Санкт-Петербургский государственный университет
аэрокосмического приборостроения

ЭКОНОМЕТРИКА

Методические указания
и контрольные задания для студентов-заочников
экономических специальностей

Санкт-Петербург - 2002

Утверждено на заседании кафедры компьютерной математики и программирования в качестве методических указаний для студентов вечернего и заочного отделений факультета экономики и менеджмента.

Составители – доктор физико-математических наук,
профессор Нарбут М.А.,
Соколовская М.В.

ПРЕДИСЛОВИЕ

Курс эконометрики появился в учебных планах по экономическим специальностям совсем недавно. Как видно из названия курса ("эконо-" - экономика, "-метрика" - измерение), он посвящен проблемам измерения экономических величин и процессов. Впрочем, некоторые авторы предпочитают название курса "эконометрия". В системе западного экономического образования курс эконометрики рассматривается как важнейшая составляющая курса экономической теории - наряду с микроэкономикой и макроэкономикой. В России до недавнего времени вопросы, относящиеся к эконометрике, изучались в курсах статистики (экономической статистики), а также в курсе математической статистики. Владение методами математической статистики является совершенно необходимым при изучении эконометрики, и на установочных сессиях мы будем далее часто их напоминать. Для повторения основных понятий теории вероятностей и математической статистики можно обратиться к учебнику В.Е.Гмурмана[2]. Что же касается собственно курса эконометрики, то для более полного его изучения можно в первую очередь рекомендовать учебные пособия [3,4,7].

Вопросы к зачету

1. Дискретные случайные величины. Математическое ожидание и дисперсия дискретной случайной величины. Свойства математического ожидания и дисперсии.

2. Непрерывные случайные величины. Математическое ожидание и дисперсия непрерывной случайной величины. Свойства математического ожидания и дисперсии.

3. Гауссово (нормальное) распределение, его плотность и функция распределения. Правило «трех сигм».

4. Ковариация двух случайных величин и коэффициент корреляции, их свойства.

5. Выборочное среднее и выборочная дисперсия.

6. Свойства статистических оценок - несмещенность, эффективность, состоятельность.

7. Доверительные интервалы для оценки математического ожидания нормального распределения при известной дисперсии.

8. Доверительные интервалы для оценки математического ожидания нормального распределения при неизвестной дисперсии.

9. Метод наименьших квадратов в случае линейной зависимости двух величин (модель парной регрессии).

10. Метод наименьших квадратов в случае линейной функции многих переменных (модель множественной регрессии).

11. Ковариационная матрица оценок коэффициентов в случае парной регрессии.

12. Логарифмические преобразования переменных. Коэффициент эластичности.

13. Проверка статистических гипотез. Ошибки первого и второго рода. Уровень значимости и мощность критерия. t-тест.

14. Коэффициент детерминации. F-тест на качество регрессии.

15. Метод максимального правдоподобия. Построение регрессионных моделей при гетероскедастичности ошибок.

16. Временные ряды. Выявление тренда и сезонных составляющих.

17. Корреляционная функция стационарного временного ряда, ее свойства.

18. Линейные стационарные модели. Процессы авторегрессии – скользящего среднего.

19. Спектральный анализ временных рядов.

20. Статистическое моделирование в эконометрических исследованиях. Генераторы случайных чисел, равномерно распределенных в заданном интервале. Моделирование дискретных и непрерывных случайных величин.

Примечание 1. Вопросы 1-8 на установочной сессии не рассматриваются – их следует повторить по учебнику [2] или какому-либо другому пособию по теории вероятностей и математической статистике.

Примечание 2. На практических занятиях студенты должны овладеть навыками работы с электронными таблицами EXCEL и программой STATISTICA.

В EXCEL необходимо уметь находить значения функций распределения вероятностей – гауссового (НОРМРАСП) и Стьюдента (СТЬЮДРАСП), а также определять параметры линейной и экспоненциальной функции одной или многих переменных по методу наименьших квадратов (ЛИНЕЙН, ЛГРФПРИБЛ).

В программе STATISTICA осваиваются модули Basic Statistics & Tables, Multiple Regression, Time Series/Forecast. Если студент не имеет возможности работать на персональном компьютере дома или на работе, такая возможность ему будет предоставлена в Университете.

Контрольное задание №1.

На зачет каждый студент должен представить решение задач 1-5 из варианта, номер которого совпадает с последней цифрой номера его студенческого билета или зачетной книжки. Здесь мы для примера приводим решения аналогичных задач. Необходимый теоретический материал можно найти в учебнике [2] и в других руководствах по теории вероятностей.

Методические указания

Обработка статистических данных совершается на основании положений теории вероятностей.

Теория вероятностей, вводя понятие вероятности случайного события, дает способ измерять числом степень возможности его осуществления и указывает приемы для определения этого числа. При этом теория вероятностей не может предсказать исход единичного события. Значение выявленных с помощью теории вероятностей закономерностей массовых явлений состоит в том, что они позволяют предвидеть, как эти события будут протекать в дальнейшем.

Случайной величиной называется *числовая характеристика, связанная с изучаемым объектом, значение которой принципиально не может быть предсказано точно в зависимости от случая*.

Формально случайная величина X – это числовая функция, заданная на некотором вероятностном пространстве (Ω,P): $X(\omega)$, $\omega\in\Omega$. Функцией распределения случайной величины X называется числовая функция числового аргумента, определяемая равенством $F(x)=P(X\le x)$, $x\in R$ (R – множество действительных чисел). Каждая функция распределения обладает следующими свойствами:

- ✓ $0\le F(x)\le 1$ при любом $x\in R$;
- ✓ $F(x)$ является неубывающей, непрерывной справа функцией;
- ✓ $\lim_{x\to-\infty} F(x)=0, \quad \lim_{x\to+\infty} F(x)=1$.

Функция распределения содержит всю вероятностную информацию о случайной величине *X*. В частности, $P(X\in(a,b])=F(b)-F(a)$ для любых чисел $a\le b$. Дискретную случайную величину удобно представлять в виде таблицы

$$X=\begin{pmatrix} x_1 & x_2 & \dots & x_k & \dots \\ p_1 & p_2 & \dots & p_k & \dots \end{pmatrix}, \qquad p_k=P(X=x_k) \qquad (1).$$

Случайная величина X называется ***непрерывной***, *если её функция распределения дифференцируема*, т.е. существует *производная* $p(x)=F'(x)$, называемая ***плотностью распределения случайной величины X***, или сокращенно ***плотностью вероятности***. В частности, $F(x)=\int_{-\infty}^{x} p(x)dy$. Плотность распределения обладает следующими свойствами:

- ✓ $p(x) \geq 0$ при любом $x \in R$;
- ✓ $\int_{-\infty}^{+\infty} p(y)dy = 1$.

<u>Математическое ожидание (среднее значение)</u> дискретной случайной величины X, имеющей распределение (1), есть по определению сумма ряда $E(x) = \sum_k x_k p_k$ *при условии его абсолютной сходимости*. Для непрерывной случайной величины X с плотностью распределения $p(x)$ математическое ожидание – это интеграл $E(x) = \int_{-\infty}^{+\infty} xp(x)dx$ также при условии, что он абсолютно сходится. Математическое ожидание имеет следующие свойства (X, Y– произвольные случайные величины, a, b – константы):

- ✓ $E(aX+bY)=aE(X)+bE(Y)$;
- ✓ Если $X \geq Y$ при всех реализациях, то $E(X) \geq E(Y)$;
- ✓ Если X – непрерывная случайная величина с плотностью распределения $p(x)$, а $g(x)$, $x \in R$ – числовая функция, то для случайной величины Y= g(X) справедливо равенство $E(Y) = \int_{-\infty}^{+\infty} g(x)p(x)dx$;
- ✓ $E(a)= a$.

