МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ
БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
«САМАРСКИЙ ГОСУДАРСТВЕННЫЙ АЭРОКОСМИЧЕСКИЙ
УНИВЕРСИТЕТ ИМЕНИ АКАДЕМИКА С.П. КОРОЛЕВА
(НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ)» (СГАУ)

ЭКОНОМЕТРИКА

Рекомендовано к изданию редакционно-издательским советом федерального государственного бюджетного образовательного учреждения высшего профессионального образования «Самарский государственный аэрокосмический университет имени академика С. П. Королева (национальный исследовательский университет) в качестве мультимедийного электронного лабораторного практикума для бакалавров в системе дистанционного обучения «MOODLE»

САМАРА
2013

ББК У9(2) 21.0
Э40

Автор-составитель: **Озерная Светлана Алексеевна**

Рецензент канд. техн. наук, доц. каф. общей информатики СГАУ, В. Г. М и х а й л о в

Эконометрика [Электронный ресурс] : мультимед. электрон. лабор. практикум для бакалавров в системе дистанц. обучения «MOODLE» / сост. С. А. Озерная. – Электрон. текстовые и граф. дан. (0,98 Мб). – Самара: Изд-во СГАУ, 2013. – 1 эл. опт. диск (CD-ROM).

В состав электронного мультимедийного пособия входят:
1. Эконометрика. Методические указания к лабораторному практикуму в системе дистанционного обучения «MOODLE».
2. Вопросы контроля выполнения лабораторного практикума.

Приводятся методы решения эконометрических задач по курсу эконометрика в 5 семестре, примеры решённых задач.

Предназначен для студентов факультета экономики и управления для работы по направлениям подготовки бакалавров 080100.62 «Экономика», 080200.62 «Менеджмент», 080500.62 «Бизнес-информатика.

Подготовлено на кафедре математических методов в экономике..

© Самарский государственный
аэрокосмический университет, 2013

ЭКОНОМЕТРИКА. МЕТОДИЧЕСКИЕ УКАЗАНИЯ К ЛАБОРАТОРНОМУ ПРАКТИКУМУ В СИСТЕМЕ ДИСТАНЦИОННОГО ОБУЧЕНИЯ «MOODLE»

СОДЕРЖАНИЕ

ЛАБОРАТОРНАЯ РАБОТА №1. ПАРНАЯ ЛИНЕЙНАЯ РЕГРЕССИЯ

Студентам необходимо осваивать основные методы, способы и средства получения, хранения, переработки информации (ОК4-10, 11-15, 20). В результате изучения и теоретических, и практических основ курса эконометрики студент должен приобрести навыки работы с компьютером как средством управления информацией, выработать способность работать с информацией в глобальных сетях (ПК2-3, 14, 20).

Цель изучения и теоретических, и практических основ курса эконометрики – получение практических навыков работы (ПК14, 20) с применением пакета ***EXCEL*** для решения математических и эконометрических задач.

В экономике каждому значению одной переменной соответствует определенное (условное) распределение другой переменной. Такая зависимость получила название статистической.

Задачами регрессионного анализа являются установление формы зависимости между переменными, оценка функции регрессии, оценка неизвестных значений (прогноз значений) зависимой переменной.

Цель работы. Пусть имеется набор значений двух переменных: y_i – эндогенная (или результат); x_i – экзогенная (или фактор).

Подобрать по данным наблюдений (x_i, y_i), $i=1,\dots,n$ функцию линейной регрессии $y_{лин\,р}(x) = a + b \cdot x$, наилучшим образом описывающую предполагаемую зависимость.

Порядок проведения работы

1. Применить метод определителей для получения параметров линейной регрессии.

На основании необходимого условия экстремума функции двух переменных после преобразования получим систему нормальных уравнений для определения параметров a и b линейной регрессии:

$$\begin{cases} na + b\sum x = \sum y \\ a\sum x + b\sum x^2 = \sum yx \end{cases}. \tag{1}$$

Согласно методу наименьших квадратов неизвестные параметры a и b выбираются таким образом, чтобы сумма квадратов отклонений эмпирических значений зависимой переменной y от значений, найденных по уравнению регрессии, была минимальной. Решить систему уравнений, т.е. получить коэффициенты a и b, методом определителей.

Коэфициенты a и b можно получить и по другим формулам:

$$b = \frac{\overline{xy} - \overline{x}.\overline{y}}{\sigma_x^2}, \qquad \sigma_x^2 = \frac{\sum x^2}{n} - \overline{x}^2,$$

$$a = \overline{y} - b\overline{x}, \qquad \sigma_y^2 = \frac{\sum y^2}{n} - \overline{y}^2.$$

где $\overline{y}$ – среднее значение y, $\overline{xy}$ – среднее значение произведения x и y, $\overline{x}$ – среднее значение x.

2. Получить линейный коэффициент корреляции r_{xy} и проанализировать тесноту связи.

Уравнение регрессии дополняется показателем тесноты связи. При использовании линейной регрессии в качестве такого показателя выступает линейный коэффициент корреляции $\boldsymbol{r_{xy}}$:

$$r_{yx} = b\frac{\sigma_x}{\sigma_y} = \frac{\overline{xy} - \bar{x}.\bar{y}}{\sigma_x.\sigma_y}. \qquad (2)$$

Линейный коэффициент корреляции находится в определенных пределах: (-1)<= $\boldsymbol{r_{xy}}$ <=(+1). При этом, чем ближе $\boldsymbol{r_{xy}}$ к нулю, тем слабее корреляция, а при $\boldsymbol{r_{xy}}$=0 линия регрессии параллельна оси $\boldsymbol{x}$; чем ближе $\boldsymbol{r_{xy}}$ к (-1) или к (+1), тем сильнее корреляция, т.е. зависимость $\boldsymbol{x}$ и $\boldsymbol{y}$ близка к линейной.

Если коэффициент регрессии $\boldsymbol{b}$>0, то 0**<=**$\boldsymbol{r_{xy}}$**<=**(+1) – это прямая корреляционная связь; если же коэффициент регрессии $\boldsymbol{b}$<0, то (-1)**<=**$\boldsymbol{r_{xy}}$**<=**0 – это обратная корреляционная связь. При прямой (при обратной) связи увеличение одной из переменных ведет к увеличению (к уменьшению) условно средней другой.

3. Проверить значимость уравнения регрессии.

Проверить статистическую значимость уравнения регрессии – значит установить, соответствует ли математическая модель, выражающая зависимость между переменными, экспериментальным данным и достаточно ли включенных в уравнение объясняющих переменных (одной или нескольких) для описания зависимой переменной. Для парной регрессии число степеней свободы уравнения регрессии $\boldsymbol{k_1=m-1}$, а число степеней свободы остаточной дисперсии $\boldsymbol{k_2=n-m}$, где $\boldsymbol{m}$ – число оцениваемых параметров уравнения регрессии***,*** $\boldsymbol{n}$ – число наблюдений.

