидание (Mean), среднее квадратическое отклонение (Std.Dev). Для этого выделить таблицу с данными – Basic Statistics – Analysis – Descriptive statistics (Базовые статистики – анализ – описательные статистики) – More statistics (Больше статистик) – активировать опции Valid N (объем выборки), Mean\_(среднее значение), Standard Deviation (среднее квадратическое отклонение), ОК – Variables (переменные) – выделить нужную переменную (в данном случае X) – ОК – ОК .

Нужная таблица (табл. 4.2) вставляется в отчет по лабораторной работе из автоотчета:

Таблица 4.2

Descriptive Statistics (lab4-1.sta)			
Variable	Valid N	Mean	Std.Dev.
x	80	2,228500	,771271

4 *Строим гистограмму.* Для этого нужно активировать таблицу данных – пункт верхнего меню Graphs – Stats2D Graphs – Histogramm (рис. 4.2). Выбрать опции Regular, Normal и число частичных интервалов (Categories) равным 10. Получим гистограмму (рис. 4.3).

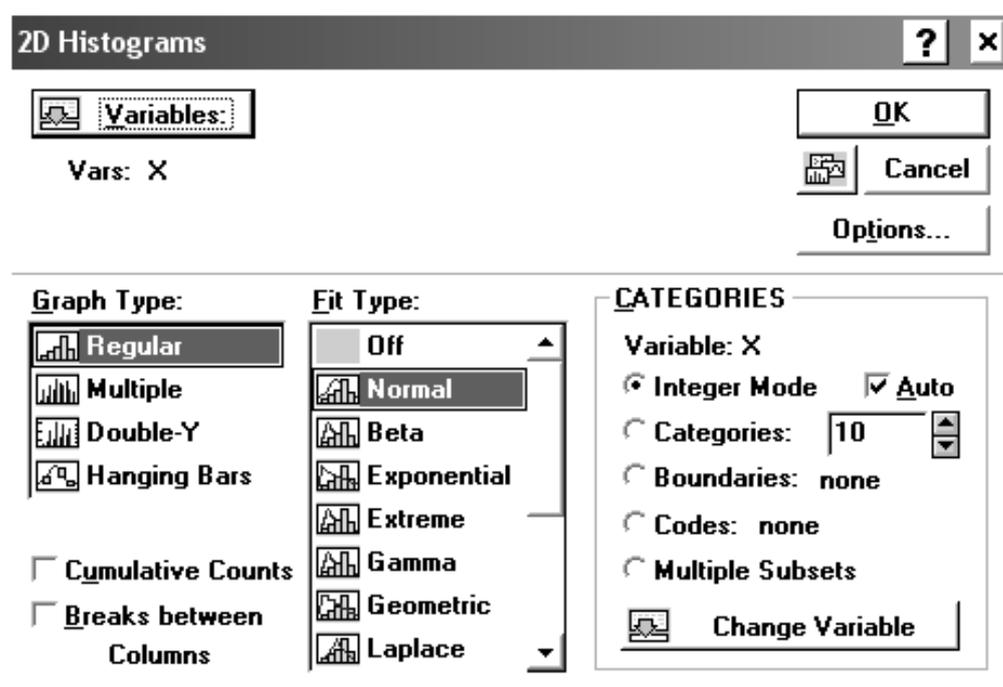


Рисунок 4.2

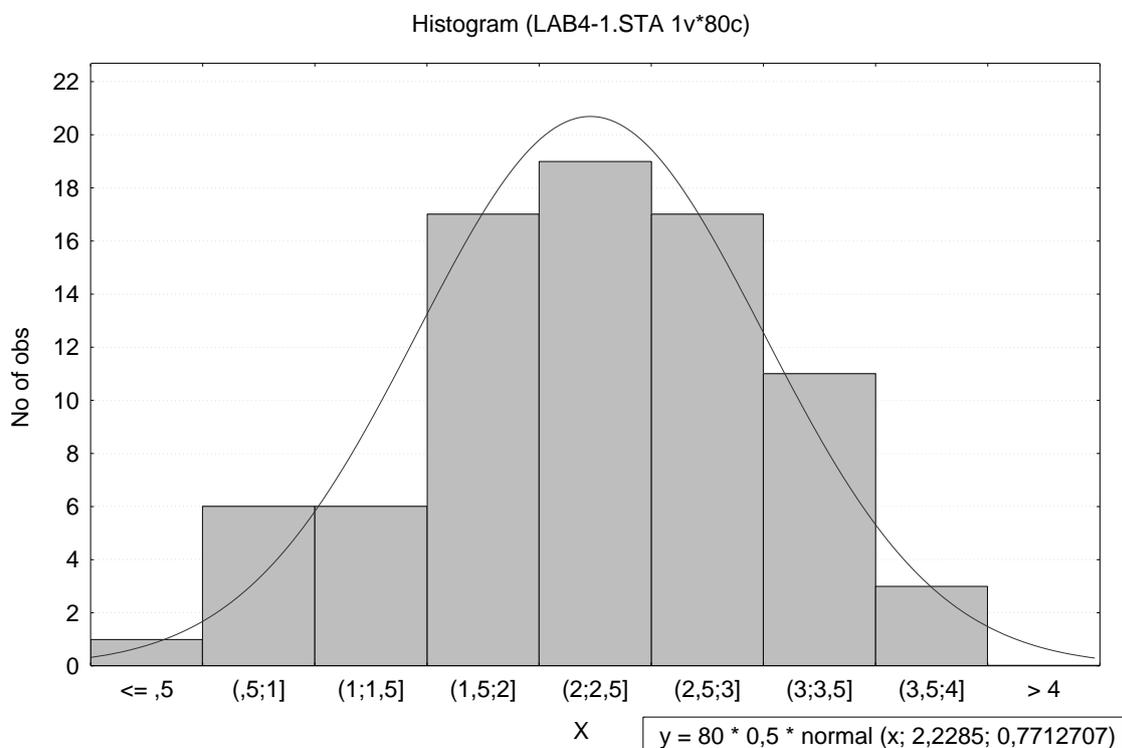


Рисунок 4.3

Аналогично строим гистограммы с кривыми для экспоненциального (Exponential) и логнормального (Lognormal) распределений (рис. 4.4, 4.5).

По виду графиков можно предварительно предположить, что наиболее подходящим является нормальное распределение.

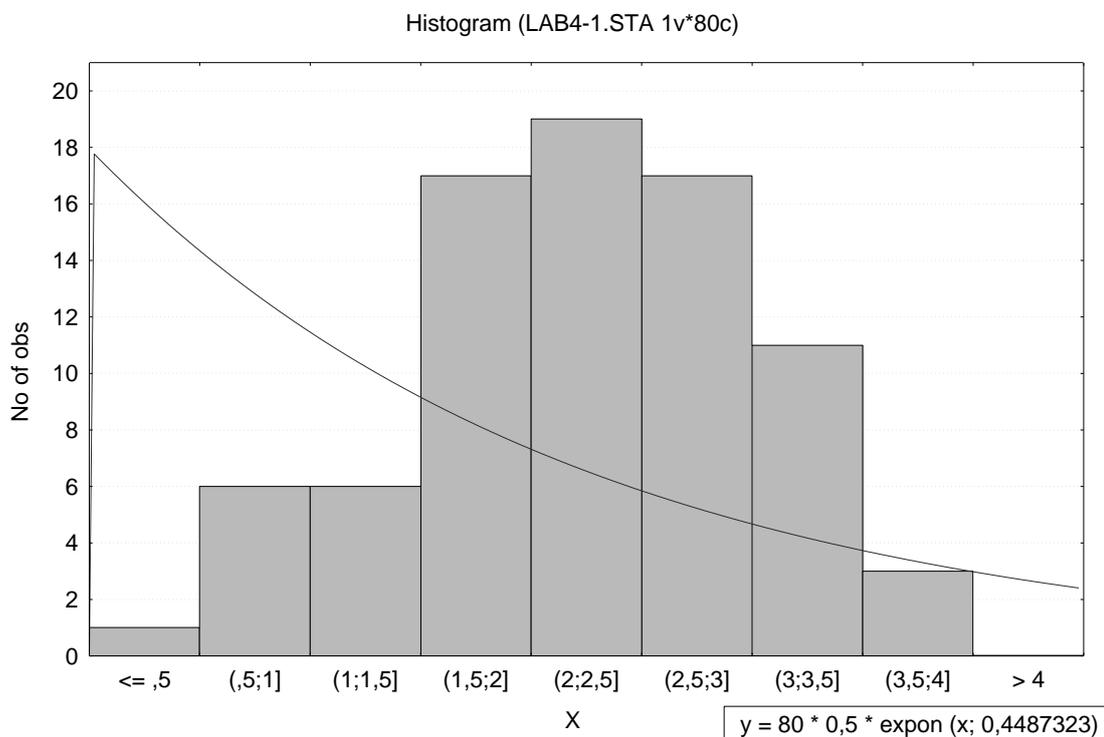


Рисунок 4.4

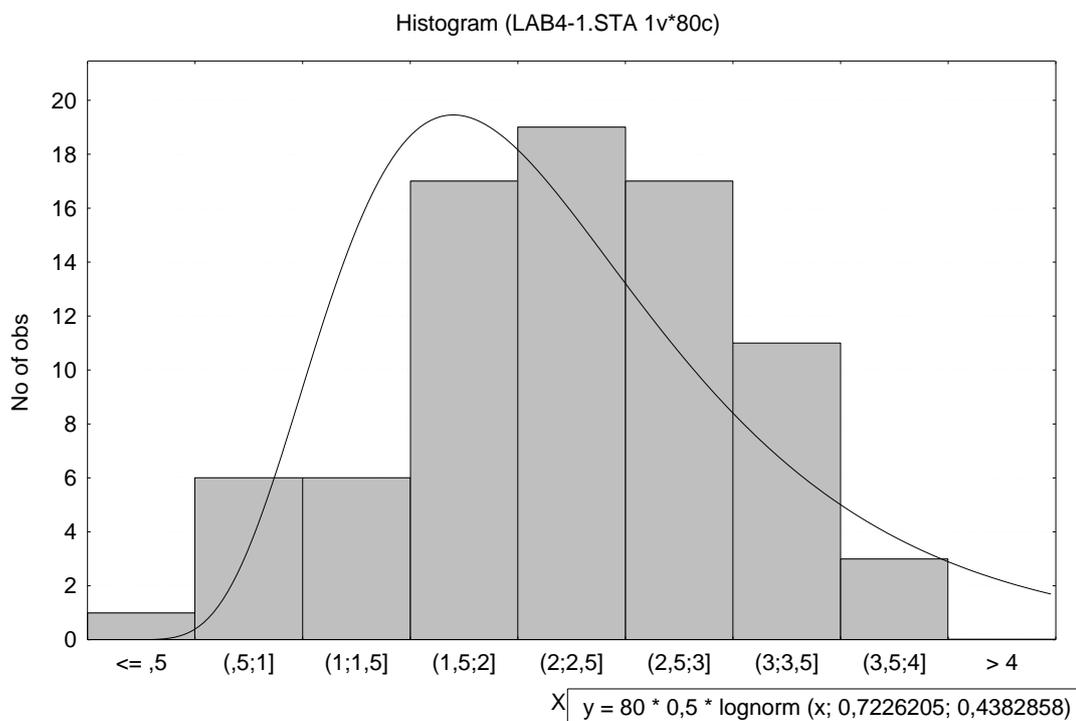


Рисунок 4.5

5 Подключаем модуль Nonparametrics/Distrib (Непараметрические статистики и подгонка распределений) (рис. 4.6).