Другой важнейшей числовой характеристикой случайной величины X является ***<u>дисперсия</u>***, отражающая *степень «разброса» случайной величины относительно среднего значения*. Она определяется равенством $D(X) = E(X - EX)^2 = E(X^2) - (EX)^2$. Дисперсия имеет следующие свойства (X, Y– ***независимые*** случайные величины, a, b – константы):

- ✓ $D(aX+bY)=a^2 D(X)+b^2 D(Y)$;
- ✓ $D(a)= 0$.

Величину $\sigma = \sqrt{D(X)}$ называют ***<u>стандартным отклонением</u>*** случайной величины X.

Рассмотрим некоторые конкретные случайные величины, часто используемые в теории вероятностей, математической статистике и их приложениях.

1. ***<u>Биноминальное распределение</u>***. Дискретная случайная величина $\xi_n(p)$, принимающая значения $k=0,1,2,\ldots,n$ с вероятностями

$$p_k = P(\xi_n(p) = k) = C_n^k p^k (1-p)^{n-k}, \quad 0 < p < 1, \quad C_n^k = \frac{n!}{k!(n-k)!}$$

называется биноминальной случайной величиной с параметрами n и p. Случайная величина с таким распределением возникает в схеме Бернулли. Если случайные величины ε_i , $i=1, \ldots, n$, независимы и принимают значение 1 с вероятностью p и 0 с вероятностью $1-p$, то $\xi_n(p) = \sum_{i=1}^{n} \varepsilon_i$, $E\xi_n(p)= np$,

$D(\xi_n(p))= np(1- p)$.

2. ***Пуассоновское распределение***. Дискретная величина *П(λ)*, принимающая значения *k=0, 1,…*, с вероятностями $p_k = \frac{\lambda^k}{k!}e^{-\lambda}, \quad k = 0, 1, ..., \quad \lambda > 0$, называется пуассоновской случайной величиной с параметром *λ*. Пуассоновское распределение широко используется в теории массового обслуживания. Число *λ* носит название интенсивность. *E(П(λ))=D(П(λ))=λ.*

3. ***Равномерное распределение***. Непрерывная случайная величина *X*, плотность распределения которой задается формулой $p(x) = \begin{cases} \frac{1}{b-a}, & x \in [a,b], \\ 0, & x \notin [a,b] \end{cases}$, называется равномерной на отрезке *[a,b]*. $E(X) = \frac{b+a}{2}, \quad D(X) = \frac{(b-a)^2}{12}$.

4. ***Показательное (экспоненциальное) распределение***. Непрерывная величина *X*, плотность распределения которой задается формулой $p(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0, \\ 0, & x < 0 \end{cases}$, называется показательной или экспоненциальной с параметром *λ*. Это распределение находит широкое применение в демографических исследованиях. $E(X) = \frac{1}{\lambda}, \quad D(X) = \frac{1}{\lambda^2}$.

5. ***Нормальное (гауссовское) распределение***. Непрерывная случайная величина X, плотность распределения которой задается формулой $p(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp(-\frac{(x-m)^2}{2\sigma^2}), \quad x \in R, \quad m \in R, \quad \sigma > 0$, называется нормальной или гауссовской с параметрами *m* и σ^2. Часто используется обозначение $X \in N(m, \sigma^2)$. Нормальная случайная величина с *m=0* и $\sigma^2=1$ называется стандартной нормальной величиной. *E(X)=m*, $D(X)=\sigma^2$.

Существуют и другие виды распределений. Более подробно о них можно узнать в учебниках [2], [7].

Пример решения задач

Задача № 1. В корзине лежат 6 красных и 6 зеленых яблок. Для гостей случайным образом выбирают 5 яблок и кладут в вазу. Количество красных яблок в вазе – случайная величина *X*. Написать ряд распределения *X*, построить график функции распределения *X*, найти *EX* и *DX*.

Решение. Из 12 яблок любые 5 можно выбрать числом способов $n = C_{12}^5$.

Это число равно: $C_{12}^5 = \frac{12!}{5!7!} = 8 * 9 * 11 = 792$.

Случайная величина *X* может принимать значения 0, 1, 2, 3, 4, 5.

Если X=0, то следует выбирать 5 зеленых яблок из 6 возможных – это можно сделать числом способов $m_0 = C_6^5 = C_6^1 = 6$.

Если $X=1$, то следует выбирать 4 зеленых яблока из 6 возможных и 1 красное из 6 возможных– это можно сделать числом способов $m_1 = C_6^4 * 6 = 90$.

Если X=2, то следует выбирать 3 зеленых яблока из 6 возможных и 2 красных из 6 возможных– это можно сделать числом способов $m_2 = C_6^3 * C_6^2 = 300$.

Если $X=3$, то следует выбирать 2 зеленых яблока из 6 возможных и 3 красных из 6 возможных– это можно сделать числом способов $m_3 = C_6^2 * C_6^3 = 300 = m_2$.

Если $X=4$, то следует выбирать 1 зеленое яблоко из 6 возможных и 4 красных из 6 возможных– это можно сделать числом способов $m_4 = m_1 = C_6^4 * 6 = 90$.

Если $X=5$, то $m_5 = m_0 = 6$.

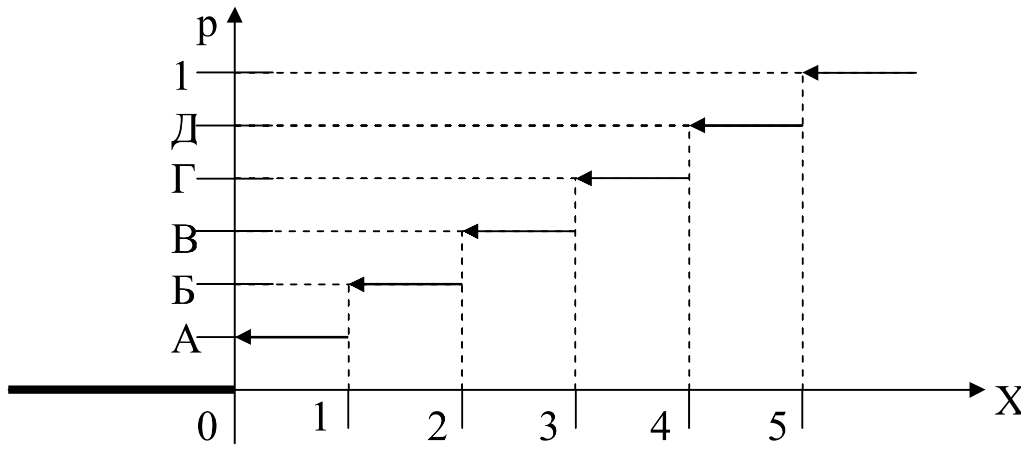
Вероятности $p(X=k)$, $k = 0,1,2,3,4,5$ равны отношениям m_k/n :

X	**0**	**1**	**2**	**3**	**4**	**5**
p	$\frac{6}{792}=\frac{1}{132}$	$\frac{90}{792}=\frac{15}{132}$	$\frac{300}{792}=\frac{50}{132}$	$\frac{300}{792}=\frac{50}{132}$	$\frac{90}{792}=\frac{15}{132}$	$\frac{6}{792}=\frac{1}{132}$

Функция распределения $F(x)=P(X<x)$ равна

- 0, если $\boldsymbol{x} \le 0$;
- 1/132, если $0 < \boldsymbol{x} \le 1$ (А);
- 16/132, если $1 < \boldsymbol{x} \le 2$ (Б);
- 66/132, если $2 < \boldsymbol{x} \le 3$ (В);
- 116/132, если $3 < \boldsymbol{x} \le 4$(Г);
- 131/132, если $4 < \boldsymbol{x} \le 5$ (Д);
- 1, если $\boldsymbol{x} > 5$.

График этой функции выглядит следующим образом:



$$EX = \sum p_k X_k = 2.5, \quad EX^2 = 7.04, \quad DX = EX^2 - (EX)^2 = 0.795 .$$

Задача № 2. Плотность вероятности случайной величины X задана соотношением $f(x)=\begin{cases}\dfrac{a}{x^4} & , \ если \quad x>10 \\ 0 & , \ иначе\end{cases}$.