F-критерий Фишера будет равен:

$$F = \frac{r_{yx}^2}{1 - r_{yx}^2}(n - 2) \qquad (3)$$

Табличное значение ***F***-критерия Фишера выбирается с учетом рассчитываемых чисел степеней свободы, а уровень значимости считается равным 0,05.

Сравнение фактического значения ***F***-критерия Фишера с табличным позволяет сделать вывод, что уравнение регрессии статистически значимо или нет.

4. Рассчитать ошибку аппроксимации.

Ошибки аппроксимации рассчитываются для каждого наблюдения (Рисунок 1), а затем средняя ошибка аппроксимации находится как средняя арифметическая из индивидуальных ошибок:

$$\bar{A} = \frac{1}{n}\sum\left|\frac{y_i - y_p}{y_i}\right| \cdot 100, \qquad (4)$$

где y_i – исходное значение для *i*-й точки,

y_p – значение соответствующей регрессии для той же точки.

5. Получить параметры линейной регрессии с помощью функции ЛИНЕЙН(y,x,1,1)

Параметры линейного приближения по методу наименьших квадратов можно получить, используя стандартную функцию.

Для этого в ячейку вводят формулу *=ЛИНЕЙН(y,x,1,1),* указав диапазон.

Известные_значения_y – диапазон, содержащий числовые значения массива объясняемой (зависимой) переменной y .

Известные_значения_x – диапазон, содержащий числовые значения массива объясняющей (независимой) переменной x.

Константа – логическое значение, указывающее на наличие или отсутствие свободного члена в уравнении, при *Константе*=1 свободный член рассчитывается обычным способом, при *Константе*=0 свободный член равен 0.

Статистика – логическое значение, указывающее на возможность вывода дополнительной информации по регрессионному анализу. При *Статистика*=1 дополнительная информация выводится, при *Статистика*=0 выводятся только оценки параметров уравнения.

Выделить группу ячеек размером 5 строк и 2 столбца с ячейкой в верхнем левом углу, содержащей формулу *=ЛИНЕЙН(y,x,1,1),* затем сначала нажать на клавиатуре клавишу *F2*, потом – комбинацию клавиш *<CTRL>+<SHIFT>+<ENTER>* для раскрытия всей таблицы дополнительной информации по регрессионному анализу:

Таблица 1. Массив результирующих ячеек

Значение коэффициента b	Значение коэффициента a
Среднеквадратическое отклонение b	Среднеквадратическое отклонение a
Коэффициент детерминации r^2	Среднеквад. отклонение y
F-статистика	Число степеней свободы
Регрессионная сумма квадратов	Остат. сумма квадратов

6. Применить функции Регрессия и Описательная статистика для получения параметров линейной регрессии.

Параметры $\boldsymbol{a}$ и $\boldsymbol{b}$ линейной регрессии $\boldsymbol{y=a+b\cdot x}$ получаются с помощью функции *Регрессия* ППП EXCEL анализа данных.

Сводную таблицу основных статистических характеристик для одного или нескольких массивов данных можно получить с помощью инструмента ППП EXCEL анализа данных *Описательная статистика**:** Данные, Анализ данных, Описательная статистика*

7. Сравнить результаты трех способов получения параметров линейной регрессии.

Пример.

Применяем метод определителей для получения параметров линейной регрессии.

Заданы два массива ***x*** и ***y*** (**Рисунок 1*),*** где ***y*** – расходы на покупку продовольственных товаров, % к общему объему расходов; ***x*** – среднемесячная заработная плата одного работающего, тыс.руб.

Параметры ***a*** и ***b*** линейной регрессии ***y=a+bx*** рассчитываем в результате решения системы нормальных уравнений относительно ***a*** и ***b***.

	A	B	C	D	E	F	G	H	I
13	Номер региона	x	y	x*y	x^2	y^2	$y_{лин\ р}$=a+b*x	y-$y_{лин\ р}$	abc((y-$y_{лин\ р}$) /y)*100
14	1	4,5	68,8	309,6	20,25	4733,44	67,1	1,724	2,51
15	2	5,9	58,3	343,97	34,81	3398,89	59,3	-1,049	1,80
16	3	5,7	62,6	356,82	32,49	3918,76	60,5	2,147	3,43
17	4	7,2	52,1	375,12	51,84	2714,41	52,2	-0,073	0,14
18	5	6,2	54,5	337,9	38,44	2970,25	57,7	-3,193	5,86
19	6	6	57,1	342,6	36	3260,41	58,8	-1,697	2,97
20	7	7,8	51	397,8	60,84	2601	48,9	2,139	4,19
21	Сумма	43,3	404,4	2463,81	274,67	23597,16	404,40033	-0,00033	20,90
22	Среднее значение	6,186	57,771	351,973	39,239	3371,023	57,771	0,000	2,985
23	Количество регионов								7

Рисунок 1. Исходные данные и промежуточные расчеты для линейной регрессии

В формулу (1) подставив результаты промежуточных расчетов (Рисунок 1), получаем следующую систему нормальных уравнений относительно ***a*** и ***b***:

$$\begin{cases} 7a + 43,3b = 404,4 \\ 43,3a + 274,67b = 2463,81 \end{cases}.$$

Решаем систему методом определителей. Получаем уравнение линейной регрессии :

$$y_{лин\ р}=91,916-5,5199\cdot x.$$

Экономический смысл коэффициента b в том, что его величина показывает среднее изменение y-результата с изменением x-фактора на одну единицу.

Вывод. Величина коэффициента b= – 5,52 означает, что с ростом заработной платы на 1 тыс. руб. доля расходов на покупку продовольственных товаров снижается в среднем на 5,52%.

Линейное уравнение регрессии дополняем расчетом линейного коэффициента корреляции по формуле (2), где σ_x=1,0668, σ_y=6.2502.

Выводы.

1. Значение r_{xy}= – 0,94215 , т.е. близко к (-1) и существует сильная корреляция ***y*** и ***x***, или иначе – зависимость ***y*** и ***x*** близка к линейной.
2. Коэффициент детерминации составит r_{yx}^2 =0.88765. Вариации ***y*** на 88,8% объясняются вариацией ***x***. На долю прочих факторов, не учитываемых в регрессии, приходится 11,2%. Коэффициент детерминации 0<=r_{xy}^2 <=(+1), чем ближе к 1, тем лучше регрессия аппроксимирует эмпирические данные.

3. F-критерий Фишера (3) будет равен F=39.5.

   Для парной регрессии число степеней свободы уравнения регрессии k_1***=m-1=1***, а число степеней свободы остаточной дисперсии k_2***=n-m=5***, где ***m*** – число оцениваемых параметров уравнения регрессии ***(m=2), n*** – число наблюдений (***n=7***).

Табличное значение ***F***-критерия Фишера при числе степеней свободы 1 и 5 и уровне значимости 0,05 составит 6,61.

Вывод. Фактическое значение ***F***-критерия Фишера превышает табличное, и можно сделать вывод, что уравнение линейной регрессии статистически значимо.

Рассчитываем ошибку аппроксимации (Рисунок 1).

Вывод. Величина средней ошибки аппроксимации 2,99% показывает хорошее соответствие.