В появившемся окне ставим переключатель на Distribution fitting (Подгонка распределения) (рис. 4.7), ОК.

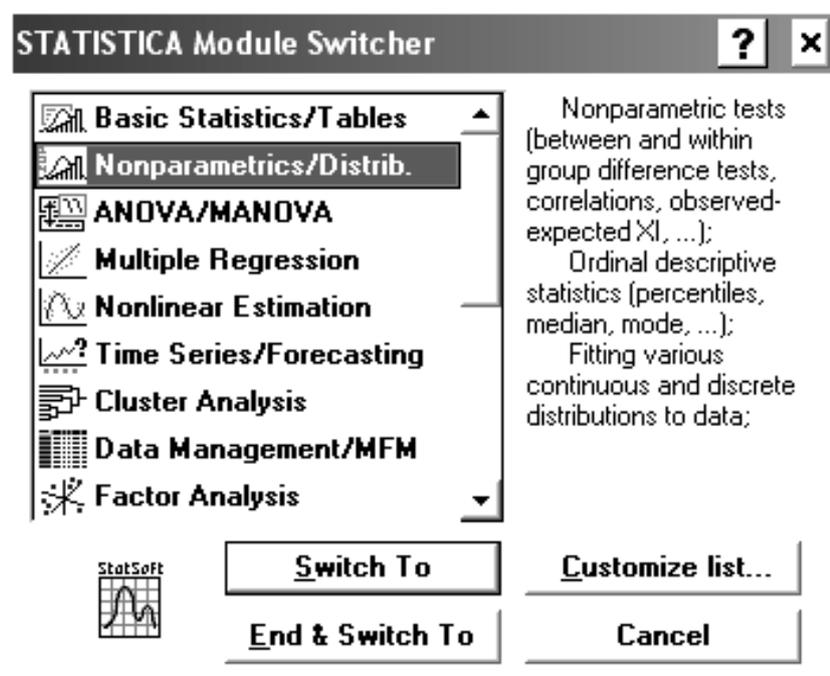


Рисунок 4.6

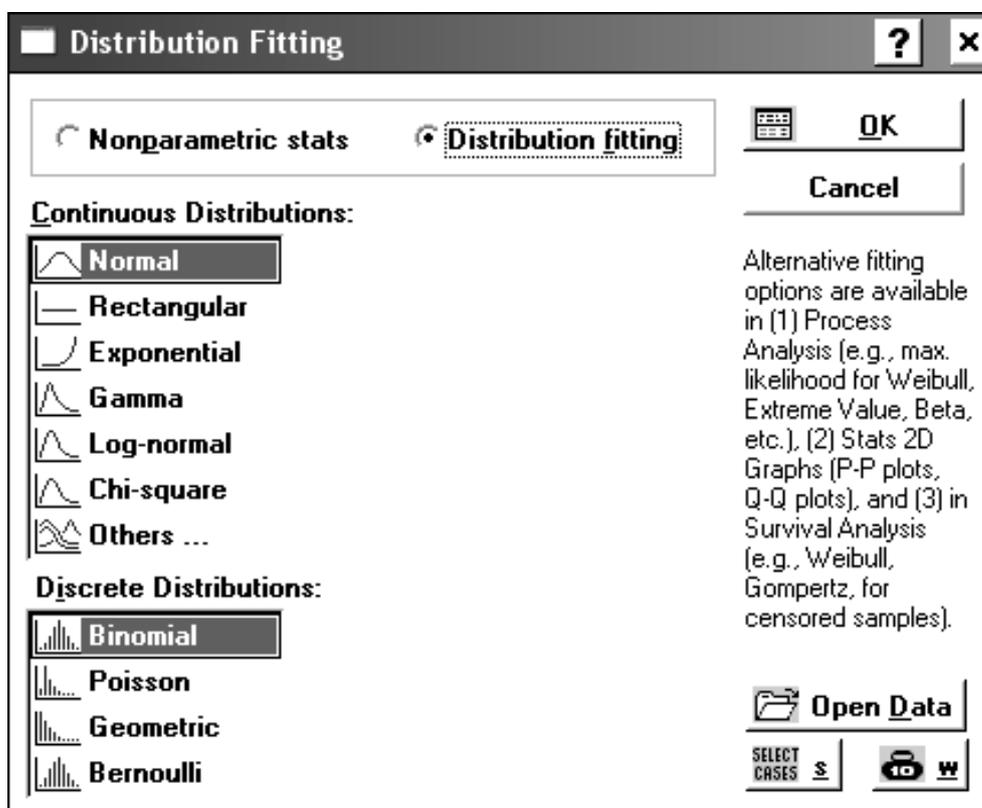


Рисунок 4.7

В окне Fitting Continuous Distributions выбираем Variables – X (переменная – X), Distribution: Normal (Распределение – нормальное), Kolmogorov-Smirnov test – Yes (continuous). ОК.(рис. 4.8).

Результат получим в виде таблицы 4.3. Значение статистики хи-квадрат небольшое – Chi-Square = 0,2986086, что свидетельствует в пользу гипотезы. Вопрос о том, что такое большое или небольшое значение статистики, снимается понятием уровня значимости. В заголовке таблицы видим, что уровень значимости  $p = 0,8613076$ , что означает, что гипотезу о согласии данных с нормальным распределением можно отвергнуть на уровне 0,8613076. Другими словами, отвергая гипотезу о нормальном распределении, мы рискуем ошибиться с вероятностью 86 %.

Окончательно заключаем, что данные хорошо согласуются с гипотезой о нормальном распределении.

Аналогичным образом получаем таблицы для проверки гипотез о согласии данных с экспоненциальным и логнормальным распределениями (табл. 4.4, 4.5).

Выводы по этим таблицам сделайте самостоятельно.

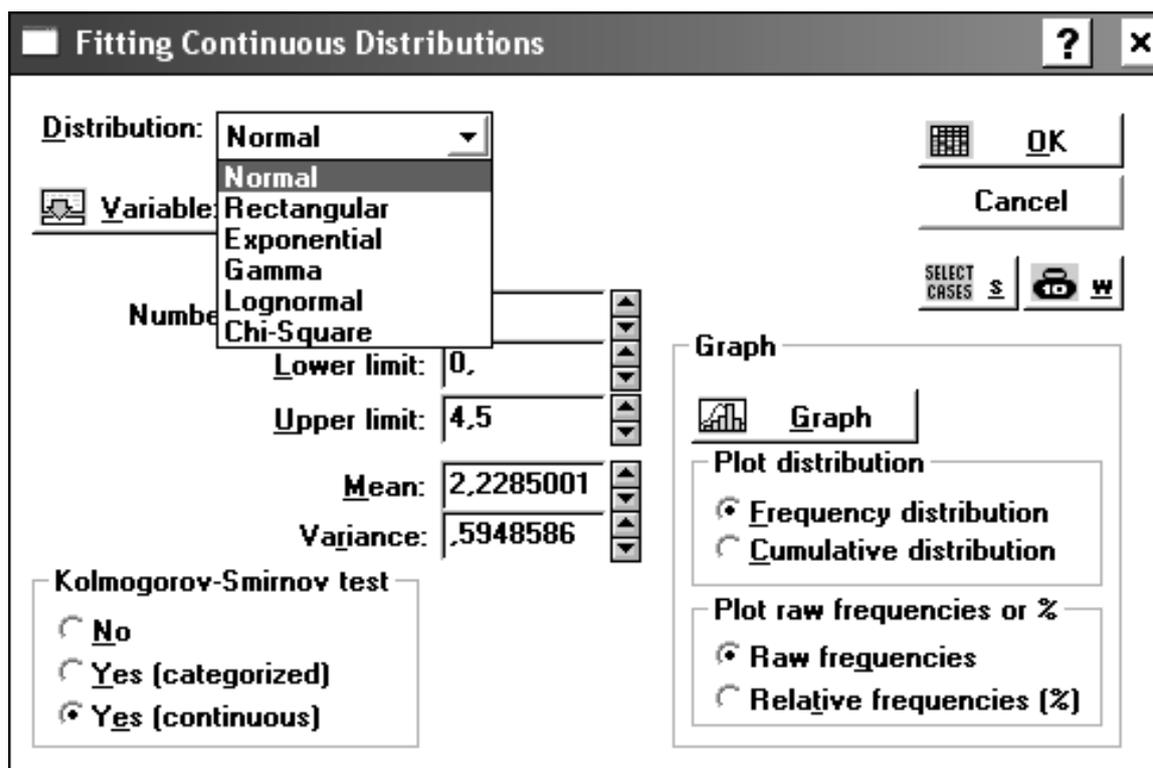


Рисунок 4.8

Таблица 4.3

Upper Boundary	observed freq-cy	cumulatv observed	percent observed	cumul. % observed	expected freq-cy	cumulatv expected
<= ,	1	1	1,25000	1,2500	1,00077	1,00077
1,0	6	7	7,50000	8,7500	3,44720	4,44797
1,5	6	13	7,50000	16,2500	9,34767	13,79564
2,0	17	30	21,25000	37,5000	16,88548	30,68112
2,5	19	49	23,75000	61,2500	20,32583	51,00696
3,0	17	66	21,25000	82,5000	16,30637	67,31333
3,5	11	77	13,75000	96,2500	8,71730	76,03063
4,0	3	80	3,75000	100,0000	3,10429	79,13493
Infinity	0	80	0,00000	100,0000	,86507	80,00000