Найти ***a, F(x)*** – функцию распределения случайной величины X, построить графики функций ***f(x)*** и ***F(x)***, вычислить EX и DX.

Решение. Постоянную ***a*** находим из условия нормировки плотности: $a\int_{10}^{\infty}\frac{dx}{x^4}=1$, откуда ***a = 3000***. Функция распределения ***F(x)=0*** при ***x ≤ 10***,

$F(x)=\int_{-\infty}^{x} f(x)dx = a\int_{10}^{x}\frac{dx}{x^4}=1-\frac{1000}{x^3}$ при ***x > 10***; $EX = a\int_{10}^{x}\frac{xdx}{x^4}=15$,

$EX^2 = a\int_{10}^{x}\frac{x^2dx}{x^4}=300$, $DX = EX^2-(EX)^2 = 75$.

Задача № 3. Случайная величина X ∈ N (1;2). Случайная величина Y связана с X функциональной зависимостью Y=6X+4. Найти g(y) – плотность вероятности случайной величины Y, EY, $DY=\sigma_y^2$. С помощью таблиц приближенно вычислить $P(|Y-EY|<2{,}75\sigma_y)$ и $P(\{0\le Y<10\}\cup\{Y\ge 25\})$.

Решение. Так как EX =1, DX =2, то EY =10, DY=36*2=72, $\sigma_y=\sqrt{72}=6\sqrt{2}\approx 8.5$. При линейном отображении распределение остается нормальным, поэтому $g(y)=\frac{1}{12\sqrt{\pi}}e^{-\frac{(y-10)^2}{144}}$. Вычисление значений функции распределения ***G(y)*** рекомендуется выполнить в EXCEL с применением статистической функции НОРМРАСП:

$$P(|Y-EY|<2{,}75\sigma_y)\approx P(-13<Y<33)=G(33)-G(-13)=0.9932$$

$$P(\{0\le Y<10\}\cup\{Y\ge 25\})=(G(10)-G(0))+(1-G(25))=0.4191$$

.

Задача № 4. Плотность вероятности случайной величины X задана соотношением $f(x)=\begin{cases}\dfrac{2}{3}x & , \ если \quad x\in(1;2) \\ 0 & , \ иначе\end{cases}$

Случайная величина связана с X функциональной зависимостью Y= X^2. Найти g(y) – плотность вероятности случайной величины Y, G(y) – функцию распределения случайной величины Y, EY, DY, $p=P(Y<\frac{EY}{3})$.

Решение. Случайная величина Y принимает ненулевые значения в промежутке (1,4), при этом плотность ***g(y)=f(ψ(y)) ψ′(y)***, где ***ψ*** — функция, обратная к заданной функции ***y=x²: ψ(y)=√y***. Поэтому $g(y) = \frac{2}{3}\sqrt{y}\,\frac{1}{2\sqrt{y}} = \frac{1}{3}$ и случайная величина Y распределена равномерно в промежутке (1,4). Функция распределения в указанном промежутке равна ***G(y)=(y-1)/3***, равна нулю при ***y ≤ 1***, равна 1 при ***y > 4***.

$$EY = \frac{5}{2}, \quad DY = \frac{3}{4}, \quad P(Y < \frac{EY}{3}) = P(Y < \frac{5}{6}) = 0 \cdot$$

Задача № 5. Случайные величины X, Y и Z независимы в совокупности. При этом X∈N(-2;6) и Y∈N(-4;2) распределены нормально, а Z – равномерно на интервале (-16;-10). Найти математическое ожидание и дисперсию случайной величины V=X—3Y—2Z+15.

Решение.

$$EX = -2, \quad EY = 4, \quad EZ = -13, \quad DX = 6, \quad DY = 2, \quad DZ = 3,$$

$$EV = EX - 3EY - 2EZ + 15 = 27, \quad DV = DX + 9DY + 4DZ = 36$$

Варианты контрольных заданий.

0. Вариант

1. В корзине лежат 2 красных и 3 зеленых яблока. Для гостей случайным образом выбирают 2 яблока и кладут в вазу. Количество красных яблок в вазе – случайная величина X. Написать ряд распределения X, построить график функции распределения X, найти EX и DX.

2. Плотность вероятности случайной величины X задана соотношением $f(x) = \begin{cases} ax^2 & , \text{ если } x \in (-1;2) \\ 0 & , \text{ иначе} \end{cases}$. Найти a, F(x) – функцию распределения случайной величины X, построить графики функций f(x) и F(x), вычислить EX и DX.

3. Случайная величина X ∈ N (1;2). Случайная величина Y связана с X функциональной зависимостью Y=2X–1. Найти g(y) – плотность вероятности случайной величины Y, EY, $DY = \sigma_y^2$. С помощью таблиц приближенно вычислить $P(|Y - EY| < 0{,}9\sigma_y)$ и $P(\{Y < 1\} \cup \{Y \geq 8\})$.

4. Плотность вероятности случайной величины X задана соотношением $f(x) = \begin{cases} \frac{1}{2}x & , \text{ если } x \in (0;2) \\ 0 & , \text{ иначе} \end{cases}$. Случайная величина связана с X функциональной зависимостью Y= X². Найти g(y) – плотность вероятности случайной величины Y, G(y) – функцию распределения случайной величины Y, EY, DY, $p = P(Y < \frac{EY}{3})$.

5. Случайные величины X, Y и Z независимы в совокупности. При этом $X \in N(-1;2)$ и $Y \in N(0;3)$ распределены нормально, а Z – равномерно на интервале (2;4). Найти математическое ожидание и дисперсию случайной величины V=2X-3Y+Z-4.

1. Вариант

1. В корзине лежат 2 красных и 4 зеленых яблока. Для гостей случайным образом выбирают 3 яблока и кладут в вазу. Количество красных яблок в вазе – случайная величина X. Написать ряд распределения X, построить график функции распределения X, найти EX и DX.

2. Плотность вероятности случайной величины X задана соотношением $f(x)=\begin{cases} ax^2 & , \ \text{если} \quad x \in (-2;3) \\ 0 & , \ \text{иначе} \end{cases}$. Найти a, F(x) – функцию распределения случайной величины X, построить графики функций f(x) и F(x), вычислить EX и DX.

3. Случайная величина $X \in N(1;2)$. Случайная величина Y связана с X функциональной зависимостью Y=3X–1. Найти g(y) – плотность вероятности случайной величины Y, EY, $DY=\sigma_y^2$. С помощью таблиц приближенно вычислить $P(|Y-EY|<1{,}0\sigma_y)$ и $P(\{1 \le Y < 20\})$.

4. Плотность вероятности случайной величины X задана соотношением $f(x)=\begin{cases} \frac{2}{9}x & , \ \text{если} \quad x \in (0;3) \\ 0 & , \ \text{иначе} \end{cases}$. Случайная величина связана с X функциональной зависимостью $Y=X^2$. Найти g(y) – плотность вероятности случайной величины Y, G(y) – функцию распределения случайной величины Y, EY, DY, $p=P(Y<\frac{EY}{3})$.

5. Случайные величины X, Y и Z независимы в совокупности. При этом $X \in N(0;2)$ и $Y \in N(-1;3)$ распределены нормально, а Z – равномерно на интервале (2;6). Найти математическое ожидание и дисперсию случайной величины V= -2X+3Y+Z-5.

2. Вариант

1. В корзине лежат 3 красных и 2 зеленых яблока. Для гостей случайным образом выбирают 2 яблока и кладут в вазу. Количество красных яблок в вазе – случайная величина X. Написать ряд распределения X, построить график функции распределения X, найти EX и DX.