Параметры линейной регрессии ***y=a+b·x*** были получены (Таблица 2) в результате решения системы нормальных уравнений относительно ***a*** и ***b*** методом определителей:

Таблица 2. Линейная регрессии методом определителей

Значение коэффициента *b*	-5,51987	Значение коэффициента *a*	91,91579
Коэффициент детерминации r^2	0,887649	Среднеквадратическое отклонение *y*	
F-статистика	39,50336	Число степеней свободы	5

Получаем параметры линейной регрессии с помощью функции ЛИНЕЙН(y,x,1,1) (
Таблица 3***).***

Таблица 3. Параметры функции ЛИНЕЙН(y,x,1,1)

Значение коэффициента *b*	-5,51987	Значение коэффициента *a*	91,91579
Среднеквадратическое отклонение *b*	0.878238	Среднеквадратическое отклонение *a*	5.501343
Коэффициент детерминации r^2	0,887649	Среднеквадратическое отклонение *y*	2.29497
F-статистика	39,50336	Число степен.свободы	5
Регрессионная сумма квадратов	208.0598	Остаточная сумма квадратов	26.33445

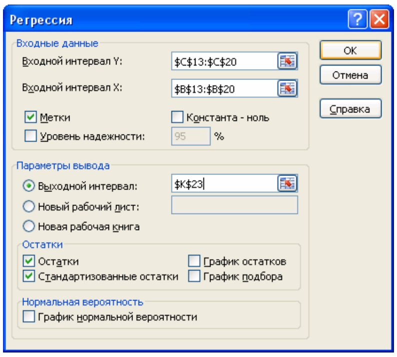
Выводы.

1. Вариации ***y*** на 88,8% объясняется вариацией ***x,*** т.к. коэффициент детерминации r^2=0,888.
2. Фактическое значение ***F***-критерия Фишера превышает табличное, и можно сделать вывод, что уравнение линейной регрессии статистически значимо.

Применяем функцию Регрессия для получения параметров линейной аппроксимации.

Для этого выполняем следующие шаги:

- вводим исходные данные,
- выполняем команду меню *Данные – Анализ данных – Регрессия,*
- при заполнении данных (Рисунок 2) диапазоны берем из исходных данных (Рисунок 1).



[redacted]

Рисунок 2. Данные для заполнения.

[redacted]

Рисунок 3. Параметры дисперсионного анализа.

Получаем уравнение линейной регрессии $\boldsymbol{y_{лин\ р}=b\cdot x+a}$, где $\boldsymbol{a}$ и $\boldsymbol{b}$ взяты из столбца *Коэффициенты* (

Рисунок **3**):

$$\boldsymbol{y_{лин\ р}=91{,}9158 - 5{,}51987\cdot x.}$$

Вывод. Величина коэффициента $b = -5{,}52$ означает, что с ростом заработной платы на 1 тыс. руб. доля расходов на покупку продовольственных товаров снижается в среднем на 5,52%.

Сводную таблицу основных статистических характеристик для двух массивов данных получаем с помощью инструмента *Описательная статистика****:*** *Данные, Анализ данных, Описательная статистика* (Рисунок 4, Рисунок 5).

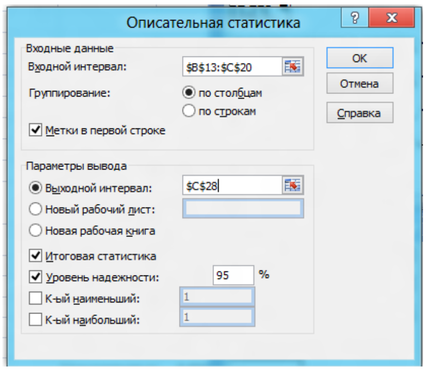


Рисунок 4. Данные для описательной статистики.

x		y	
Среднее	6,18571429	Среднее	57,7714286
Стандартная ошибка	0,40321835	Стандартная ошибка	2,36237515
Медиана	6	Медиана	57,1
Мода	#Н/Д	Мода	#Н/Д
Стандартное отклонение	1,06681547	Стандартное отклонение	6,25025714
Дисперсия выборки	1,13809524	Дисперсия выборки	39,0657143
Минимум	4,5	Минимум	51
Максимум	7,8	Максимум	68,8
Сумма	43,3	Сумма	404,4
Счет	7	Счет	7
Уровень надежности(95,0%)	0,987	Уровень надежности(95,0%)	5,78

Рисунок 5. Результаты описательной статистики.

Линейное уравнение регрессии дополняем линейным коэффициентом корреляции r_{yx}= -0,94215 (Рисунок 2), при этом σ_x=1,067 – стандартное отклонение по *x*, а σ_y=6,25 – стандартное отклонение по *y*, взятые из результатов *Описательной статистики* (Рисунок 5), а коэффициент *b* – из таблиц *Регрессии* (

Рисунок *3*).

Выводы

1. Значение r_{yx}= -0,94215 (Рисунок 2), т.е. близок к (-1) и существует сильная корреляция *y* и *x*, или другими словами – зависимость *y* и *x* близка к линейной.
2. Коэффициент регрессии b=-5,51987, т.е. b<0 и (-1)<= r_{yx} <=0 – это *обратная* корреляционная связь. При обратной связи увеличение одной из переменных ведет к уменьшению в 5,5 раз условно средней другой.
3. Коэффициент детерминации (Рисунок 2) r^2_{yx}= 0,887649. Вариации ***y*** на 88,8% объясняются вариацией ***x***. На долю прочих факторов, не учитываемых в регрессии, приходится 11,2%. Коэффициент детерминации 0<=$\boldsymbol{r_{yx}^2}$ <=(+1) и чем ближе к 1, тем лучше математическая модель в виде линейной регрессии аппроксимирует эмпирические данные.
   F-критерий Фишера равен 39,5, причем значение взято из таблиц *Регрессии* (
   Рисунок *3*).

Табличное значение F-критерия Фишера при числе степеней свободы $k1=m-1=1$ и $k2=n-m=5$, где m– число оцениваемых параметров уравнения регрессии (m=2), n – число наблюдений (n=7), и уровне значимости 0,05 составит $\boldsymbol{F_{табл}}$=6,61.

Вывод. Фактическое значение F-критерия Фишера $F_{факт}$=39,5 превышает табличное $F_{табл}$=6,61, и можно сделать вывод, что полученное с помощью функции *Регрессия* линейной уравнение $\boldsymbol{y=a+b\cdot x}$ статистически значимо.

Таблица 4. Результаты функции *Регрессия*

Значение коэффициента b	-5,51987	Значение коэффициента a	91,91579
Коэффициент детерминации r^2	0,887649	Среднеквадратическое отклонение y	
F-статистика	39,503384	Число степеней свободы	5

Вывод. Для получения параметров *a* и *b линейной* регрессии были применены три способа:

- метод определителей (Таблица 2),
- с помощью стандартной функции *ЛИНЕЙН*(***y,x***,1,1) (
- Таблица **3**),
- с помощью функции *Регрессия (*Таблица 4).

Сравнение результатов вполне удовлетворительное.