Upper Boundary	percent expected	cumul. % expected	observed-expected
<= ,	1,25096	1,2510	-,00077
1,0	4,30900	5,5600	2,55280
1,5	11,68459	17,2446	-3,34767
2,0	21,10685	38,3514	,11452
2,5	25,40729	63,7587	-1,32583
3,0	20,38297	84,1417	,69363
3,5	10,89663	95,0383	2,28270
4,0	3,88037	98,9187	-,10429
Infinity	1,08134	100,0000	-,86507

Таблица 4.4

Upper Boundary	observed freq-cy	cumulatv observed	percent observed	cumul. % observed	expected freq-cy	cumulatv expected
<= ,	1	1	1,25000	1,2500	16,07820	16,07820
1,0	6	7	7,50000	8,7500	12,84684	28,92504
1,5	6	13	7,50000	16,2500	10,26492	39,18996
2,0	17	30	21,25000	37,5000	8,20190	47,39186
2,5	19	49	23,75000	61,2500	6,55350	53,94536
3,0	17	66	21,25000	82,5000	5,23640	59,18176
3,5	11	77	13,75000	96,2500	4,18400	63,36575
4,0	3	80	3,75000	100,0000	3,34311	66,70886
Infinity	0	80	0,00000	100,0000	13,29114	80,00000

Upper Boundary	percent expected	cumul. % expected	observed-expected
<= ,	20,09775	20,0977	-15,0782
1,0	16,05855	36,1563	-6,8468

Таблица 4.5

Upper Boundary	observed freq-cy	cumulativ observed	percent observed	cumul. % observed	expected freq-cy	cumulativ expected
<=,58889	2	2	2,50000	2,5000	,17112	,17112
1,07778	7	9	8,75000	11,2500	5,40683	5,57795
1,56667	5	14	6,25000	17,5000	15,71636	21,29430
2,05556	19	33	23,75000	41,2500	18,55463	39,84893
2,54444	18	51	22,50000	63,7500	14,96128	54,81022
3,03333	16	67	20,00000	83,7500	10,10206	64,91228
3,52222	10	77	12,50000	96,2500	6,24996	71,16224
4,01111	3	80	3,75000	100,0000	3,70312	74,86536
Infinity	0	80	0,00000	100,0000	5,13463	80,00000

Upper Boundary	percent expected	cumul. % expected	observed-expected
<=,58889	,21390	,2139	1,8289
1,07778	6,75853	6,9724	1,5932
1,56667	19,64545	26,6179	-10,7164
2,05556	23,19328	49,8112	,4454
2,54444	18,70161	68,5128	3,0387
3,03333	12,62758	81,1404	5,8979
3,52222	7,81245	88,9528	3,7500
4,01111	4,62891	93,5817	-,7031
Infinity	6,41829	100,0000	-5,1346

## 4.5 Индивидуальные задания к лабораторной работе

Индивидуальные задания к лабораторной работе взять из лабораторной работы 2.

## Приложение А

### КРАТКИЕ СВЕДЕНИЯ О СИСТЕМЕ STATISTICA

#### А.1 Структура пакета Statistica

Пакет Statistica состоит из нескольких независимо работающих модулей. После загрузки пакета на экране появляется переключатель модулей (*Module Switcher*) (рис. А.1). Необходимо выбрать нужный модуль из списка и щелкнуть на кнопке *Switch to*.

В каждом модуле собраны логически связанные между собой статистические процедуры. Загрузить можно сразу несколько модулей, но активным может быть только один. Кнопки загруженных модулей находятся на панели задач, кнопка активного модуля более светлого цвета.

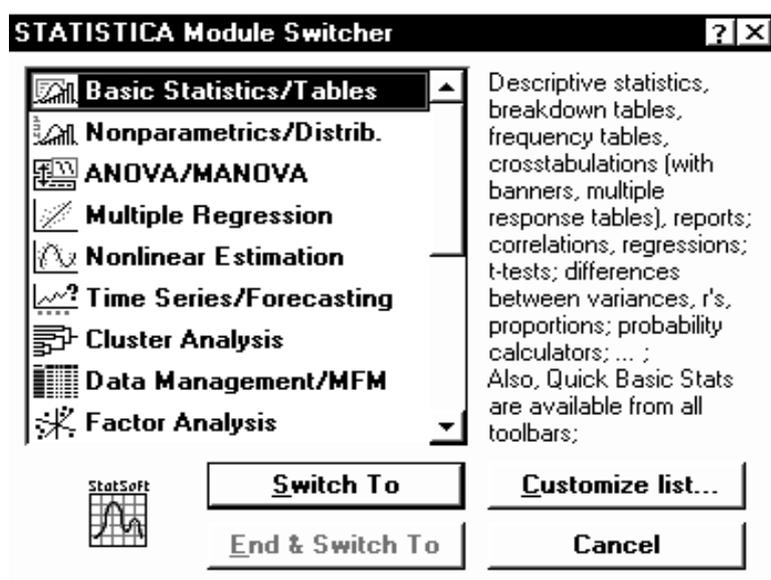


Рисунок А.1

Переход от модуля к модулю – щелчком левой клавишей мыши по соответствующей кнопке.

#### А.2 Типы файлов в Statistica

В пакете Statistica будем использовать 4 типа файлов:

- \*.sta – исходные данные;
- \*.scr – результаты статистической обработки данных;
- \*.stg – график;
- \*.rtf – отчет.

Длина имени файла, как и любого другого идентификатора, не более 8 символов.

В одном сеансе работы с пакетом может быть создано или открыто сколько угодно файлов \*.scr, \*.stg, \*.rtf, но только один файл с исходными данными \*.sta.

### А.3 Создание новой таблицы данных

Выбрать пункт меню *File – New Data*, в появившемся диалоговом окне набрать имя файла (указать полный путь к файлу) – *Ok*. Откроется пустая электронная таблица размером 10x10 (рис. А.2).

В столбцах расположены переменные (*Vars*), в строках – случаи (*Cases*).

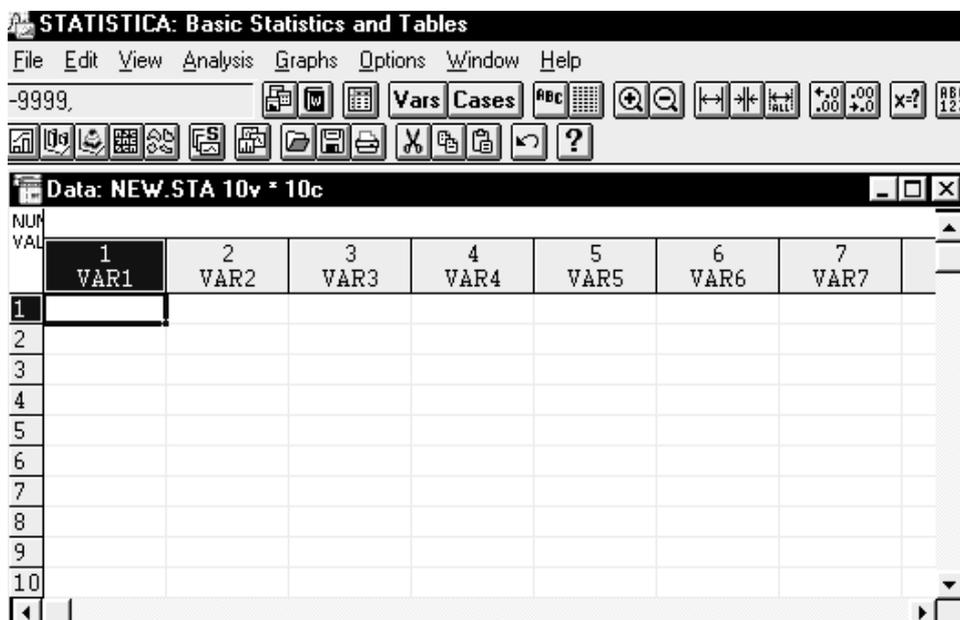


Рисунок А.2

### А.4 Удаление и добавление новых переменных и случаев

Удаление и добавление новых переменных и случаев производится командами *Add* (добавить) и *Delete* (удалить) после выделенной строки или столбца. Нажать кнопку *Vars*, если добавляются или удаляются переменные, и указать, сколько элементов добавляется или удаляется, *Ok*. Для случаев – аналогично, но с кнопкой *Cases*.

### А.5 Корректировка таблицы

Корректировка таблицы сводится к корректировке названий переменных, корректировке содержимого столбцов полностью и корректировке отдельных клеток. Для этого щелчок по имени столбца, щелчок по кнопке *Vars* (переменные), по пункту меню *Current Specs* (текущие спецификации). После этого можно задать новое имя столбцу в строке *Name* (имя) и новые значения случаям с помощью формулы в окошке *Long name* (длинное имя) (рис. А.3).

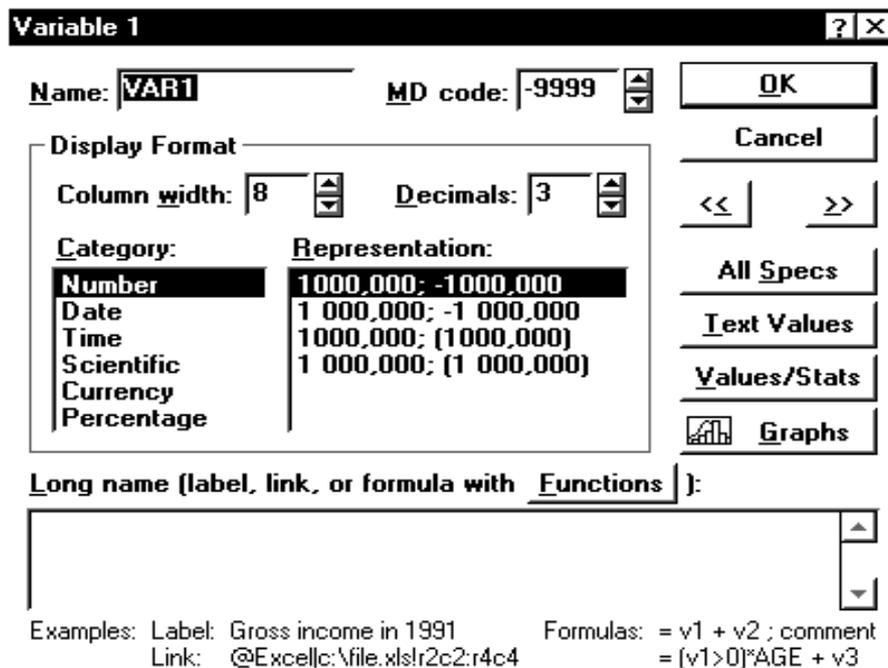


Рисунок А.3

Так же можно изменить представление числа (автоматически – 8 позиций (Column width), число позиций после запятой (Decimals) – 3). Корректировка числовых значений в отдельных клетках производится как в EXCEL.