2. Плотность вероятности случайной величины X задана соотношением $f(x)=\begin{cases} ax^2 & , \ \text{если} \quad x \in (-2;1) \\ 0 & , \ \text{иначе} \end{cases}$. Найти a, F(x) – функцию распределения случайной величины X, построить графики функций f(x) и F(x), вычислить EX и DX.

3. Случайная величина $X \in N(1;2)$. Случайная величина Y связана с X функциональной зависимостью Y=3X+2. Найти g(y) – плотность вероятности

случайной величины Y, EY, $DY = \sigma_y^2$. С помощью таблиц приближенно вычислить $P(|Y - EY| < 1{,}1\sigma_y)$ и $P(\{Y < 1{,}5\} \cup \{Y \ge 5\})$.

4. Плотность вероятности случайной величины X задана соотношением $f(x) = \begin{cases} \frac{1}{8}x \quad , \quad если \quad x \in (0;4) \\ \quad 0 \quad , \quad иначе \end{cases}$. Случайная величина связана с X функциональной зависимостью Y= X^2. Найти g(y) – плотность вероятности случайной величины Y, G(y) – функцию распределения случайной величины Y, EY, DY, $p = P(Y < \frac{EY}{3})$.

5. Случайные величины X, Y и Z независимы в совокупности. При этом X∈N(-2;1) и Y∈N(-3;2) распределены нормально, а Z – равномерно на интервале (-2;0). Найти математическое ожидание и дисперсию случайной величины V=2X-3Y-Z+4.

3. *Вариант*

1. В корзине лежат 3 красных и 2 зеленых яблока. Для гостей случайным образом выбирают 3 яблока и кладут в вазу. Количество красных яблок в вазе – случайная величина X. Написать ряд распределения X, построить график функции распределения X, найти EX и DX.

2. Плотность вероятности случайной величины X задана соотношением $f(x) = \begin{cases} ax^2 \quad , \quad если \quad x \in (-3;2) \\ \quad 0 \quad , \quad иначе \end{cases}$. Найти a, F(x) – функцию распределения случайной величины X, построить графики функций f(x) и F(x), вычислить EX и DX.

3. Случайная величина X ∈ N (1;2). Случайная величина Y связана с X функциональной зависимостью Y=-X–1. Найти g(y) – плотность вероятности случайной величины Y, EY, $DY = \sigma_y^2$. С помощью таблиц приближенно вычислить $P(|Y - EY| < 1{,}2\sigma_y)$ и $P(\{-2 \le Y < 40\})$.

4. Плотность вероятности случайной величины X задана соотношением $f(x) = \begin{cases} \frac{2}{25}x \quad , \quad если \quad x \in (0;5) \\ \quad 0 \quad , \quad иначе \end{cases}$. Случайная величина связана с X функциональной зависимостью Y= X^2. Найти g(y) – плотность вероятности случайной величины Y, G(y) – функцию распределения случайной величины Y, EY, DY, $p = P(Y < \frac{EY}{3})$.

5. Случайные величины X, Y и Z независимы в совокупности. При этом X∈N(-2;2) и Y∈N(-1;3) распределены нормально, а Z – равномерно на интервале (0;2). Найти математическое ожидание и дисперсию случайной величины V=—2X-3Y+Z+5.

4. Вариант

1. В корзине лежат 3 красных и 4 зеленых яблока. Для гостей случайным образом выбирают 2 яблока и кладут в вазу. Количество красных яблок в вазе – случайная величина X. Написать ряд распределения X, построить график функции распределения X, найти EX и DX.

2. Плотность вероятности случайной величины X задана соотношением $f(x)=\begin{cases} ax^2 & , \textit{если} \quad x\in(-3;3) \\ \quad 0 & , \textit{иначе} \end{cases}$. Найти a, F(x) – функцию распределения случайной величины X, построить графики функций f(x) и F(x), вычислить EX и DX.

3. Случайная величина X $\in$ N (1;2). Случайная величина Y связана с X функциональной зависимостью Y=2X–3. Найти g(y) – плотность вероятности случайной величины Y, EY, $DY=\sigma_y^2$. С помощью таблиц приближенно вычислить $P(|Y-EY|<1{,}3\sigma_y)$ и $P(\{Y<-6\}\cup\{Y\geq 0{,}5\})$.

4. Плотность вероятности случайной величины X задана соотношением $f(x)=\begin{cases} \frac{1}{18}x & , \textit{если} \quad x\in(0;6) \\ \quad 0 & , \textit{иначе} \end{cases}$. Случайная величина связана с X функциональной зависимостью Y= X^2. Найти g(y) – плотность вероятности случайной величины Y, G(y) – функцию распределения случайной величины Y, EY, DY, $p=P(Y<\frac{EY}{3})$.

5. Случайные величины X, Y и Z независимы в совокупности. При этом X$\in$N(-3;2) и Y$\in$N(0;3) распределены нормально, а Z – равномерно на интервале (-1;3). Найти математическое ожидание и дисперсию случайной величины V= -X+2Y-Z+6.

5. Вариант

1. В корзине лежат 3 красных и 4 зеленых яблока. Для гостей случайным образом выбирают 2 яблока и кладут в вазу. Количество красных яблок в вазе – случайная величина X. Написать ряд распределения X, построить график функции распределения X, найти EX и DX.

2. Плотность вероятности случайной величины X задана соотношением $f(x)=\begin{cases} ax^2 & , \textit{если} \quad x\in(-4;2) \\ \quad 0 & , \textit{иначе} \end{cases}$. Найти a, F(x) – функцию распределения случайной величины X, построить графики функций f(x) и F(x), вычислить EX и DX.

3. Случайная величина X $\in$ N (1;2). Случайная величина Y связана с X функциональной зависимостью Y=2X–4. Найти g(y) – плотность вероятности случайной величины Y, EY, $DY=\sigma_y^2$. С помощью таблиц приближенно вычислить $P(|Y-EY|<1{,}4\sigma_y)$ и $P(\{-4\leq Y<10\})$.

4. Плотность вероятности случайной величины X задана соотношением $f(x)=\begin{cases}\frac{2}{49}x & , \text{ } если \quad x\in(0;7)\\ 0 & , \text{ } иначе\end{cases}$. Случайная величина связана с X функциональной зависимостью Y= X^2. Найти g(y) – плотность вероятности случайной величины Y, G(y) – функцию распределения случайной величины Y, EY, DY, $p=P(Y<\frac{EY}{3})$.

5. Случайные величины X, Y и Z независимы в совокупности. При этом $X\in N(-2;3)$ и $Y\in N(-2;6)$ распределены нормально, а Z – равномерно на интервале (-1;5). Найти математическое ожидание и дисперсию случайной величины V=X-2Y-Z+6.

6. Вариант

1. В корзине лежат 3 красных и 4 зеленых яблока. Для гостей случайным образом выбирают 4 яблока и кладут в вазу. Количество красных яблок в вазе – случайная величина X. Написать ряд распределения X, построить график функции распределения X, найти EX и DX.

2. Плотность вероятности случайной величины X задана соотношением $f(x)=\begin{cases}ax^2 & , \text{ } если \quad x\in(-2;4)\\ 0 & , \text{ } иначе\end{cases}$. Найти a, F(x) – функцию распределения случайной величины X, построить графики функций f(x) и F(x), вычислить EX и DX.

3. Случайная величина X $\in$ N (1;2). Случайная величина Y связана с X функциональной зависимостью Y=–2X+3. Найти g(y) – плотность вероятности случайной величины Y, EY, $DY=\sigma_y^2$. С помощью таблиц приближенно вычислить $P(|Y-EY|<1{,}5\sigma_y)$ и $P(\{-1\le Y<22\})$.

4. Плотность вероятности случайной величины X задана соотношением $f(x)=\begin{cases}\frac{1}{32}x & , \text{ } если \quad x\in(0;8)\\ 0 & , \text{ } иначе\end{cases}$. Случайная величина связана с X функциональной зависимостью Y= X^2. Найти g(y) – плотность вероятности случайной величины Y, G(y) – функцию распределения случайной величины Y, EY, DY, $p=P(Y<\frac{EY}{3})$.