Следовательно, любой из названных методов можно использовать для получения ***уравнения регрессии.***

Лабораторная работа №2.
ПАРНАЯ ПОКАЗАТЕЛЬНАЯ РЕГРЕССИЯ

Пусть имеется набор значений двух переменных $\boldsymbol{y_i}$ (эндогенная или результат) и $\boldsymbol{x_i}$ (экзогенная или фактор).

Цель работы.

Подобрать по данным наблюдений $\boldsymbol{(x_i,y_i)}$, $\boldsymbol{i=1,\dots,n}$ функцию показательной регрессии $\boldsymbol{y_{пок\,р}(x)=ab^x}$, наилучшим образом описывающую заданную зависимость.

Порядок проведения работы.

1. Применить метод определителей для получения показательной регрессии.

Для оценки параметров модели в виде показательной функции $\boldsymbol{y=ab^x}$ линеаризуем модель путем логарифмирования $\boldsymbol{\ln y = \ln a + x*\ln b}$. Обозначим $\boldsymbol{\ln a = A}$, $\boldsymbol{\ln b = B}$. Применяем метод наименьших квадратов и получаем систему уравнений относительно $\boldsymbol{\ln a}$ и $\boldsymbol{\ln b}$:

$$\begin{cases} n\ln a + \ln b\sum x = \sum \ln y \\ \ln a\sum x + \ln b\sum x^2 = \sum x\ln y \end{cases}. \qquad (5)$$

2. Получить индекс корреляции.

Показательное уравнение регрессии дополняется расчетом индекса корреляции:

$$R = \sqrt{1 - \frac{\sum (y - y_{\text{стр}})^2}{\sum (y - \bar{y})^2}}$$

$$\sum (y - \bar{y})^2 = n * \sigma_y^2 = \mathbf{234{,}4}$$

R= 0,965

$\sigma_x =$ 0,987679

$\sigma_y =$ 5,786614

с листа линейной регрессии

(6)

3. Проверить значимость уравнения регрессии.

Критерий Фишера определяется по формуле (3).

4. Рассчитать ошибку аппроксимации.

Средняя ошибка аппроксимации находится как средняя арифметическая из индивидуальных ошибок по формуле (4).

5. Получить параметры регрессии с помощью функции ЛГРФПРИБЛ(y,x,1,1)

Параметры приближения в виде показательной функции по методу наименьших квадратов можно получить, используя стандартную функцию ***ЛГРФПРИБЛ(y,x,1,1).***

Для этого в ячейку вводят формулу ***=ЛГРФПРИБЛ(y,x,1,1),*** указав диапазоны.

Известные_значения_у – диапазон, содержащий числовые значения массива объясняемой (зависимой) переменной *y* ,

Известные_значения_х – диапазон, содержащий числовые значения массива объясняющей (независимой) переменной *x*,

Константа – логическое значение, указывающее на наличие или отсутствие свободного члена в уравнении, при *Константе*=1 свободный член рассчитывается обычным способом, при *Константе*=0 свободный член равен 0.

Статистика – логическое значение, указывающее на возможность вывода дополнительной информации по регрессионному анализу. При *Статистика*=1 дополнительная информация выводится, при *Статистика*=0 выводятся только оценки параметров уравнения.

Выделить группу ячеек размером 5 строк и 2 столбца с ячейкой в верхнем левом углу, содержащей формулу ***=ЛГРФПРИБЛ(y,x,1,1),*** затем сначала нажать на клавиатуре клавишу *F2*, потом – комбинацию клавиш *<CTRL>+<SHIFT>+<ENTER>* для раскрытия всей таблицы дополнительной информации по регрессионному анализу. В массиве ячеек в Таблица *3* показано, как будут размещаться параметры.

6. Применить метод решения с помощью функции Регрессия.

Параметры a и b показательной регрессии ***$y=ab^x$*** рассчитываются с помощью функции *Регрессия* ППП EXCEL анализа данных. Но в ***качестве входного интер-***

вала у используется диапазон ln(y), а столбец x является входным интервалом x, причем метка (заголовок) столбца выделяется, вместе со значениями x или y.

7. Сравнить результаты трех способов получения параметров показательной регрессии.

Пример.

Применяем метод определителей для получения параметров показательной регрессии

Заданы два массива ***x*** и ***y*** (Рисунок 6):

Номер региона	x	y	x^2	ln(y)	x*ln(y)	$y_{пок\ р}$=a*b^x	$(y\text{-}y^{сред\ з})^2$	$(y\text{-}y_{пок\ р})^2$	abc((y-$y_{пок\ р}$)/y)*100
1	4,5	68,8	20,25	4,2312	19,0404	67,345	121,629	2,12	2,12
2	5,9	58,3	34,81	4,0656	23,9871	59,053	0,279	0,57	1,29
3	5,7	62,6	32,49	4,1368	23,5796	60,172	23,315	5,90	3,88
4	7,2	52,1	51,84	3,9532	28,4628	52,270	32,165	0,03	0,33
5	6,2	54,5	38,44	3,9982	24,7888	57,413	10,702	8,49	5,35
6	6	57,1	36	4,0448	24,2688	58,501	0,451	1,96	2,45
7	7,8	51	60,84	3,9318	30,6682	49,408	45,852	2,53	3,12
Сумма	43,3	404,4	274,7	28,362	174,8	404,2	234,39429	21,59551	18,5
Среднее значение	6,186	57,771	39,239	4,052	24,971	57,737	33,485	3,085	2,648

Рисунок 6. Исходные данные и расчеты.

y – расходы на покупку продовольственных товаров, % к общему объему расходов;

x – среднемесячная заработная плата одного работающего, тыс.руб.

Система нормальных уравнений имеет следующие значения коэффициентов при ***A=lna*** и ***B=lnb.*** Решаем систему методом определителей.

$$\begin{cases} 7\ln a + 43{,}3\ln b = 28{,}362 \\ 43{,}3\ln a + 274{,}77\ln b = 174{,}8 \end{cases}.$$

Параметры ***A*** и ***B*** рассчитываем в результате решения системы нормальных уравнений относительно **ln*a*** и **ln*b***, и получаем ***A***=4,63214 и ***B***=-0,09384.

Выполнив потенцирование, получим:

$$a = e^{4,63214} = 102{,}734003;\ b = e^{-0,09384} = 0{,}91042;$$

таким образом, имеем показательное уравнение регрессии:

$$y_{пок\ р} = 102{,}734003 \cdot 0{,}91042^x.$$

Показательное уравнение регрессии дополняется расчетом по формуле (6) индекса корреляции R=0.94996:

Коэффициент детерминации составит R^2=0.90242.

Выводы

1. Вариация ***y*** на 90,8% объясняется вариацией ***x***. На долю прочих факторов, не учитываемых в регрессии, приходится 9,2%.
2. Критерий Фишера по формуле (3) будет равен 46,24. Фактическое значение ***F***-критерия Фишера ***F***=46.24 превышает табличное 6,61, и можно сделать вывод, что <u>показательное уравнение</u> <u>статистически значимо</u>.
3. Величина средней ошибки аппроксимации (Рисунок 6) 2,65% показывает хорошее соответствие.