## А.6 Создание автоотчета

Автоотчет необходимо создавать при каждом сеансе работы с пакетом, для того, чтобы все результаты работы (таблицы и графики) помещались в автоотчет.

Активировать таблицу \*.sta. Далее указан путь для создания краткого автоотчета: File, Page/Output Setup → появится двухстраничное меню со страницами Text/... и Graphs (файл – страница/создание отчета – страницы Текст и График). На странице Text/... активировать опции: Output-Off (отключение принтера и записи в файл), Window (вывод отчета на экран), Suppl.Info-Brief (краткая форма отчета) и Auto-report--Automatically Print All Scrollsheets (Auto-report) (автоматическая запись статистической таблицы в отчет) (рис. А.4).

На странице Graphs активировать опции Output-Off, Window, Metafile Mode--Screen resolution (результаты выводить на экран) и Auto-report--Automatically Print All Gaphs (автоматически вносить графики в отчет) (рис. А.5).

После этого нажать *Ok*. Появляется пустое окно автоотчета. Его можно свернуть. Все расчетные таблицы и графики помещаются в автоотчет.

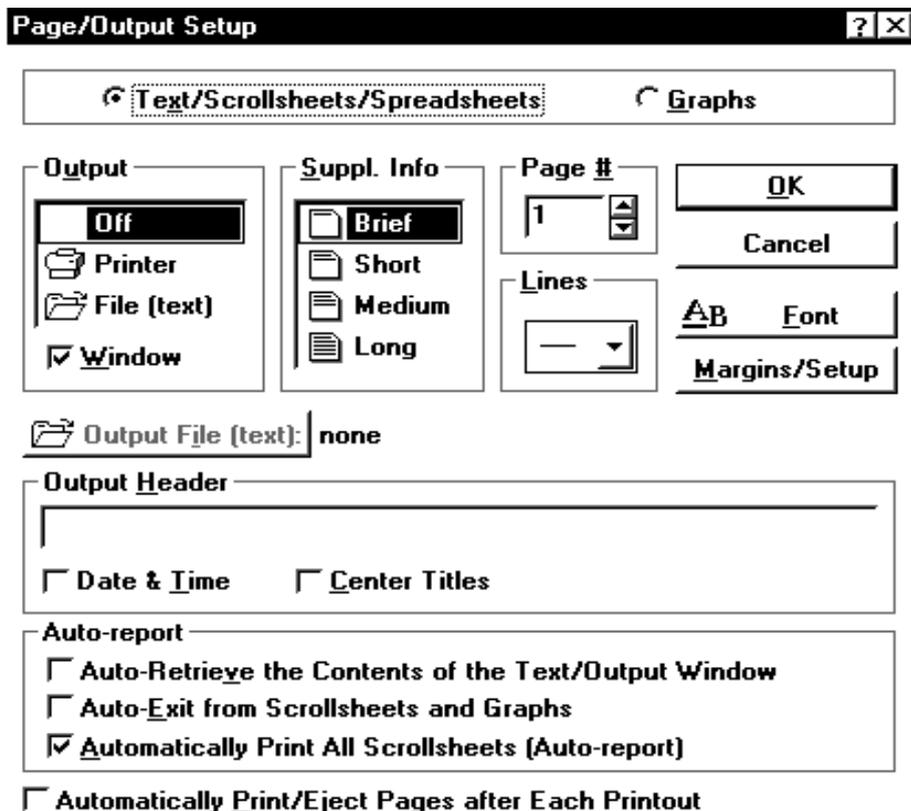


Рисунок А.4

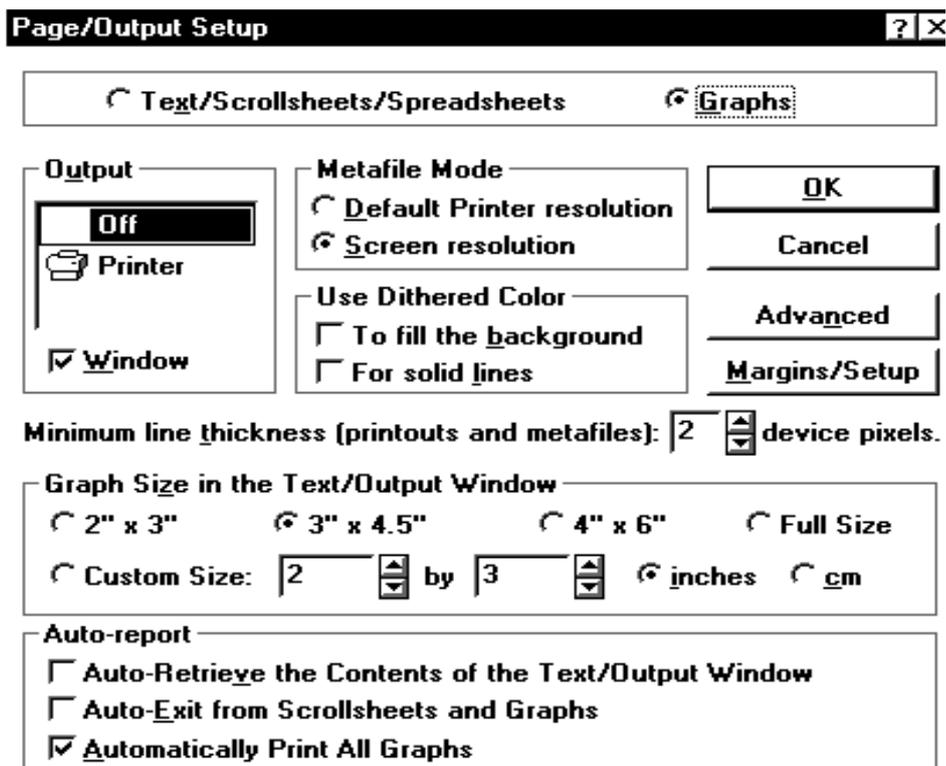


Рисунок А.5

**Примечание.** Файл отчета имеет расширение \*.rtf, его можно редактировать в WORD только после того, как файл полностью создан, т. к. после редактирования в WORD файл отчета в пакете Statistica открыть невозможно.

## A.7 Вычисление статистических характеристик для значений переменных

Вычисление статистических характеристик для значений переменных (например, максимальное, минимальное, средние значения, дисперсия и т.п.) производится следующим образом: активировать таблицу данных, затем активировать пункты меню Analysis – Descriptive Statistics – More Statistics (анализ – описательные статистики – больше статистик) (рис. А.6), выбрать из списка нужные статистические характеристики (например, min или mean), *Ok*, выделить переменные, для которых ищутся характеристики (кнопка *Variables*), *Ok*, *Ok*.

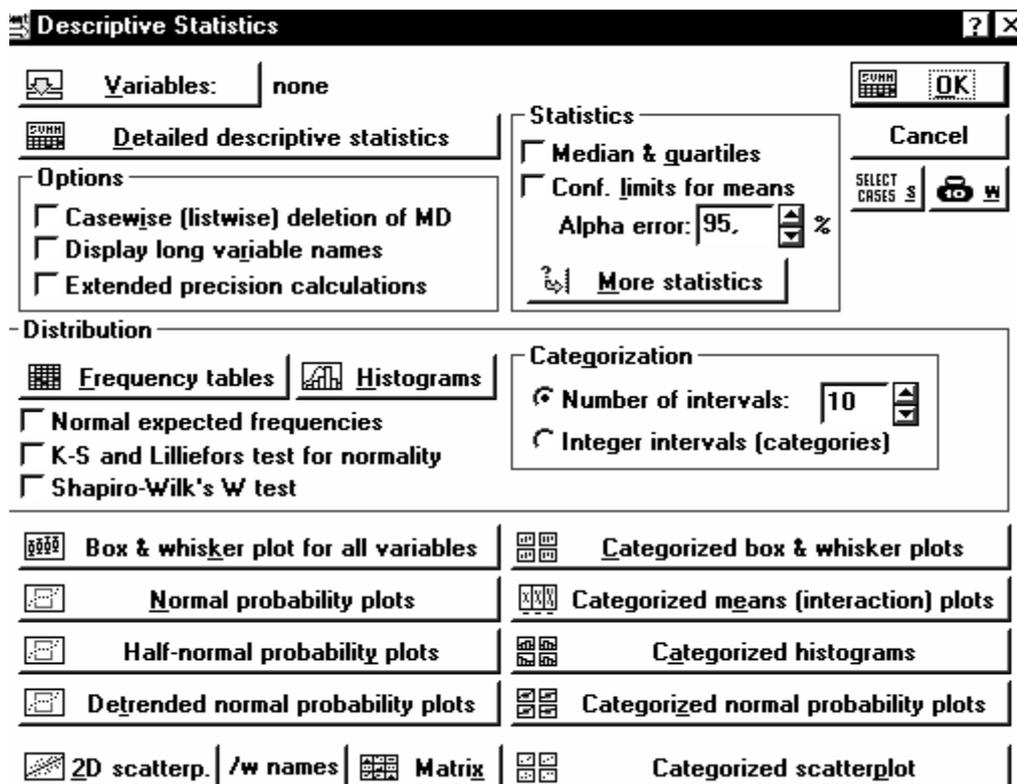


Рисунок А.6

На экран выведется таблица с нужными характеристиками.

**Примечание.** Min – минимальное значение, Max – максимальное значение, Valid N – объем выборки, Mean – среднее значение, Standard Deviation – среднее квадратическое отклонение, Variance – дисперсия.

## А.8 Получение графика и уравнения линейной регрессии

Порядок получения графика и уравнения линейной регрессии следующий. Активировать таблицу. Выбрать пункт меню Graphs, Stats 2D Graphs, Scatterplots (Графики, Статистические 2-мерные графики, Точечный график), выбрать опции Regular, Linear, *Ok* (регулярный, линейный), выбрать переменные Variables (для аргумента – X и функции – Y), *Ok* (рис. А.7). Появится график линейной регрессии, над которым записано уравнение регрессии (рис. А.8).