5. Случайные величины X, Y и Z независимы в совокупности. При этом $X\in N(-1;4)$ и $Y\in N(-2;4)$ распределены нормально, а Z – равномерно на интервале (-2;4). Найти математическое ожидание и дисперсию случайной величины V=X--2Y+Z+6.

7. Вариант

1. В корзине лежат 3 красных и 4 зеленых яблока. Для гостей случайным образом выбирают 5 яблок и кладут в вазу. Количество красных яблок в вазе – случайная величина X. Написать ряд распределения X, построить график функции распределения X, найти EX и DX.

2. Плотность вероятности случайной величины X задана соотношением $f(x)=\begin{cases} ax^2 & , \text{ если } x\in(-5;-3) \\ 0 & , \text{ иначе} \end{cases}$. Найти a, F(x) – функцию распределения случайной величины X, построить графики функций f(x) и F(x), вычислить EX и DX.

3. Случайная величина X ∈ N (1;2). Случайная величина Y связана с X функциональной зависимостью Y=–5X+6. Найти g(y) – плотность вероятности случайной величины Y, EY, $DY=\sigma_y^2$. С помощью таблиц приближенно вычислить $P(|Y-EY|<1{,}6\sigma_y)$ и $P(\{-1\le Y<18\})$.

4. Плотность вероятности случайной величины X задана соотношением $f(x)=\begin{cases} \frac{2}{81}x & , \text{ если } x\in(0;9) \\ 0 & , \text{ иначе} \end{cases}$. Случайная величина связана с X функциональной зависимостью Y= X^2. Найти g(y) – плотность вероятности случайной величины Y, G(y) – функцию распределения случайной величины Y, EY, DY, $p=P(Y<\frac{EY}{3})$.

5. Случайные величины X, Y и Z независимы в совокупности. При этом X∈N(-4;1) и Y∈N(2;2) распределены нормально, а Z – равномерно на интервале (-3;1). Найти математическое ожидание и дисперсию случайной величины V= -X-2Y+Z+6.

8. Вариант

1. В корзине лежат 5 красных и 3 зеленых яблока. Для гостей случайным образом выбирают 2 яблока и кладут в вазу. Количество красных яблок в вазе – случайная величина X. Написать ряд распределения X, построить график функции распределения X, найти EX и DX.

2. Плотность вероятности случайной величины X задана соотношением $f(x)=\begin{cases} ax^2 & , \text{ если } x\in(-5;-1) \\ 0 & , \text{ иначе} \end{cases}$. Найти a, F(x) – функцию распределения случайной величины X, построить графики функций f(x) и F(x), вычислить EX и DX.

3. Случайная величина X ∈ N (1;2). Случайная величина Y связана с X функциональной зависимостью Y=7X–5. Найти g(y) – плотность вероятности случайной величины Y, EY, $DY=\sigma_y^2$. С помощью таблиц приближенно вычислить $P(|Y-EY|<1{,}7\sigma_y)$ и $P(\{Y<3{,}2\})$.

4. Плотность вероятности случайной величины X задана соотношением $f(x)=\begin{cases} \frac{1}{50}x & , \text{ если } x\in(0;10) \\ 0 & , \text{ иначе} \end{cases}$. Случайная величина связана с X функциональной зависимостью Y= X^2. Найти g(y) – плотность вероятности случайной величины Y, G(y) – функцию распределения случайной величины Y, EY, DY, $p=P(Y<\frac{EY}{3})$.

5. Случайные величины X, Y и Z независимы в совокупности. При этом X∈N(-2;3) и Y∈N(-1;2) распределены нормально, а Z – равномерно на интервале (-5;-1). Найти математическое ожидание и дисперсию случайной величины V=2X+5Y-Z-8.

9. Вариант

1. В корзине лежат 4 красных и 3 зеленых яблок. Для гостей случайным образом выбирают 3 яблока и кладут в вазу. Количество красных яблок в вазе – случайная величина X. Написать ряд распределения X, построить график функции распределения X, найти EX и DX.

2. Плотность вероятности случайной величины X задана соотношением $f(x)=\begin{cases} ax^2 \quad , \quad \text{если} \quad x\in(-5;1) \\ 0 \quad , \quad \text{иначе} \end{cases}$. Найти a, F(x) – функцию распределения случайной величины X, построить графики функций f(x) и F(x), вычислить EX и DX.

3. Случайная величина X ∈ N (1;2). Случайная величина Y связана с X функциональной зависимостью Y=4X–2. Найти g(y) – плотность вероятности случайной величины Y, EY, $DY=\sigma_y^2$. С помощью таблиц приближенно вычислить $P(|Y-EY|<1{,}8\sigma_y)$ и $P(\{-20\le Y<3{,}5\})$.

4. Плотность вероятности случайной величины X задана соотношением $f(x)=\begin{cases} -\dfrac{1}{2}x^2 \quad , \quad \text{если} \quad x\in(-2;0) \\ 0 \quad , \quad \text{иначе} \end{cases}$. Случайная величина связана с X функциональной зависимостью Y= X^2. Найти g(y) – плотность вероятности случайной величины Y, G(y) – функцию распределения случайной величины Y, EY, DY, $p=P(Y<\dfrac{EY}{3})$.

5. Случайные величины X, Y и Z независимы в совокупности. При этом X∈N(-2;3) и Y∈N(-1;2) распределены нормально, а Z – равномерно на интервале (-5;-1). Найти математическое ожидание и дисперсию случайной величины V= -2X+5Y-Z+8.

Контрольное задание №2.

Теоретический материал по данной теме будет дополнительно изложен на установочной сессии.

Практические и лабораторные занятия с использованием программ EXCEL и STATISTICA проводятся в компьютерном классе. Однако желающие могут освоить данную тему самостоятельно по любому пособию – см. книги [3,4,6,7] в списке литературы.

Предлагаемые задачи выполняются «вручную» или же (по возможности!) с применением электронных таблиц EXCEL. Студенту следует выбрать для решения задачу, номер которой совпадает с последней цифрой его номера в студенческом билете или зачетной книжки. Подробные методические указания по решению задач этого типа можно найти в учебном пособии [5].

Методические указания

Метод наименьших квадратов (МНК) предназначен для решения избыточной системы нормальных уравнений. В эконометрике такая система образуется при оценивании параметров линейной регрессии.

Пусть по группе предприятий, выпускающих один и тот же вид продукции, необходимо определить зависимость затрат на производство (y – определяемая переменная) от объема выпуска продукции (x – определяющая величина). Эта задача решается методами линейной регрессии, которая сводится к решению уравнения вида

$$y=a+bx+\varepsilon \qquad (1).$$

В нашем случае x – объем выпуска продукции; y – затраты на производство; ε – случайная величина, характеризующая отклонение реального значения определяемого признака от теоретического, получаемого по формуле

$$y=a+bx \qquad (1`).$$

Значения y и x варьируются в зависимости от предприятия. Поэтому для определения теоретического значения этих величин необходимо в уравнении (1) найти коэффициенты a и b. Тогда для любого аналогичного предприятия можно будет спрогнозировать затраты на производство по заданному объему продукции.