Параметры показательной регрессии $\boldsymbol{y=ab^x}$ были получены в результате решения системы нормальных уравнений методом определителей (Таблица 5):

Таблица 5. Показательная регрессии методом определителей

Значение коэффициента *b*	0,9104	102,734	Значение коэффициента *a*
Коэффициент детерминации r^2	0,902		Среднеквадратическое отклонение *y*
F-статистика	46,24	5	Число степеней свободы

По методу решения с помощью стандартной функции ЛГРФПРИБЛ(y,x,1,1) получаем параметры логарифмического приближения (Таблица 6).

Таблица 6. Показательная регрессии методом логарифмического приближения

Значение коэффициента b	0,910425	Значение коэффициента a	102,734
Среднеквадратическое отклонение b	0,013801	Среднеквад. отклонение a	0,086448
Коэффициент детерминации r2	0,902419	Среднеквад. отклонение y	0,036063
F-статистика	46,2395	Число степеней свободы	5
Регрессионная сумма квадратов	0,060137	Остаточная сумма квадратов	0,006503

Выводы.

1. Вариации ***y*** на 90,2% объясняется вариацией ***x***, т.к. коэффициент детерминации r^2=0,902.
2. Фактическое значение ***F***-критерия Фишера превышает табличное, и можно сделать вывод, что уравнение показательной регрессии статистически значимо.

Применяем функция Регрессия для получения параметров показательной аппроксимации.

Для этого выполняем следующие шаги:

- вводим исходные данные;
- выполняем команду *Меню, Данные, Анализ данных, Регрессия.*

В качестве входного интервала y используем столбец ln(y), а столбец x является входным интервалом x, причем метку (заголовок) столбца выделяем вместе со значениями x или y (Рисунок 7)**.**

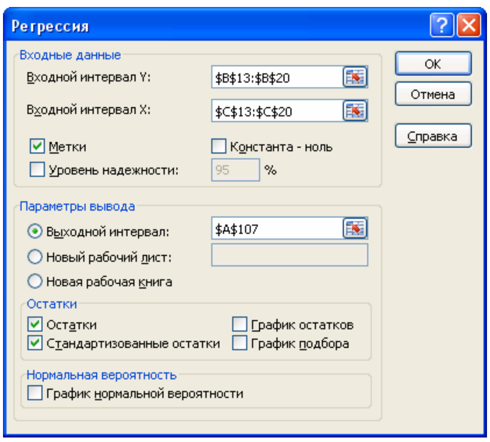


Рисунок 7. Окно *Регрессия* для ввода данных

[redacted]

Рисунок 8. Параметры дисперсионного анализа

Получаем уравнение:

$$Ln\ y_{пок\ р}=4{,}63214 - 0{,}09384 \cdot x,$$

причем ***a*** и ***b*** взяты (
Рисунок **8**) из столбца *Коэффициенты*.

Выполнив потенцирование, получим уравнение показательной регрессии $\boldsymbol{y_{пок\ р}}$:

$$a = e^{4,63214}; \qquad b=e^{-0,09384},$$

$$y_{покр} = 102{,}73403 \cdot 0{,}91042^{x}.$$

Показательное уравнение регрессии дополняем индексом корреляции: ***R***=0,94996 (Рисунок 7).

Коэффициент детерминации (Рисунок **7**) составит $\boldsymbol{R^2}$=0.90242.

Вывод. Вариации ***y*** на 90,2% объясняется вариацией ***x***. На долю прочих факторов, не учитываемых в регрессии, приходится 9,8%.

F-критерий Фишера будет (3) равен 46,2.

Параметры ***a*** и ***b*** показательной регрессии $\boldsymbol{y=ab^x}$ были получены с помощью функции *Регрессия* (Таблица 7).

Таблица 7. Показательная регрессия через функцию *Регрессия*

Значение коэффициента	0,9104	Значение коэффициента	102,7340
Коэффициент детерминации r^2	0,90242	Среднеквадр. отклонение y	
F-статистика	46,239	Число степ. свободы	5

Вывод. Фактическое значение ***F***-критерия Фишера 46,2 превышает табличное $\boldsymbol{F}_{табл}$=6,61, и можно сделать вывод, что уравнение <u>показательной регрессии статистически значимо</u>.

Вывод. Для получения параметров *a* и *b* ***показательной*** регрессии были применены три способа:

- метод определителей (Таблица 5),
- с помощью стандартной функции ЛГРФПРИБЛ(y,x,1,1) (Таблица 6),
- с помощью функции *Регрессия (*Таблица 7).

Сравнение результатов вполне удовлетворительное.

Следовательно, любой из названных методов можно использовать для получения уравнения показательной регрессии.

ЛАБОРАТОРНАЯ РАБОТА №3.
ПАРНАЯ СТЕПЕННАЯ И ПАРНАЯ ЭКСПОНЕНЦИАЛЬНАЯ РЕГРЕССИИ

Для получения моделей в виде парной степенной функции и в виде парной экспоненциальной функции ниже приведены расчеты с применением метода определителей, а функцию *Регрессия* следует применить для проверки.

Пусть имеется набор значений двух переменных y_i и x_i.

Цель работы.

Подобрать по данным наблюдений (x_i,y_i), $i=1,\dots,n$ функцию степенной регрессии $y_{степ\,р}(x)=ax^b$ и в виде экспоненциальной функции $y_{экс\,р}=ae^{bx}$, наилучшим образом описывающие предполагаемую зависимость.

Порядок проведения работы.

1. Применить метод определителей для получения параметров степенной регрессии.

Для оценки параметров модели в виде степенной функции $y=ax^b$ линеаризуем модель путем логарифмирования $\ln y = \ln a + b\cdot\ln x$. Обозначим $\ln a = A$, $\ln x = X$, $\ln y = Y$. Тогда получим $Y=A+b\cdot X$. Применяем метод наименьших квадратов и получаем систему уравнений.

$$\begin{cases} nA + b\sum X = \sum Y \\ A\sum X + b\sum X^2 = \sum YX \end{cases} \qquad (7)$$

1. Применить метод определителей для получения параметров экспоненциальной регрессии.

Для оценки параметров модели в виде экспоненциальной функции $y=ae^{bx}$ линеаризуем модель путем логарифмирования $\ln y = \ln a + b*x$. Обозначим $\ln a = A$. Применяем метод наименьших квадратов и получаем систему уравнений.

$$\begin{cases} n\ln a + b\sum x = \sum \ln y \\ \ln a\sum x + b\sum x^2 = \sum x\ln y \end{cases} \qquad (8)$$

2. Получить индекс корреляции.

Уравнение регрессии дополнить расчетом индекса корреляции по формуле (6) для степенной и экспоненциальной функций:

3. Проверить значимость уравнения регрессии.

Критерий Фишера определяется по формуле (3).

4. Рассчитать ошибку аппроксимации.

Средняя ошибка аппроксимации находится как средняя арифметическая из индивидуальных ошибок по формуле (4).

5. Функцию Регрессия следует применить для проверки полученных ранее параметров степенной экспоненциальной регрессий.