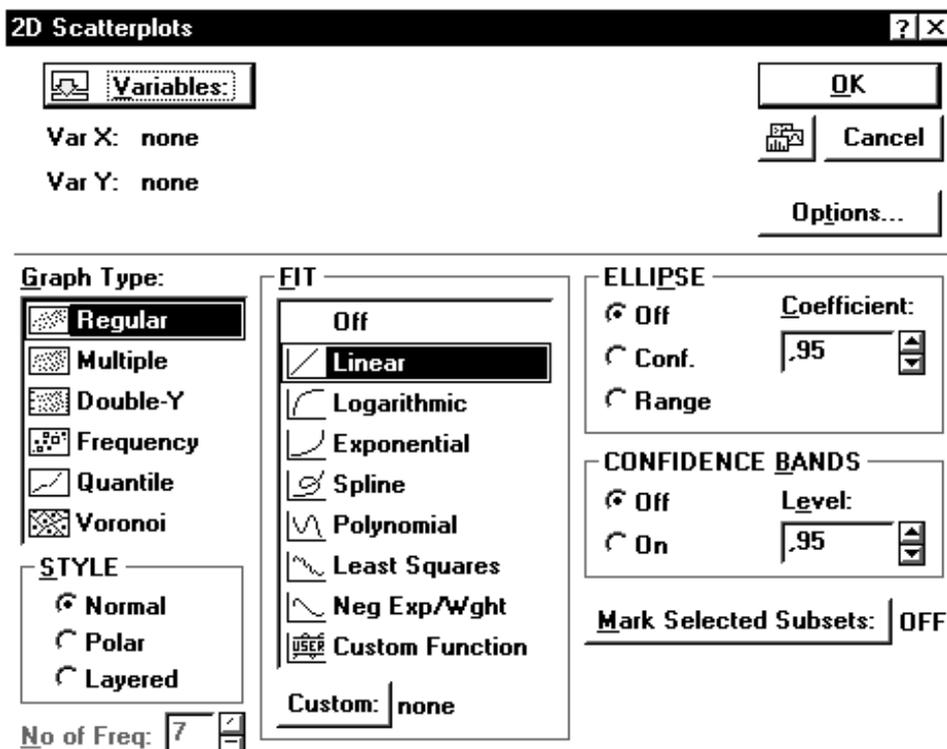


Рисунок А.7

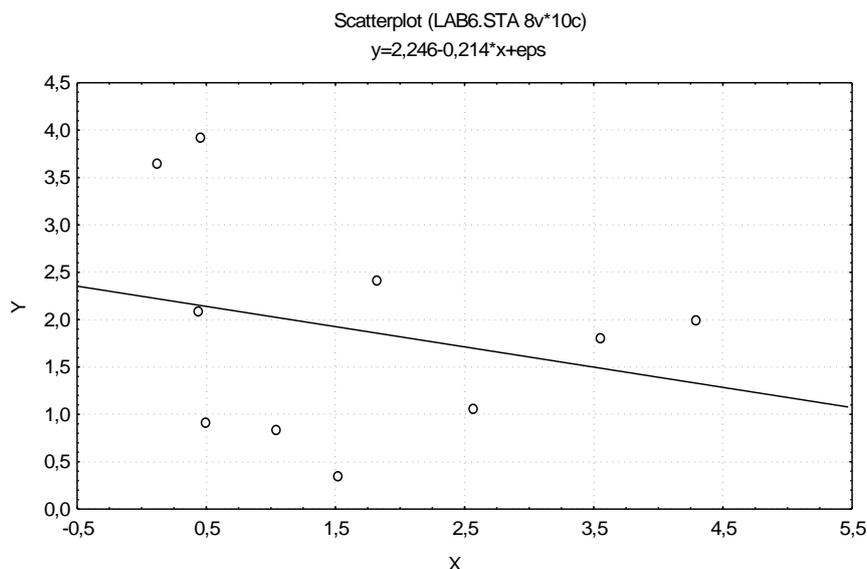


Рисунок А.8

## А.9 Особенности работы в системе Statistica 6

Основным отличием программы *Statistica 6* является то, что вычислительные модули собраны в меню *Statistics*.

При работе с различными модулями можно использовать одну таблицу данных и сохранять результаты в один файл автоотчета.

Дальше описан путь создания сжатого автоотчета в программе *Statistica 6*: *File – Output Manager*. Появится многостраничное меню. На странице *Output Manager* нужно отметить опции: *Single Workbook*, *Place results in Workbook automatically*, *Also send to Report Window*, *Single Report* (рис. А.9). На странице *Workbook*, кроме уже заданных опций, в поле *Add to Workbook performs* отметить опцию *Copy*. На странице *Report*, кроме уже заданных опций, в поле *Add to Report performs* отметить опцию *Copy*. После этого нажать *Ok*. Теперь все расчетные таблицы и графики автоматически заносятся в отчет.

Файл отчета нужно сохранить с расширением *\*.rtf*. Теперь его можно редактировать в *WORD*.

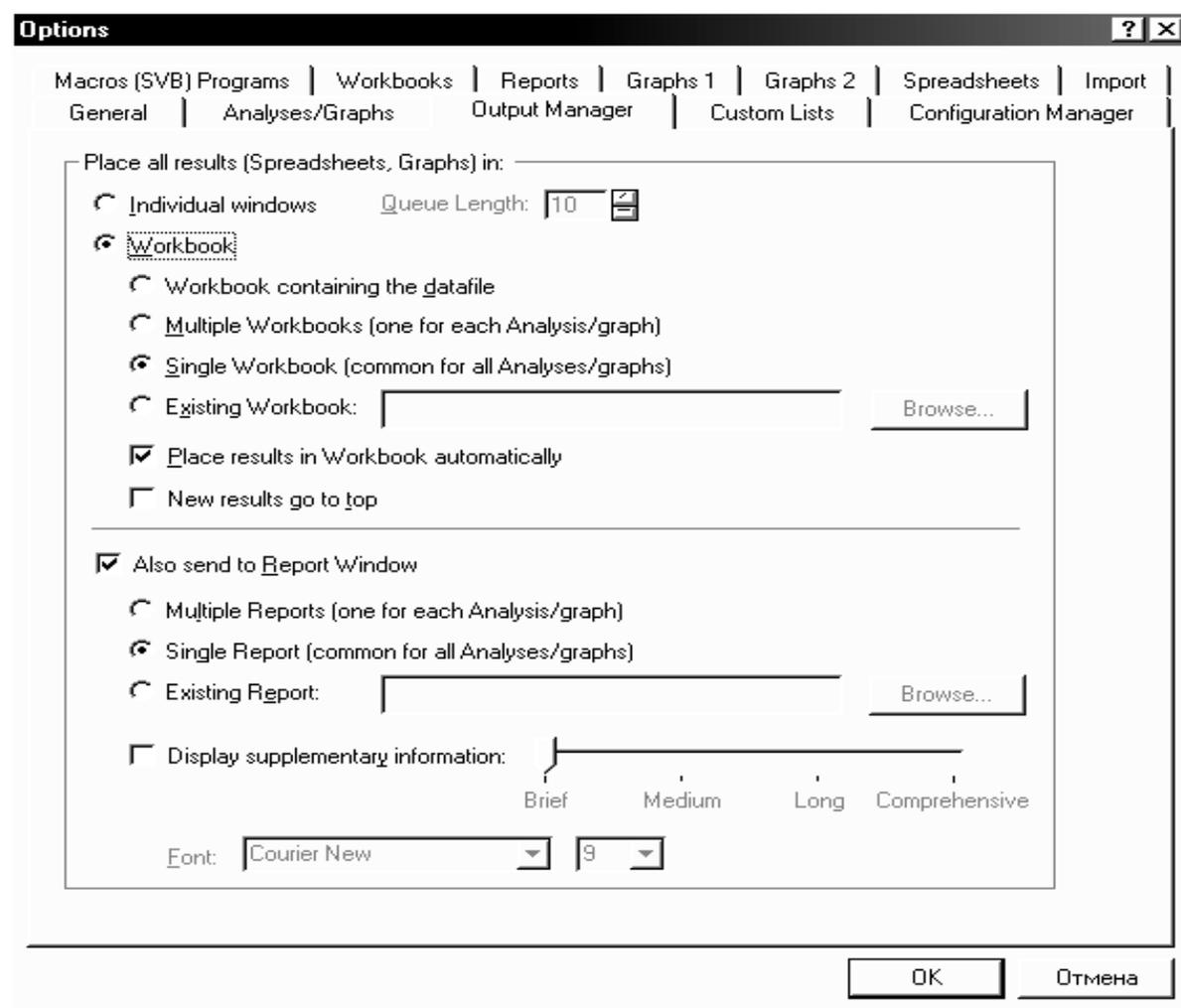


Рисунок А.9

## ПРИЛОЖЕНИЕ Б

### Пример выполнения лабораторной работы 2 в среде MathCad.

1 Средствами редактора Блокнот создаем файл данных dan.dat.

2 Рассчитываем следующие значения: объем выборки  $n$ , математическое ожидание  $\text{mean}$ , размах выборки  $R = x_{\min} - x_{\max}$ , среднеквадратическое отклонение  $\sigma$ , асимметрию  $Sk$ , эксцесс  $Ex$ .

```
ORIGIN := 1
i := 1 .. 80
ξi := READ("dan.dat")
xmax := max(ξ)      xmin := min(ξ)   xmax = 3.75   xmin = 0.09
ξ := sort(ξ)         n := length(ξ)   n = 80   R := xmax - xmin
mean := mean(ξ)      mean = 2.03
disp := var(ξ) ·  $\frac{n}{n-1}$    disp = 0.574
σ := √disp           σ = 0.758
μ3 :=  $\left(\frac{1}{n}\right) \cdot \sum_{i=1}^n (\xi_i - \text{mean})^3$    μ4 :=  $\frac{1}{n} \cdot \sum_{i=1}^n (\xi_i - \text{mean})^4$ 
Sk :=  $\frac{\mu_3}{\sigma^3}$            Sk = -0.173
Ex :=  $\left(\frac{\mu_4}{\sigma^4}\right) - 3$            Ex = -0.288
```

По значениям асимметрии и эксцесса делаем заключение, значительно ли отличается распределение случайной величины от нормального.