Пусть имеется n предприятий. Тогда уравнение (1) превратится в систему уравнений

$$\begin{cases} y_1 = a + bx_1 + \varepsilon_1 \\ y_2 = a + bx_2 + \varepsilon_2 \\ \ldots \\ y_n = a + bx_n + \varepsilon_n \end{cases} \qquad (2),$$

где x_i, y_i – затраты на производство и объем выпущенной продукции i-го предприятия. Или в матричной форме:

$$У=a+вX+E \qquad (2`),$$

где

$$Y = \begin{vmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{vmatrix}; \quad X = \begin{vmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{vmatrix}; \quad E = \begin{vmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \dots \\ \varepsilon_n \end{vmatrix}; \quad .$$

Система (2), (2`) – избыточная: неизвестных два – a и b, а уравнений n.Она решается МНК, который заключается в следующем. Система (2) решается, исходя из условия

$$\sum_{i=1}^{n} \varepsilon_i^2 \to \min;$$

если обозначить сумму квадратов *Φ(a,b),* то

$$\Phi(a,b) = \sum_{i=1}^{n} (a + bx_i - y_i)^2 \to \min . \tag{3}$$

Чтобы найти минимум функции (3), необходимо вычислить частные производные по каждому из параметров a и b и приравнять их нулю:

$$\frac{\partial \Phi}{\partial a} = 0, \ \frac{\partial \Phi}{\partial b} = 0 .$$

Вычисляя производные, приходим к системе уравнений

$$\begin{cases} na + b\sum x_i = \sum y_i \\ a\sum x_i + b\sum x_i^2 = \sum x_i y_i \end{cases} \tag{4}.$$

Решая систему нормальных уравнений (4) либо методом последовательного исключения переменных, либо методом определителей, найдем искомые оценки параметров a и b:

$$a = \overline{y} - b\overline{x}\ , \quad b = \frac{\overline{yx} - \overline{x} * \overline{y}}{\frac{1}{n}\sum \ x^2 - \overline{x}^2} \tag{5},$$

где

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{1} x_i$$
$$\overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \ .$$

<u>Параметр *b* называется коэффициентом регрессии</u>. Его величина показывает среднее изменение результата с изменением фактора на одну единицу.

Величина в знаменателе параметра b

$$\sigma_x^2 = \frac{1}{n}\sum x^2 - \overline{x}^2 \tag{6}$$

называется дисперсией признака x. Коэффициент регрессии можно записать иначе

$$b = \frac{\overline{yx} - \overline{x} * \overline{y}}{\sigma_x^2} .$$

Уравнение регрессии всегда дополняется показателем тесноты связи. При использовании линейной регрессии в качестве такого показателя выступает линейный коэффициент корреляции

$$r_{xy} = b\frac{\sigma_x}{\sigma_y} = \frac{\overline{yx} - \overline{x} * \overline{y}}{\sigma_x \sigma_y} \tag{7}.$$

Линейный коэффициент корреляции находится в границах: $-1 \le r_{xy} \le 1$.

Если коэффициент регрессии $b>0$, то $0 \le r_{xy} \le 1$, и, наоборот, при $b<0$, $-1 \le r_{xy} \le 0$. Близость модуля линейного коэффициента корреляции к 1 означает, что зависимость величин x и y близка к линейной.

Пример решения задач «вручную».

Пусть для 7 предприятий дан объем выпуска продукции (x) и затраты на его производство (y). Найти затраты на производство при изменении объема выпускаемой продукции на исследуемых предприятиях. Исходная информация и результаты расчета сведены в таблицу 1:

Таблица1

Расчетная таблица

№ предприятия	Выпуск продукции (x)	Затраты на производство (y)	y*x	x^2	y^2	Расчетные затраты на производство
1.	1	30	30	1	900	31,1
2.	2	70	140	4	4900	67,9
3.	4	150	600	16	22500	141,6
4.	3	100	300	9	10000	104,7
5.	5	170	850	25	28900	178,4
6.	3	100	300	9	10000	104,7
7.	4	150	600	16	22500	141,6
итого	**22**	**770**	**2820**	**80**	**99700**	**770,0**

Система нормальных уравнений будет иметь вид $\begin{cases} 7*a+22*b=70 \\ 22*a+80*b=2820 \end{cases}$.Решая её, получим: $a=-5,79$ и $b=36,84$. Тогда теоретическое уравнение регрессии (1’)примет вид $y=-5,79+36,84x$.

Коэффициент регрессии ($b=36,84$)покажет на сколько единиц изменятся затраты на производство (y) при изменении объема продукции (x) на единицу. Пользуясь полученной формулой, можно спрогнозировать изменение затрат на производство (y) при изменении объема производимой продукции (x).

Далее имеем: $\begin{matrix} \bar{x}=3,14; & \sigma_x=1,25; \\ \bar{y}=110; & \sigma_y=46,29; \end{matrix}$ $r_{xy}=0,991$, что достаточно близко к 1 и означает наличие очень тесной зависимости затрат на производство от величины объема выпущенной продукции.

Пример решения задач с применением электронных таблиц EXCEL.

1. Встроенная статистическая функция **ЛИНЕЙН** определяет параметры линейной регрессии $y=a+bx$. Порядок вычисления следующий:
   1.1. Введите исходные данные или откройте существующий файл, содержащий анализируемые данные.

1.1.1. В нашей задаче в колонку А заносим параметры Х, в колонку В заносим параметры У (Рис.1):

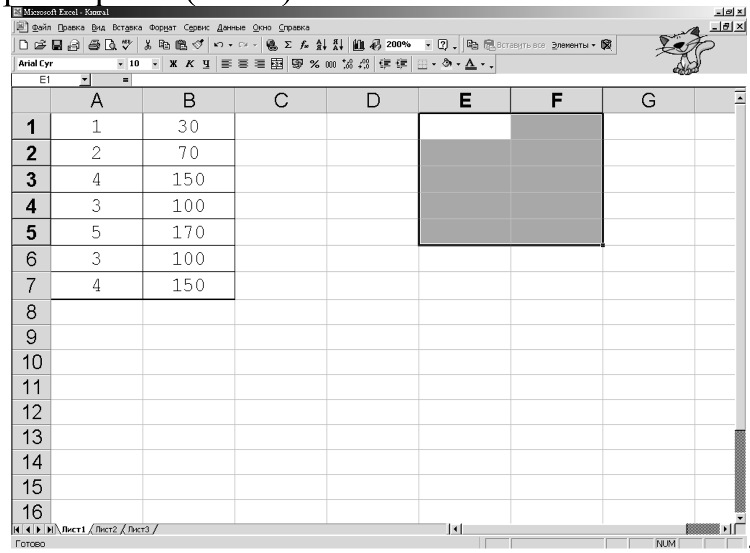


Рис.1

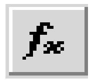
1.2. Выделите область пустых ячеек –2 столбца и 5 строк для результатов регрессионной статистики.

1.2.1. На Рис.1 это область E1:F5.

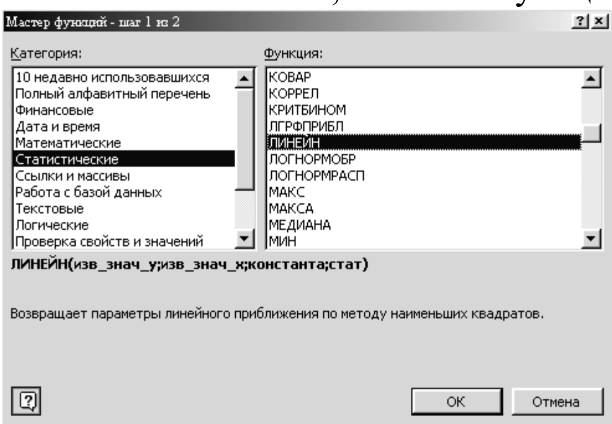
1.3. Активизируйте Мастер функций любым из способов:

1.3.1. В главном меню выберите **Вставка/Функция**.

1.3.2. На панели инструментов **Стандартная** щелкнуть по кнопке **Вставка функции**.



1.4. В появившемся Мастере функций (Рис.2) в окне **Категории** выбрать **Статистические**, а в окне Функция – **ЛИНЕЙН**. Щелкните по кнопке **ОК**.