Пример 1.

Применяем метод определителей для получения параметров степенной регрессии $y=ax^b$.

Заданы два массива $\boldsymbol{x}$ и $\boldsymbol{y}$ (

Рисунок **9**), где $\boldsymbol{y}$ – расходы на покупку продовольственных товаров, % к общему объему расходов; $\boldsymbol{x}$ – среднемесячная заработная плата одного работающего, тыс.руб.

Номер региона	x	y	X=lnx	Y=lny	Y*X	X^2	Y^2	$y_{степ\ р}=e^a*x^b$	$(y-y_{степ\ р})^2$	abc((y-$y_{степ\ р}$)/y)*100
1	4,5	68,8	1,50	4,23	6,3641	2,262	17,903	68,521	0,08	0,406
2	5,9	58,3	1,77	4,07	7,2162	3,150	16,529	58,635	0,11	0,574
3	5,7	62,6	1,74	4,14	7,1999	3,029	17,113	59,809	7,79	4,458
4	7,2	52,1	1,97	3,95	7,8039	3,897	15,628	52,289	0,04	0,362
5	6,2	54,5	1,82	4,00	7,2949	3,329	15,986	56,985	6,18	4,560
6	6,0	57,1	1,79	4,04	7,2473	3,210	16,360	58,070	0,94	1,699
7	7,8	51,0	2,05	3,93	8,0765	4,219	15,459	49,936	1,13	2,087
Сумма	43,3	404,4	12,664	28,362	51,203	23,098	114,978	404,245	16,265	14,146
Среднее значение	6,186	57,771	1,809	4,052	7,315	3,300	16,425	57,749	2,324	2,021
Количество регионов										7

Рисунок 9. Исходные данные для степенной регрессии

Система нормальных уравнений (7) имеет следующие значения коэффициентов при ***A*** и ***b.*** Решаем систему методом определителей.

$$\begin{cases} 7a + 12{,}664b = 28{,}362 \\ 12{,}664\ A + 23{,}098\ b = 51{,}203 \end{cases}$$

Получаем уравнение**:**

$$\boldsymbol{Ln\ y_{степ\ р}} = 5{,}09233 - 0{,}57523 \cdot \boldsymbol{ln\ x}.$$

Выполнив потенцирование, получим:

$$\boldsymbol{y_{степ\ р}} = \boldsymbol{e}^{5{,}09233}\boldsymbol{x}^{-0.57523} \text{ или } \boldsymbol{y_{степ\ р}} = 162{,}76791 \cdot \boldsymbol{x}^{-0{,}57523}.$$

Вывод. Параметр ***b*** =-0.57523 означает коэффициент эластичности, который показывает, что с ростом зарплаты на 1% доля расходов на продовольствие снижается в среднем на 0,58%.

Степенное уравнение регрессии дополняем расчетом индекса корреляции:

$$R = \sqrt{1 - \frac{\sum (y - y_{стр})^2}{\sum (y - \bar{y})^2}}$$

$$\sum (y - \bar{y})^2 = n * \sigma_y^2 = \mathbf{234{,}4}$$

$$\boxed{\mathbf{R= \quad 0{,}965}}$$

$\boxed{\sigma_x =}$ 0,987679

$\boxed{\sigma_y =}$ 5,786614

с листа линейной регрессии

(6)

Вывод. Индекс корреляции ***R***=0,965, что свидетель-ствует о тесноте связи рассматриваемых признаков.

Расчет коэффициента корреляции между ***ln y*** и ***ln x***

не показывает тесноту связи:

$$r_{Lny\ Lnx} = \frac{\overline{LnxLny} - \overline{Lnx} * \overline{Lny}}{\sigma_{Lnx} * \sigma_{Lny}}$$, т.е. $r_{lny\ lnx}$=-7,9·10^{-06}.

$$\sigma^2_{Lny} = \frac{\sum Lny^2}{n} - \overline{Lny}^2$$ =98,5.

$$\sigma^2_{Lnx} = \frac{\sum Lnx^2}{n} - \overline{Lnx}^2$$ =19,6.

Коэффициент детерминации составит R^2=0,930609.

Выводы

1. Вариация ***y*** на 93,1% объясняется вариацией ***x***. На долю прочих факторов, не учитываемых в регрессии, приходится 6,9%.
2. Фактическое значение ***F***-критерия Фишера (3) ***F***=67,1 превышает табличное 6,61, и можно сделать вывод, что степенное уравнение статистически значимо.
3. Величина средней ошибки аппроксимации 2,02% показывает хорошее соответствие (
4. Рисунок *9*).

Параметры степенной регрессии $y=ax^b$ были получены (

Таблица *8*) в результате решения системы нормальных уравнений методом определителей:

Таблица 8.Степенная регрессия методом определителей

Значение коэффициента *b*	-0,57523	Значение коэффициента *a*	5,092
Коэффициент детерминации r^2	0,930609	Среднеквадратическое отклонение *y*	
F-статистика	67,1	Число степен.свободы	5

Пример 2.

Применяем метод определителей для получения параметров парной экспоненциальной регрессии $y=ae^{bx}$.

Заданы два массива ***x*** и ***y*** (Рисунок 10):

Номер региона	x	y	x^2	ln(y)	x ln(y)	$y_{экс\ р}=ae^{bx}$	$(y-y_{экс\ р})^2$	abc((y-$y_{экс\ р}$) /y)*100
1	4,5	68,8	20,25	4,2312	19,040	67,347	2,11	2,111
2	5,9	58,3	34,81	4,0656	23,987	59,056	0,57	1,297
3	5,7	62,6	32,49	4,1368	23,580	60,175	5,88	3,874
4	7,2	52,1	51,84	3,9532	28,463	52,274	0,03	0,334
5	6,2	54,5	38,44	3,9982	24,789	57,417	8,51	5,352
6	6,0	57,1	36,00	4,0448	24,269	58,504	1,97	2,460
7	7,8	51,0	60,84	3,9318	30,668	49,412	2,52	3,114
Сумма	43,3	404,4	274,67	28,362	174,80	404,185	21,595	18,541
Среднее значение	6,186	57,771	39,239	4,052	24,971	57,741	3,085	2,649

Рисунок 10 Исходные данные для парной экспоненциальной регрессии

Система нормальных уравнений (8) имеет следующие значения коэффициентов при ***lna*** и ***b.*** Решаем систему методом определителей.

$$\begin{cases} 7\ ln\ a + 43{,}3\ b = 28{,}362 \\ 43{,}3\ ln\ a + 274{,}77\ b = 174{,}8 \end{cases}.$$

Коэфициенты ***ln a*** и ***b*** подставляем в уравнение:

$$\boldsymbol{ln\ y_{экс\ р} = ln\ a + b \cdot x} = 4{,}63214 - 0{,}09384 \cdot x.$$

Выполнив потенцирование, получим:

$$\boldsymbol{y_{экс\ р}} = \boldsymbol{e}^{4{,}63214} \cdot \boldsymbol{x}^{-0{,}09384} \quad \text{или} \quad \boldsymbol{y_{экс\ р}} = 102{,}73403 \cdot \boldsymbol{e}^{-0{,}09384 * \boldsymbol{x}}.$$

Экспоненциальное уравнение регрессии дополняем расчетом индекса корреляции:

$$R = \sqrt{1 - \frac{\sum (y - y_{экс\ р})^2}{\sum (y - \bar{y})^2}} = 0{,}9628211,$$

где σ_x=1,0668 σ_y=6.2502 $\sum (y-\bar{y})^2 = n * \sigma_y^2 =$ 234,394.