3 Строим гистограмму (рис. Б.1), задав число частичных интервалов равным 10.

```
m := 10 h :=  $\frac{R}{m}$            h = 0.366
j := 1 .. m           k := 1 .. m - 1
xj := xmin +  $\left(\frac{h}{2}\right) \cdot (2 \cdot j - 1)$ 
f := hist(x, ξ)
```

Высота столбцов равна количеству точек, попавших в соответствующий частичный интервал. Данная гистограмма требованиям, предъявляемым к гистограммам, не удовлетворяет (объяснить).

Уменьшаем число частичных интервалов до 7 и строим новую гистограмму (рис. Б.2).

Ответьте на вопрос: удовлетворяет ли данная гистограмма требованиям, предъявляемым к гистограммам? Можно ли еще уменьшить число частичных интервалов?

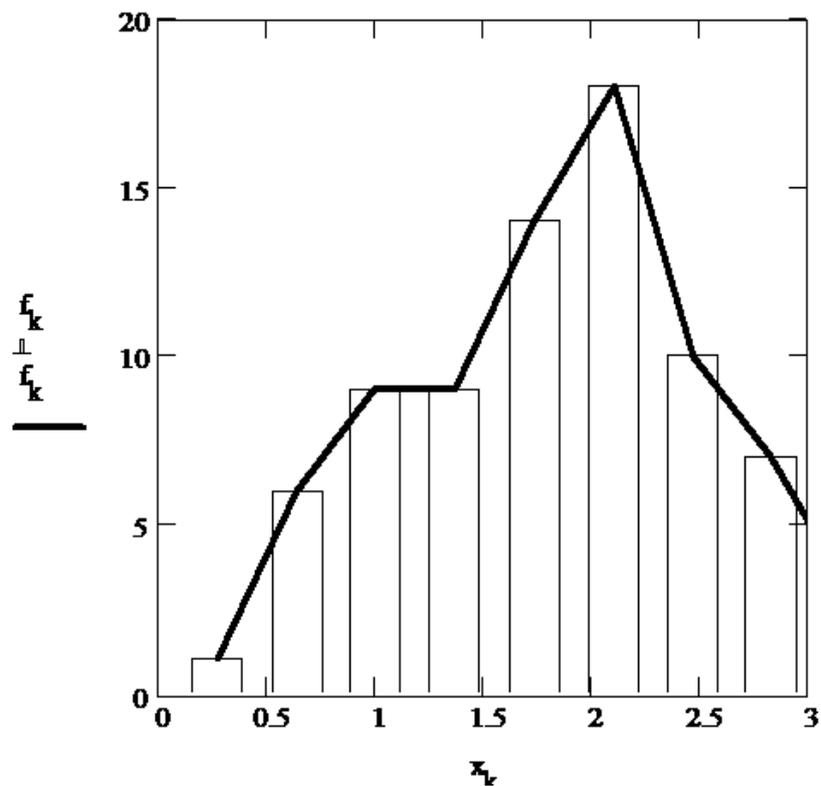


Рисунок Б.1

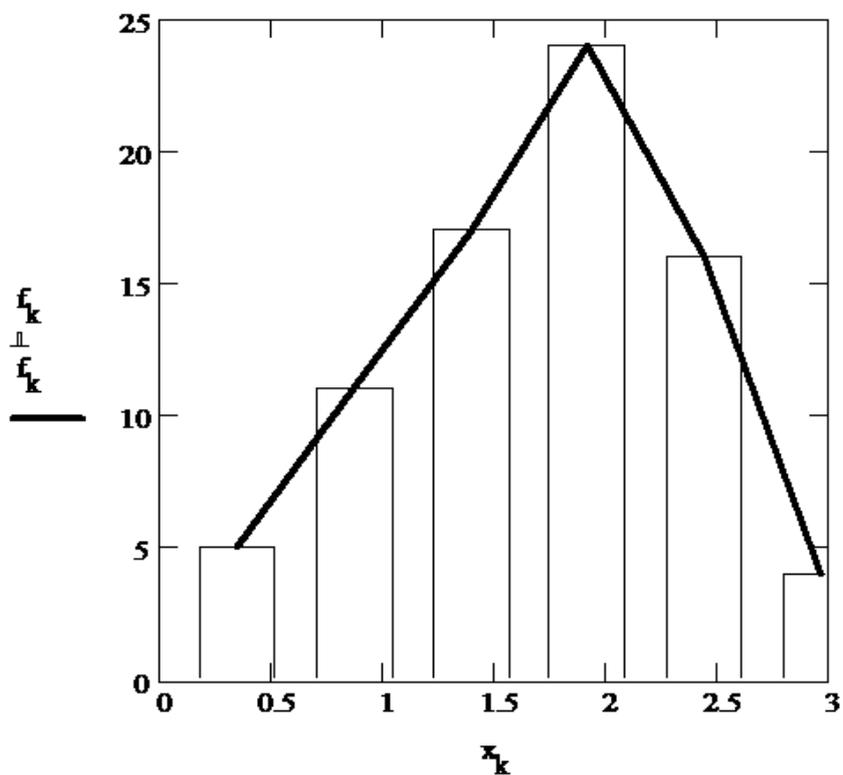


Рисунок Б.2

4 Вероятность попадания случайной величины в промежуток  $(a, b)$  рассчитывается по формуле  $P(a < X < b) = F(b) - F(a)$ , где  $F(x)$  – функция распределения случайной величины.

Значения функции распределения находим с помощью встроенной функции  $\text{pnorm}(x, \text{mean}, \sigma)$ :

$$\begin{aligned} \text{pnorm}(3.2, \text{mean}, \sigma) - \text{pnorm}(2.1, \text{mean}, \sigma) &= 0.402 \\ P(2.1 < X < 3.2) &= F(3.2) - F(2.1) = 0.402. \end{aligned}$$

5 Так как технологический процесс отрегулирован правильно, то выборочное среднее  $\bar{x}$  можно принять за значение параметра, заданного в технической документации. Десятипроцентное отклонение находим по формуле  $\delta = 0.1 \cdot \bar{x} = 0.1 \cdot 2.03$ . Затем вычисляется вероятность

$$P(|x - \bar{x}| < \delta) = 2F\left(\frac{\delta}{\sigma}\right).$$

$$\delta := 0.1 \cdot \text{mean} \qquad \delta = 0.203$$

$$2 \cdot \text{pnorm}\left(\frac{\delta}{\sigma}, \text{mean}, \sigma\right) = 0.02$$

Получаем ответ: выход годной продукции при заданном допуске  $0.1\bar{x}$  составляет 2 % от всей продукции.

## ПРИЛОЖЕНИЕ В

### Пример выполнения лабораторной работы 3 в среде MathCad.

Средствами редактора Блокнот создаем файлы данных dan\_x.dat и dan\_y.dat.

Данные будут представлены в виде рис. В.1, В.2.

```
ORIGIN := 1          N := 15          i := 1..N
```

```
xi := READ("dan_x.dat")
```

$$x^T =$$

	1	2	3	4	5	6	7	8	9	10
1	7.89	14.41	6.01	9.17	6.78	8.91	6.17	10.11	5.98	6.1

Рисунок В.1

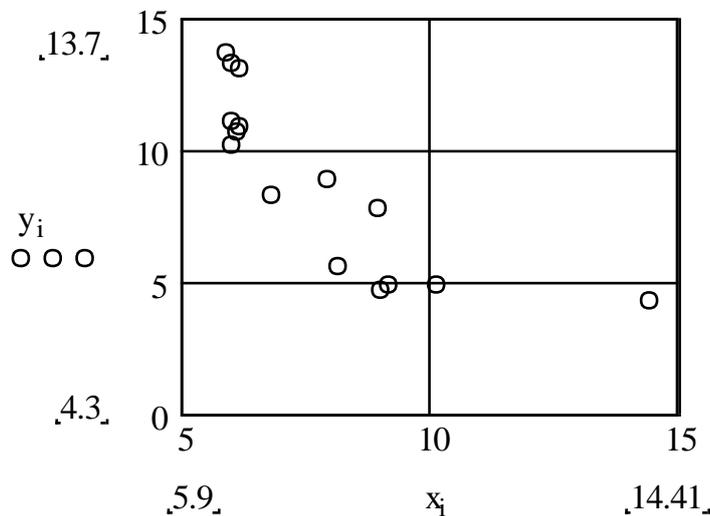
```
yi := READ("dan_y.dat")
```

$$y^T =$$

	1	2	3	4	5	6	7	8	9	10
1	8.9	4.3	10.2	4.9	8.3	7.8	13.1	4.9	13.3	10.7

Рисунок В.2

Строим корреляционное поле (рис. В.3) и находим коэффициент корреляции.



$$\text{corr}(x, y) = -0.808$$

Рисунок В.3

По виду корреляционного поля можно предположить, что зависимость между  $X$  и  $Y$  существует. Для определения тесноты линейной связи находим коэффициент корреляции:

Коэффициент корреляции для переменных  $X, Y$  равен  $-0,808$ . Так как  $0,6 < |-0,808| < 0,9$ , то линейная связь между этими переменными достаточная.

По методу наименьших квадратов находим коэффициенты модели  $y = b_0 + b_1x$ :

$$b_0 := \text{intercept}(x, y) \quad b_0 = 17.818$$

$$b_1 := \text{slope}(x, y) \quad b_1 = -1.157$$

$$y_i := b_0 + b_1 \cdot x_i$$

Находим координаты центра рассеяния и область прогнозов.  
Среднее значение фактора  $X$ :

$$X_{\text{mean}} := \text{mean}(x) \quad X_{\text{mean}} = 7.78$$

Среднее значение фактора  $Y$ :

$$Y_{\text{mean}} := \text{mean}(y) \quad Y_{\text{mean}} = 8.82$$

Средние значения (Mean) дают координаты центра рассеяния  $(\bar{x}, \bar{y}) = (7.78, 8.82)$ .

Рассчитывают значение дисперсии:

$$S^2 := \left( \frac{1}{N-1} \right) \cdot \sum_{k=1}^N (y_k - y_{\bar{k}})^2 \quad S^2 = 3.88$$

Область прогнозов задается в виде  $X_{\text{min}} \leq X \leq X_{\text{max}}$ , где минимальное и максимальное значения находим следующим образом:

$$X_{\text{min}} := \text{min}(x) \quad X_{\text{min}} = 5.9 \quad X_{\text{max}} := \text{max}(x) \quad X_{\text{max}} = 14.41$$

Область прогнозов задается интервалом  $(X_{\text{min}}; X_{\text{max}})$ , в данном примере  $(5,9; 14,41)$ .