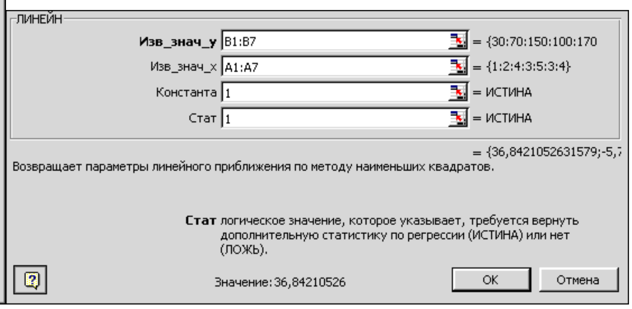


Рис.2 Рис.3

1.5. Заполните аргументы функции (Рис.3):

1.5.1. *Известные значения у* – диапазон, содержащий данные результативного признака;

Известные значения х – диапазон, содержащий данные факторов независимого признака;

Константа – логическое значение, которое указывает на наличие или на отсутствие свободного члена в уравнении; если *Константа*=1, то свободный член рассчитывается обычным образом, если *Константа*=0, то свободный член равен 0;

Статистика – логическое значение, которое указывает выводить дополнительную информацию по регрессионному анализу или нет. Если *Статистика*=1, то дополнительная информация выводится, если *Статистика*=0, то выводятся только оценки параметров уравнения.

Щелкните по кнопке **ОК**.

1.5.2. Заполняя аргументы функции рассматриваемой задачи, получаем следующие данные (Рис.3):

Известные значения y – B1:B7;
Известные значения x – A1:A7;
Константа – 1;
Статистика – 1.

1.6. В левой верхней ячейке выделенной области появится первый элемент итоговой таблицы. Чтобы раскрыть таблицу необходимо нажать <F2>, а затем – комбинацию клавиш <CTRL>+<SHIFT>+<ENTER>.

1.6.1.

b	36,84211	a	-5,78947
S_b	2,201737	S_a	7,443229
r^2	0,982456	S_y	7,254763
F	280	N	5
RSS	14736,84	ESS	263,1579

2. С помощью инструмента анализа данных Регрессия, помимо результатов регрессионной статистики, дисперсионного анализа и доверительных интервалов, можно получить остатки и графики подбора линий регрессии, остатков и нормальной вероятности. Порядок действий следующий:

2.1. Необходимо проверить доступ к пакету анализа. В главном меню последовательно выберите **Сервис/Надстройки**. Установите флажок **Пакет анализа** (Рис.4);

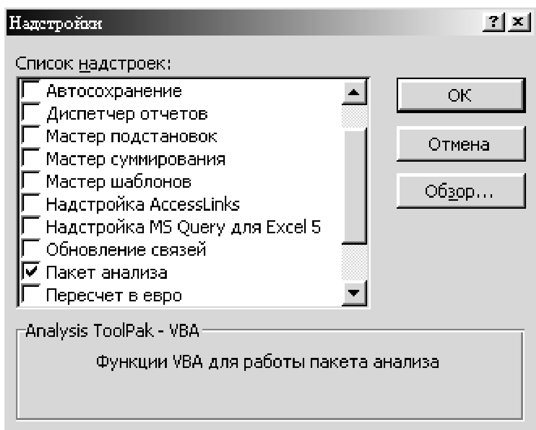


Рис.4

2.2. В главном меню необходимо выбрать **Сервис/Анализ** данных/**Регрессия**. Щелкнуть по кнопке **ОК**;

2.3. Заполнение диалогового окна ввода данных и параметров вывода (Рис.5):

2.3.1. *Входной интервал у* – диапазон, содержащий данные результативного признака;

Входной интервал х – диапазон, содержащий данные факторов независимого признака;

Метки – флажок, который указывает, содержит ли первая строка названия столбцов или нет*;*

Константа—ноль – флажок, указывающий на наличие или отсутствие свободного члена в уравнении*;*

Выходной интервал – достаточно указать левую верхнюю ячейку будущего диапазона;

Новый рабочий лист – можно задать произвольное имя нового листа.

Если необходимо получить информацию и графики остатков, необходимо установить соответствующие флажки в диалоговом окне.

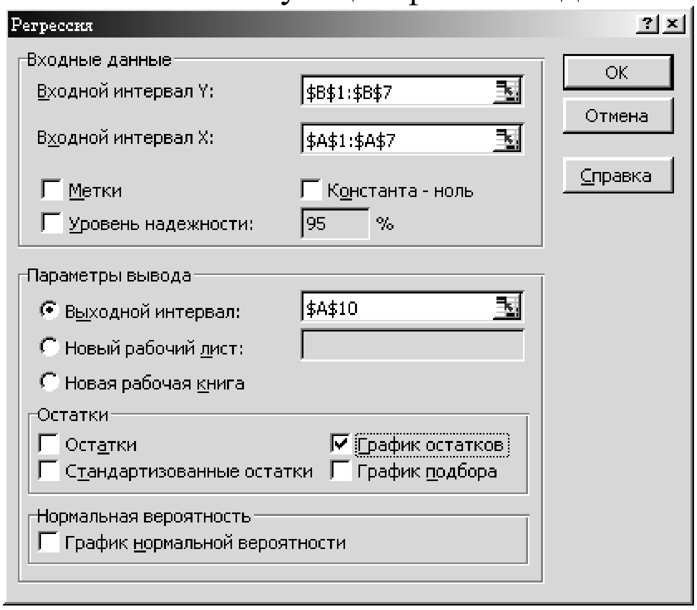
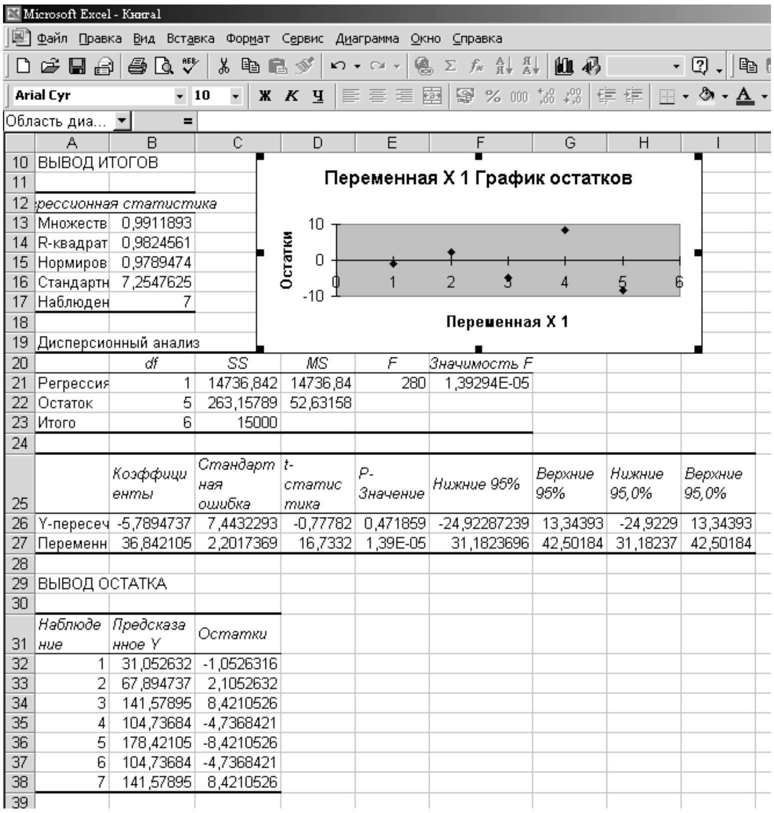


Рис.5

2.3.2. Данные для рассматриваемой задачи представлены на Рис.5.

2.4. На Рис.6 показаны результаты, получившиеся вследствие применения инструмента **Анализа данных Регрессия**.