Выводы.

1. Индекс корреляции ***R***=0,96, что свидетельствует о тесноте связи рассматриваемых признаков.
2. Коэффициент детерминации составит $\boldsymbol{R^2}$=0,907868. Это говорит о том, 90,79% вариации ***y*** связано с вариацией ***x***.
3. Фактическое значение (3) ***F***-критерия Фишера ***F***=49,3 превышает табличное 6,61, и можно сделать вывод, что экспоненциальное уравнение статистически значимо.
4. Величина средней ошибки аппроксимации 2,7% показывает хорошее соответствие.

Параметры экспоненциальной регрессии $\boldsymbol{y_{экс\ р}} = 102{,}73403 \cdot \boldsymbol{e}^{-0{,}09384 * \boldsymbol{x}}$ были получены в результате решения системы нормальных уравнений методом определителей:

Таблица 9. Экспоненциальная регрессия методом определителей

Значение коэффициента *b*	-0,0938	102,734	Значение коэффициента *a*
Коэффициент детерминации r^2	0,90787		Среднеквад. отклонение *y*
F-статистика	49,3	5	Число степеней свободы

ЛАБОРАТОРНАЯ РАБОТА №4.
ПАРНАЯ РЕГРЕССИЯ В ВИДЕ ГИПЕРБОЛЫ И ОБРАТНОЙ ФУНКЦИИ

Для получения моделей в виде равносторонней гиперболы и в виде обратной функции ниже приведены примеры с применением метода определителей, а функцию *Регрессия* следует применить для проверки.

Пусть имеется набор значений двух переменных $\boldsymbol{y_i}$ и $\boldsymbol{x_i}$.

Цель работы. Подобрать по данным наблюдений (x_i,y_i), $i=1,\dots,n$ функцию регрессии в виде равносторонней гиперболы $y_{степ\ p}(x)=ax^b$ и в виде обратной функции $y_{экс\ p}=ae^{bx}$, наилучшим образом описывающие предполагаемую зависимость.

Порядок проведения работы.

1. Применить метод определителей для получения параметров регрессии в виде равносторонней гиперболы.

Для оценки параметров модели в виде равносторонней гиперболы $y=a+b/x$ линеаризуем модель путем замены $z=1/x$ тогда $y = a + b*z$. Применяем метод наименьших квадратов и получаем систему уравнений.

$$\begin{cases} n\,a + b\sum z = \sum y \\ a\sum z + b\sum z^2 = \sum yz \end{cases} \qquad (9)$$

2. Применить метод определителей для получения параметров регрессии в виде обратной функции.

Для оценки параметров регрессии в виде обратной функции $y=1/(a+b*x)$ линеаризуем модель путем замены $Y=1/y$, тогда $Y = a + b*x$. Получаем систему уравнений.

$$\begin{cases} n\,a + b\sum x = \sum Y \\ a\sum x + b\sum x^2 = \sum Yx \end{cases}. \qquad (10)$$

3. Получить индекс корреляции.

Уравнения регрессии в виде равносторонней гиперболы и в виде обратной функции дополнить расчетом индекса корреляции по формуле (6):

4. Проверить значимость уравнения регрессии.

Критерий Фишера определяется по формуле (3).

5. Рассчитать ошибку аппроксимации.

Средняя ошибка аппроксимации находится как средняя арифметическая из индивидуальных ошибок по формуле (4).

6. Функцию Регрессия следует применить для проверки полученных ранее параметров в виде равносторонней гиперболы и параметров в виде обратной функции.

Пример 1.

Для парной регрессии в виде равносторонней гиперболы применяем метод определителей.

Пусть имеется набор значений двух переменных y_i и x_i, где y_i – расходы на покупку продовольственных товаров, % к общему объему расходов; x_i – среднемесячная заработная плата одного работающего, тыс. руб. Расчеты сведены в таблицу:

Таблица 10. Исходные данные для равносторонней гиперболы

Номер региона	x	y	z=1/x	y*z	z^2	ln(y)	x ln(y)	$y_{гип\ р}$=a+b/x	$(y-y_{гип\ р})^2$	abc((y-$y_{гип\ р}$)/y)*100
1	4,5	68,8	0,222	15,29	0,049	4,2312	19,040	69,035	0,06	0,342
2	5,9	58,3	0,169	9,88	0,029	4,0656	23,987	58,465	0,03	0,284
3	5,7	62,6	0,175	10,98	0,031	4,1368	23,580	59,658	8,66	4,700
4	7,2	52,1	0,139	7,24	0,019	3,9532	28,463	52,331	0,05	0,444
5	6,2	54,5	0,161	8,79	0,026	3,9982	24,789	56,822	5,39	4,260
6	6,0	57,1	0,167	9,52	0,028	4,0448	24,269	57,899	0,64	1,400
7	7,8	51,0	0,128	6,54	0,016	3,9318	30,668	50,190	0,66	1,589
Сумма	43,3	404,4	1,162	68,234	0,198	28,362	174,80	404,400	15,479	13,018
Среднее значение	6,186	57,771	0,166	9,748	0,028	4,052	24,971	57,771	2,211	1,860

Система нормальных уравнений (9) имеет следующие значения коэффициентов при ***a*** и ***b*.** Решаем систему методом определителей.

$$\begin{cases} 7a + 1,1622\,b = 404,4 \\ 1,1622\,a + 0,1984\,b = 68,234 \end{cases}.$$

Получаем уравнение регрессии в виде равносторонней гиперболы

$$\boldsymbol{y}_{гип\ р} = 24,49081+200,45059/\boldsymbol{x}.$$

Определяем (6) индекс корреляции ***R***=0,96642**,** и по нему судим о тесноте связей.

Выводы.

1. Коэффициент детерминации составит R2=0,933961. Это говорит о том, 93,4% вариации y связано с вариацией x.
2. Фактическое значение (3) F-критерия Фишера F=70,7 превышает табличное 6,61, и можно сделать вывод, что уравнение регрессии в виде равносторонней гиперболы статистически значимо.
3. Величина средней ошибки аппроксимации 1,86% показывает хорошее соответствие.

Параметры регрессии в виде равносторонней гиперболы ***y = a+b/x*,** полученные в результате решения системы нормальных уравнений относительно ***a*** и ***b*** методом определителей, сведены в таблицу:

Таблица 11. Регрессия в виде равносторонней гиперболы

Значение коэффициента *b*	200,45	24,491	Значение коэффициента *a*
Коэффициент детерминации r^2	0,9339		Среднеквадратическое отклонение *y*
F-статистика	70,7	5	Число степеней свободы

Пример 2.