Строим графики линейной регрессии с 80 %, 95 % и 99 % доверительными областями.

Получим три графика (рис. В.4, В.5, В.6).

Уровень доверия  $\gamma = 80\%$ .

$$\alpha := 0.20 \quad t := \text{qt} \left[ 1 - \left( \frac{\alpha}{2} \right), N - 2 \right]$$

$$\delta_i := t \cdot \sqrt{S^2} \cdot \sqrt{\left( \frac{1}{N} \right) + \frac{(x_i - X_{\text{mean}})^2}{\sum_{k=1}^N (x_k - X_{\text{mean}})^2}}$$

$$y_{\text{left}_i} := y_{r_i} - \delta_i$$

$$y_{\text{right}_i} := y_{r_i} + \delta_i$$

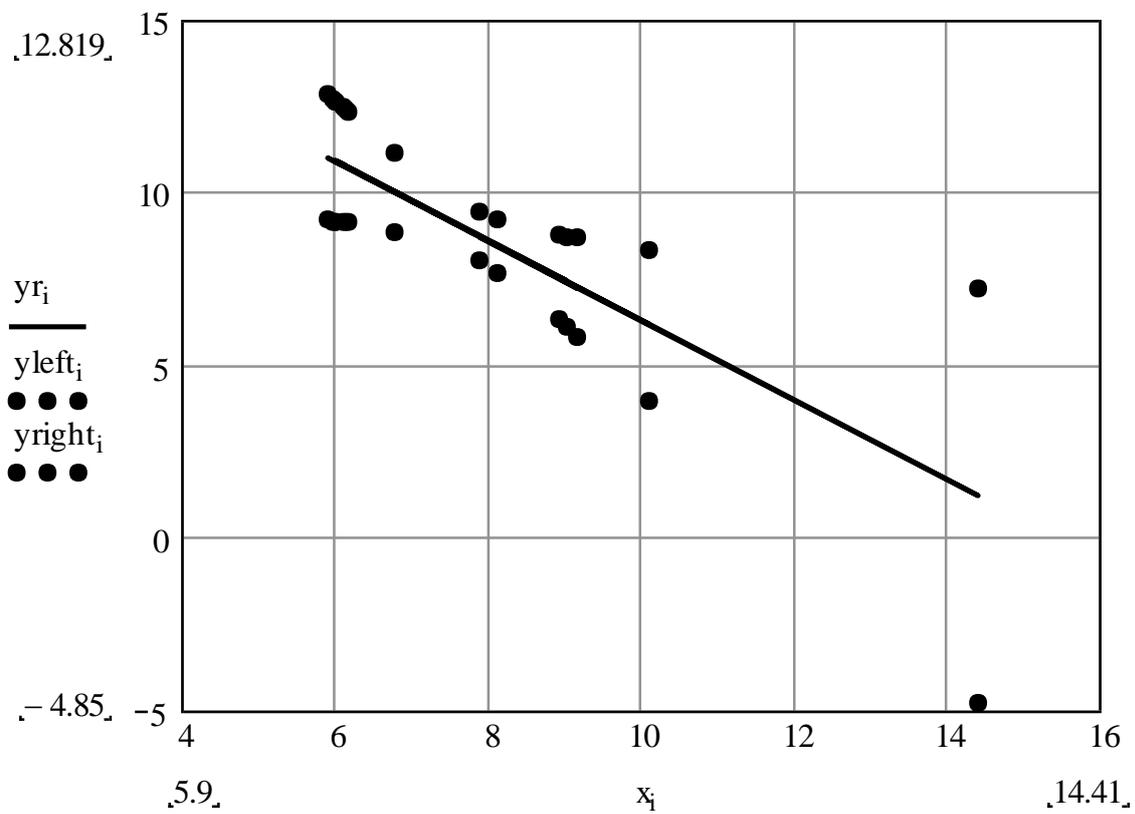


Рисунок В.4

Уровень доверия  $\gamma = 95\%$

$$\alpha := 0.05 \quad t := \text{qt} \left[ 1 - \left( \frac{\alpha}{2} \right), N - 2 \right]$$

$$\delta_i := t \cdot \sqrt{S^2} \cdot \sqrt{\left( \frac{1}{N} \right) + \frac{(x_i - X_{\text{mean}})^2}{\sum_{k=1}^N (x_k - X_{\text{mean}})^2}}$$

$$y_{\text{left}_i} := y_{r_i} - \delta_i$$

$$y_{\text{right}_i} := y_{r_i} + \delta_i$$

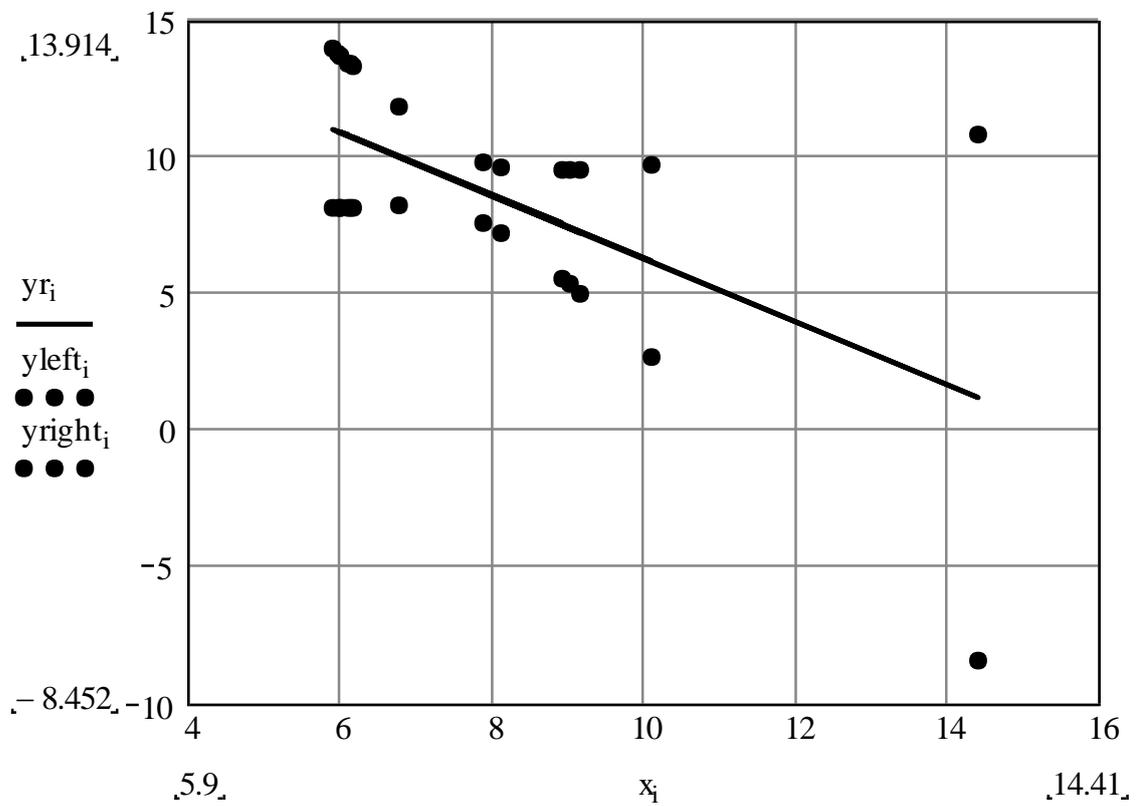


Рисунок В.5

Уровень доверия  $\gamma = 99\%$ .

$$\alpha := 0.01 \quad t := qt\left[1 - \left(\frac{\alpha}{2}\right), N - 2\right]$$

$$\delta_i := t \cdot \sqrt{S^2} \cdot \sqrt{\left(\frac{1}{N}\right) + \frac{(x_i - X_{\text{mean}})^2}{\sum_{k=1}^N (x_k - X_{\text{mean}})^2}}$$

$$y_{\text{left}_i} := y_{r_i} - \delta_i$$

$$y_{\text{right}_i} := y_{r_i} + \delta_i$$

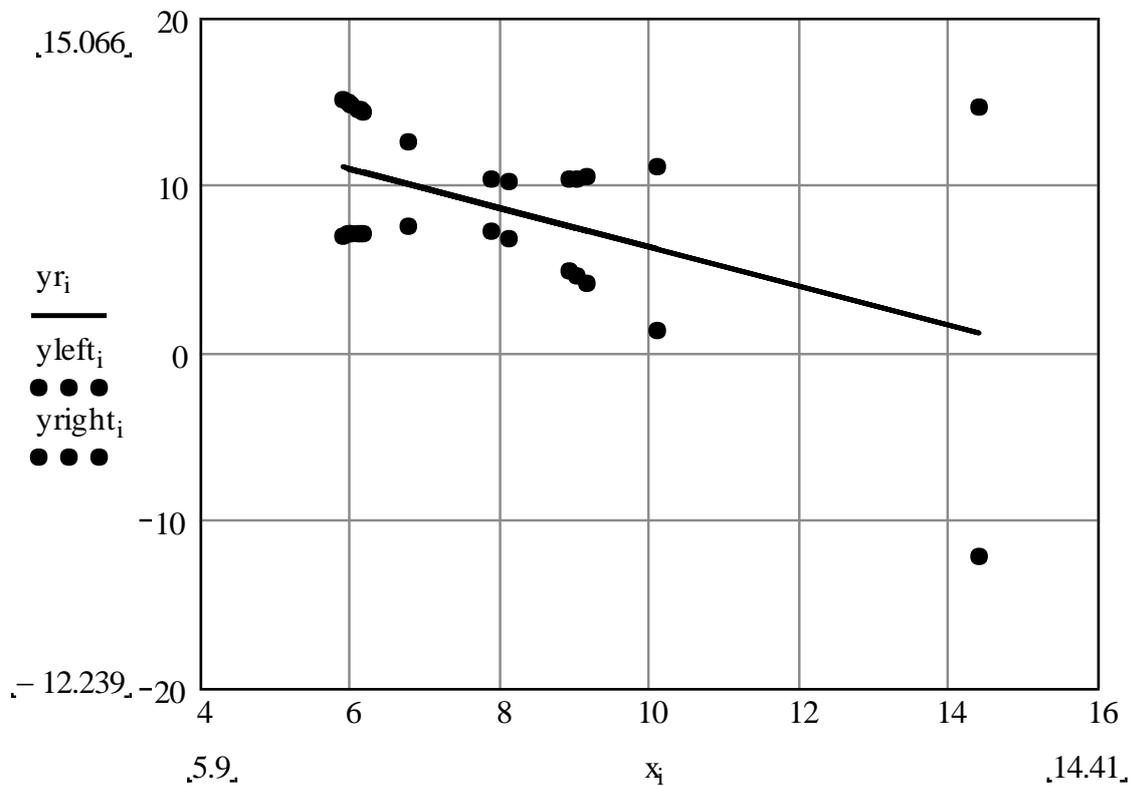


Рисунок В.6

## ПРИЛОЖЕНИЕ Г

### Тематический план изучения дисциплины, распределение времени на ее изучение, тематика контрольных работ

#### Г.1 Цель и задачи изучения дисциплины

Цель изучения дисциплины – научить студента использовать идеи вероятностно-статистического моделирования для решения инженерных задач, использовать прикладные системы (STATISTICA), а также дать студентам необходимую теоретическую базу для дальнейшего самостоятельного освоения вероятностно-статистических идей.