	A	B	C	D	E	F	G	H	I
10	ВЫВОД ИТОГОВ								
11									
12	рессионная статистика								
13	Множеств	0,9911893							
14	R-квадрат	0,9824561							
15	Нормиров	0,9789474							
16	Стандартн	7,2547625							
17	Наблюден	7							
18									
19	Дисперсионный анализ								
20		df	SS	MS	F	Значимость F			
21	Регрессия	1	14736,842	14736,84	280	1,39294E-05			
22	Остаток	5	263,15789	52,63158					
23	Итого	6	15000						
24									
25		Коэффициенты	Стандартная ошибка	t-статистика	P-Значение	Нижние 95%	Верхние 95%	Нижние 95,0%	Верхние 95,0%
26	Y-пересеч	-5,7894737	7,4432293	-0,77782	0,471859	-24,92287239	13,34393	-24,9229	13,34393
27	Переменн	36,842105	2,2017369	16,7332	1,39E-05	31,1823696	42,50184	31,18237	42,50184
28									
29	ВЫВОД ОСТАТКА								
30									
31	Наблюдение	Предсказанное Y	Остатки						
32	1	31,052632	-1,0526316						
33	2	67,894737	2,1052632						
34	3	141,57895	8,4210526						
35	4	104,73684	-4,7368421						
36	5	178,42105	-8,4210526						
37	6	104,73684	-4,7368421						
38	7	141,57895	8,4210526						
39									

Рис.6

Если студент не имеет возможности работать на персональном компьютере дома или на работе, такая возможность ему будет предоставлена в Университете.

Варианты контрольных заданий.

0. *Вариант*

В таблице представлены статистические данные о размере товарооборота X и суммы издержек обращения Y по девяти магазинам:

Товарооборот X	480	510	530	540	570	590	680	640	650
Издержки обращения Y	31	25	31	28	29	32	36	36	37

Составить уравнение линии регрессии. Предсказать сумму издержек обращения при товарообороте 500 и 640, а также изменение суммы издержек обращения при изменении товарооборота не единицу. Найти коэффициент корреляции.

1. *Вариант*

Обозначим через X цену оптовой продажи некоторого товара, через Y— цену его розничной продажи. Составить уравнение прямой линии регрессии, вычислить коэффициент корреляции и предсказать цену розничной продажи товара при оптовой цене 70 и 80. Найти среднее изменение в Y, соответствующее единичному изменению в X.

X	80	79	77	76	76	76	74	72	70	71	69	70
Y	84	82	81	82	81	86	83	82	82	82	82	81

2. *Вариант*

В таблице приводится связь между валовым национальным продуктом (ВНП) на душу населения (X) и процентным показателем (Y) грамотности взрослого населения для 22 стран и территорий в середине 1970-х гг. Написать уравнение прямой линии регрессии, вычислить коэффициент корреляции, установить изменение показателя грамотности, соответствующее единичному изменению ВНП. Предсказать уровень грамотности населения для стран с ВНП, равным 40 и 1000.

	X	Y		X	Y
Непал	45	5,0	Брит.Гвиана	235	74,0
Бирма	57	47,5	Гонконг	272	57,5
Уганда	64	27,5	Панама	329	65,7
Южн.Вьетнам	76	17,5	Ливан	362	47,5
Таиланд	96	68,0	Сингапур	400	50,0
Гаити	105	10,5	Аргентина	490	86,4
Индонезия	131	17,5	Исландия	572	98,5
Южн.Корея	144	77,0	Чехословакия	680	97,5
Гана	172	22,5	Франция	943	96,4
Перу	179	47,5	Новая Зеландия	1310	98,5
Сальвадор	219	39,4	Канада	1947	97,5

3. *Вариант*

Таблица содержит данные о росте (X) и массе (Y) 25 выбранных наугад студентов. Найти линию регрессии и коэффициент корреляции, предсказать массу студентов, имеющих рост 175 и 180, а также среднее изменение массы студента при изменении роста на единицу.

X	175	188	178	165	175	185	183	175	183	193	188	183	185
Y	63	95	67	66	83	75	70	77	79	70	84	84	77

X	173	178	180	173	185	165	185	188	163	183	183	170
Y	75	100	84	82	77	61	79	82	68	77	75	66

4. *Вариант*

На основе данных о годовой производительности труда в расчете на одного рабочего (Y) и энерговооруженности труда (X) на предприятиях одной отрасли построить линию регрессии, найти коэффициент корреляции, предсказать производительность труда при энерговооруженности 9 и 10, а также изменение производительности труда при единичном изменении энерговооруженности.

X	6,7	6,9	7,2	7,3	8,4	8,8	9,1	9,8	10,6	10,7	11,1	11,8	12,1	12,4
Y	2,8	2,2	3,0	3,5	3,2	3,7	4,0	4,8	6,0	5,4	5,2	5,4	6,0	9,0

5. *Вариант*

Случайная выборка 10 фармацевтических фирм показала следующее соотношение между прибылью (Y) и затратами на научные исследования (X):

X	20	40	40	30	10	40	50	30	20	50
Y	50	50	60	40	10	30	50	50	30	60

Составить уравнение прямой линии регрессии, найти коэффициент корреляции. Спрогнозировать прибыль, если затраты на научные исследования составили 50 и 45, а также изменение прибыли при увеличении затрат на единицу.

6. *Вариант*

На 10 территориях были измерены процентный показатель перенаселенности (X) и показатель детской смертности (Y). По представленным в таблице данным определить уравнение прямой линии регрессии, коэффициент корреляции, а также предсказать показатели детской смертности для территорий с показателями перенаселенности 5 и 10.

X	13	33	12	40	12	7	20	4	15	26
Y	124	151	124	156	128	78	127	104	127	144

7. *Вариант*

Найти уравнение прямой линии регрессии и коэффициент корреляции для фондоотдачи оборудования (X) и удельного веса продукции высшей категории качества (Y):

X	1,47	1,25	1,82	1,45	1,75	1,37	1,61	1,93	1,68	1,66
Y	34,08	35,89	36,93	32,31	34,91	30,20	31,23	48,13	30,08	42,86

8. *Вариант*

В таблице содержатся данные, показывающие связь между количеством дней (X), проведенных пациентами в больнице, и затратами больницы (Y), которые компенсируются страховой компанией.

X	1	3	6	7	2	4	12	15	5	9
Y	50	175	180	200	60	140	420	540	170	300

Вычислить коэффициент корреляции. По уравнению линии регрессии спрогнозировать затраты больницы при количестве дней 5 и 10. каково изменение затрат при увеличении дней госпитализации больного на единицу.

9. *Вариант*

Крупная корпорация проверяет соотношение между прибылью (Y) и процентом используемых производственных мощностей (X) для каждого из 12 заводов, входящих в корпорацию. Найти уравнение линии регрессии, коэффициент корреляции. Предсказать прибыль при проценте использования производственных мощностей 60 и 70.

X	50	57	61	68	77	80	82	85	89	91	95	99
Y	2,5	6,2	3,1	4,6	7,3	4,5	6,1	11,6	10,0	14,2	16,1	19,5

Литература

1. Айвазян С.А., Мхитарян В.С. Прикладная статистика. Основы эконометрики. В 2 т.- М. ЮНИТИ. 2001 - Т.1 - 656 с., Т.2 - 432 с.

2. Гмурман В.Е. Теория вероятностей и математическая статистика. М. Высшая школа, 1999 480 с.

3. Доугерти К. Введение в эконометрику. М. Инфра-М. 2001. - 402 с.

4. Елисеева И.И., Курышева С.В., Костеева Т.В. и др.Эконометрика. Под ред. И.И.Елисеевой.- М. «Финансы и статистика», 2001.- 344 с.

5. Елисеева И.И., Курышева С.В., Гордиенко Н.М.и др. Практикум по эконометрике. Под ред. И.И.Елисеевой - М. «Финансы и статистика», 2001 - 192 с.

6. Кремер Н.Ш., Путко Б.А. Эконометрика. - М. ЮНИТИ-ДАНА, 2002 - 311 с.

7. Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика. Начальный курс- М. Дело, 2001 - 400 с.

8. Орлов А.И. Эконометрика. Начальный курс- М. Изд-во "Экзамен", 2002 - 576 с.