##### *Задачи изучения дисциплины*

После изучения дисциплины студенты должны **знать**:

- операции в алгебре событий;
- различные определения вероятности событий, правила вычисления вероятностей;
- понятие непрерывной и дискретной случайных величин, законы распределения, их графическое изображение;
- числовые характеристики дискретных и непрерывных случайных величин;
- понятие повторных независимых опытов, биномиальный закон распределения;
- нормальный закон распределения, его графическое изображение и числовые характеристики;
- понятие точечных оценок числовых характеристик случайной величины и интервалов доверия с заданным уровнем значимости;
- понятие статистической гипотезы и статистического критерия;
- понятие независимых и зависимых случайных величин, линии регрессии;
- область доверия линии регрессии;
- метод Монте-Карло, понятие скорости статистической сходимости.

После изучения дисциплины студент должен получить **навыки**:

- вычисления вероятности событий;
- работы с системой STATISTICA для вычисления числовых характеристик, построения гистограмм и нахождения линий регрессии и их доверительных областей.

После изучения дисциплины студент должен **уметь**:

- вычислять надежность схем через надежность ее составляющих;
- пользоваться формулой Бернулли и ее асимптотическими вариантами;
- вычислять числовые характеристики дискретных и непрерывных случайных величин;
- находить точечные оценки числовых характеристик случайных величин и доверительные интервалы с заданным уровнем значимости;
- находить уравнения линейной и нелинейной регрессии.

## Г.2 Распределение учебного времени по темам

Таблица Г.1

Наименование разделов, тем	Распределение по видам занятий					
	Всего	Лекции	Лабораторные работы	Практические занятия	Контроль знаний	СРС
<b>Триместры 6, 9</b>						
<b>Модуль 1. Случайные события</b>	<b>Объем модуля – 2 кредита ECTS</b>					
<b>Тема 1.1.</b> Вероятность. Алгебра событий. Вероятность суммы и произведения событий	20	2	3	2		13
<b>Тема 1.2.</b> Формулы полной вероятности и Байесса	16	2		1		13
<b>Тема 1.3.</b> Формула Бернулли и асимптотические формулы для нее	16	2		1		13
<b>Тема 1.4.</b> Простейший поток событий	17	2		1		14
Тестирование	3				3	
<b>Всего по модулю 1</b>	<b>72</b>	<b>8</b>	<b>3</b>	<b>5</b>	<b>3</b>	<b>53</b>
<b>Модуль 2. Случайные величины и математическая статистика</b>	<b>Объем модуля – 2 кредита ECTS</b>					
<b>Тема 2.1.</b> Случайные величины, их классификация и законы распределения	6	2				4
<b>Тема 2.2.</b> Числовые характеристики случайных величин, их свойства	7	2		1		4
<b>Тема 2.3.</b> Равномерный и нормальный распределения	7	2		1		4
<b>Тема 2.4.</b> Элементы математической статистики: основные определения, обработка одномерного статистического массива	9	2	3			4
<b>Тема 2.5.</b> Распределения биномиальное и Пуассона. Показательное распределения. Распределение Вейбулла	6	2				4
<b>Тема 2.6.</b> Точечные и интервальные оценки параметров распределения	7	2		1		4
<b>Тема 2.7.</b> Двумерные случайные величины, законы их распределения	6	2				4
<b>Тема 2.8.</b> Числовые характеристики, коэффициент корреляции	7	2		1		4
<b>Тема 2.9.</b> Линейная и нелинейная регрессии	9	2	3			4
<b>Тема 2.10.</b> Статистические гипотезы	4	1				3
Тестирование	4				4	
<b>Всего по модулю 2</b>	<b>72</b>	<b>19</b>	<b>6</b>	<b>4</b>	<b>4</b>	<b>39</b>
<b>Всего за 6-й, 9-й триместры</b>	<b>144</b>	<b>27</b>	<b>9</b>	<b>9</b>	<b>7</b>	<b>92</b>

### Г.3 Тестирование и оценка знаний по дисциплине

Тестирование, контрольные работы и защита задач предназначены для текущего контроля знаний студентов согласно действующему в Академии положению об организации учебного процесса. Их цель – выявить уровень усвоения знаний и умений, предусмотренных программой дисциплины.

В условиях адаптации идей ECTS к системе высшего образования Украины контрольные работы могут заменяться тестированием. Для повышения уровня тестирования и автоматизации этой процедуры на кафедре разработано программное обеспечение системы компьютерного тестирования знаний студентов **Testing** [10]. Эта система позволяет достаточно быстро проверить уровень усвоения знаний студентов.

**Защита модуля** требует выполнения студентом комплекса заданий, запланированных кафедрой и необходимых для получения положительной оценки по модулю. Для оценки уровня подготовки студентов используется 100-балльная рейтинговая накопительная система, разрабатывается плане-график выполнения работ “Состав модулей дисциплины” (табл. Г.2), в котором приводится перечень составных частей (лабораторных работ, практических занятий, контрольных точек) каждого модуля с указанием сроков защиты и количества баллов, которые можно набрать при выполнении каждой составляющей.

Защита составляющих модуля выполняется в сроки, указанные в плане-графике. Модуль считается выполненным, если выполнены все его составляющие.

Итоговая оценка за триместр по дисциплине определяется с учетом весовых коэффициентов каждого из модулей.

**Г.4 Состав модулей дисциплины “Прикладная математика”,  
распределение времени на освоение, сроки контроля**

*Таблица Г.2*

Краткое содержание модуля	Триместр	Общее количество часов	Кредиты ECTS	Количество ауд. часов	Формы и методы контроля	Неделя проведения
1 Случайные события	6, 9	72	2	16	Выполнение и защита лабораторной работы	1–7
					Выполнение и защита практических работ	
					Тестирование	
2 Случайные величины и математическая статистика	6, 9	72	2	23	Выполнение и защита лабораторной работы	8–15
					Выполнение и защита практических работ	
					Тестирование	
		144	4	45		

## СПИСОК РЕКОМЕНДУЕМОЙ ЛИТЕРАТУРЫ

1 **Гмурман, В. Е.** Теория вероятностей и математическая статистика : учебное пособие / Гмурман В. Е. – 9-е изд., стер. – М. : Высшая школа, 2003. – 480 с.

2 **Гмурман, В. Е.** Руководство к решению задач по теории вероятностей и математической статистике : учебное пособие для студентов вузов / Гмурман В. Е – М. : Высшая школа, 2002. – 404 с.

3 **Смирнов, Н. В.** Курс теории вероятностей и математической статистики для технических приложений / Н. В.Смирнов, И. С. Дунин-Барковский. – М. : Наука, 1969. – 512 с.

4 **Боровиков, В.** STATISTICA : искусство анализа данных на компьютере для профессионалов / Боровиков В. – СПб. : Питер, 2001. – 656 с. : ил. – ISBN 5-318-00302-8.

5 Математическая статистика : учебник / под ред. В. С. Зарубина и А. П. Крищенко. – М. : МГТУ им. Баумана, 2001. – 424 с.

6 **Герасимович, А. И.** Математическая статистика : учеб. пособие для инж.-техн. спец. вузов / Герасимович А. И. – 2-е изд., перераб. и доп. – Минск : Выш. школа, 1983. – 280 с.

7 **Горелова, Г. В.** Теория вероятностей и математическая статистика в примерах и задачах с применением EXCEL : учебное пособие / Горелова Г. В., Кацко И. А. – Изд. 2-е, испр. и доп. – Ростов н/Дону : Феникс, 2002. – 400 с.

8 **Колемаев, В. А.** Теория вероятностей и математическая статистика : учебник / Колемаев В. А., Калинина В. Н. – М. : ИНФРА-М, 1999. – 304 с.

9 **Кремер, Н. Ш.** Теория вероятностей и математическая статистика : учебник / Кремер Н. Ш. – М. : ЮНИТИ, 2001. – 544 с.

10 **Васильева, Л. В.** Курс лекцій та контрольні завдання з дисципліни “Прикладна математика” : навчальний посібник для студентів вищих технічних навчальних закладів / Васильева Л. В., Гетьман І. А., Топтунова Л. М., Чорномаз В. М. – Краматорськ : ДДМА, 2006. – 132 с. – ISBN 966-379-074-1.

*Навчальний посібник*

**ВАСИЛЬЄВА Людмила Володимирівна,  
ЧОРНОМАЗ Володимир Миколайович,  
МАЛИГІНА Світлана Валеріївна,  
КЛЬОВАНІК Олена Анатоліївна**

## **ПРИКЛАДНА МАТЕМАТИКА**

### **ЛАБОРАТОРНИЙ ПРАКТИКУМ**

**для студентів вищих навчальних закладів**

*(Російською мовою)*

Редактор                      О. М. Болкова

Комп'ютерна верстка                      О. П. Ордіна

34/2008. Підп. до друку 16.11.09. Формат 60x84/16.  
Папір офсетний. Ум. друк. арк. 3,72. Обл.-вид. арк. 3,46.  
Тираж 70 прим. Зам. №114

Донбаська державна машинобудівна академія  
84313, м. Краматорськ, вул. Шкадінова, 72.  
Свідоцтво про внесення суб'єкта видавничої справи  
до Державного реєстру  
серія ДК № 1633 від 24.12.2003