Данные расчетов ***методом определителей для парной регрессии в виде обратной функции приведены на*** (Рисунок 11).

Номер региона	x	y	Y=1/y	Y*x	x^2	$y_{обр\ р}$= 1/(a+b*x)	$(y-y_{обр\ р})^2$	abc((y-$y_{обрп\ р}$) /y)*100
1	4,5	68,8	0,01453	0,0654	20,25	67,719	1,17	1,572
2	5,9	58,3	0,01715	0,1012	34,81	58,770	0,22	0,807
3	5,7	62,6	0,01597	0,0911	32,49	59,901	7,28	4,311
4	7,2	52,1	0,01919	0,1382	51,84	52,347	0,06	0,475
5	6,2	54,5	0,01835	0,1138	38,44	57,152	7,03	4,866
6	6,0	57,1	0,01751	0,1051	36,00	58,221	1,26	1,963
7	7,8	51,0	0,01961	0,1529	60,84	49,834	1,36	2,287
Сумма	43,3	404,4	0,12233	0,7676	274,67	403,943	18,386	16,281
Среднее значение	6,186	57,77	0,017	0,110	39,239	57,706	2,627	2,326

Рисунок 11. Исходные данные для регрессии в виде обратной функции

Система нормальных уравнений (10) имеет следующие значения коэффициентов при ***a*** и ***b*.** Решаем систему.

$$\begin{cases} 7a + 43{,}3\,b = 0{,}12233 \\ 43{,}3\,a + 274{,}67\,b = 0.7676 \end{cases}.$$

Получаем уравнение ***Y = a + b*x*,** тогда уравнение регрессии в виде обратной функции:

$$y_{обр\ р} = 1/(0{,}00754+0{,}001606\,x).$$

Определяем (6) индекс корреляции ***R***=0,959981**,** и по нему судим о тесноте связей.

Выводы.

1. Коэффициент детерминации составит R^2=0,921563. Это говорит о том, 92,2% вариации ***y*** связано с вариацией ***x***.
2. Фактическое значение (3) ***F***-критерия Фишера ***F***=58,7 превышает табличное 6,61, и можно сделать вывод, что уравнение регрессии в виде обратной функции статистически значимо.
3. Величина средней ошибки аппроксимации 2,32% показывает хорошее соответствие.

Параметры регрессии в виде обратной функции ***y = 1/(a+b·x)*,** полученные в результате решения системы нормальных уравнений относительно ***a*** и ***b*** методом определителей, приводятся в таблице:

Таблица 12. Регрессия в виде обратной функции

Значение коэффициента *b*	0,00161	0,00754	Значение коэффициента *a*
Коэффициент детерминации r^2	0,92156		Среднеквадратическое отклонение *y*
F-статистика	58,7	5	Число степеней свободы

ВЫБОР НАИЛУЧШЕЙ МОДЕЛИ

В лабораторных работах 1-4 найдены следующие модели и выбрана лучшая.

№	Название регрессии	Уравнение регрессии	Коэффициент детерминации	F-критерий Фишера	Статистическая значимость уравнения	Средняя ошибка аппроксимации в %
1	Линейная	$y=a+bx$	0,8876	39,50	значимо	2,99
2	Показательная	$y=ab^x$	0.9024	46.24	значимо	2,65
3	Степенная	$y=ax^b$	0,9306	67,1	значимо	2,02
4	Экспоненциальная	$y=ae^{bx}$	0,9078	49,3	значимо	2,7
5	Равносторонняя гипербола	$y=a+b/x$	0,9339	70,7	значимо	1,86
6	Обратная функция	$y=1/(a+b\cdot x)$	0,9216	58,7	Значимо	2,32

Вывод. Регрессия в виде равносторонней гиперболы имеет максимальный коэффициент детерминации 0,9339 и при этом эта регрессия дает минимальную погрешность **1,86.**

ЛАБОРАТОРНАЯ РАБОТА №5.
ВЫЧИСЛЕНИЕ ПАРАМЕТРОВ ЛИНЕЙНОГО УРАВНЕНИЯ МНОЖЕСТВЕННОЙ РЕГРЕССИИ МЕТОДОМ СТАНДАРТИЗАЦИИ ПЕРЕМЕННЫХ

Множественная регрессия широко используется в решении проблем спроса, доходности акций, при изучении функции издержек производства, в макроэкономических расчётах и целого ряда других вопросов эконометрики. Основная цель множественной регрессии – построить модель с большим числом факторов, определив при этом влияние каждого из них в отдельности, а также совокупное воздействие их на моделируемый показатель. Специфика множественной регрессии заключается в исследовании комплексного воздействия факторов в условиях их независимости друг от друга.

Цель работы.

Поскольку одним из условий построения линейного уравнения множественной регрессии $\hat{y} = a + b_1 x_1 + b_2 x_2$ является независимость действия факторов, то сначала необходимо:

1. Оценить показатели вариации каждого признака и сделать вывод о возможностях применения метода наименьших квадратов (МНК) для их изучения.

Сводную таблицу основных статистических характеристик для нескольких массивов данных можно получить с помощью инструмента EXCEL анализа данных *Описательная статистика.*

2. Проанализировать линейные коэффициенты парной и частной корреляции.

Значения линейных коэффициентов парной корреляции определяют тесноту попарно связанных переменных, использованных в уравнении множественной регрессии. Линейные коэффициенты частной корреляции оценивают тесноту связи значений двух переменных, исключая влияние всех других переменных, представленных в уравнении множественной регрессии.

Матрицу парных коэффициентов корреляции переменных можно получить, используя инструмент анализа данных *Корреляция***.**

В ППП EXCEL нет специального инструмента для расчёта линейных коэффициентов частной корреляции.

Их можно рассчитать по рекуррентной формуле через коэффициенты парной корреляции

$$r_{yx1/x2} = \frac{r_{yx1} - r_{yx2} \cdot r_{x1x2}}{\sqrt{(1 - r^2_{yx2}) \cdot (1 - r^2_{x1x2})}}$$

$$r_{yx2/x1} = \frac{r_{yx2} - r_{yx1} \cdot r_{x1x2}}{\sqrt{(1 - r^2_{yx1}) \cdot (1 - r^2_{x1x2})}}$$

$$r_{x1x2/y} = \frac{r_{x2x1} - r_{yx2} \cdot r_{yx1}}{\sqrt{(1 - r^2_{yx2}) \cdot (1 - r^2_{yx1})}} \qquad (11)$$

где используется обозначение $r_{yx2/x1}$, т.е. ***x1*** фиксируем.

3. Составить методом стандартизации переменных уравнение множественной регрессии, оценить значимость его параметров, пояснить их экономический смысл.

Стандартизованные частные коэффициенты регрессии β-коэффициенты показывают, на какую часть своего среднеквадратического отклонения изменится признак-результат y с увеличением соответствующего фактора x_i на величину своего среднеквадратического отклонения при неизменном влиянии прочих факторов модели.

Для вычисления коэффициентов множественной регрессии применим метод стандартизации переменных.

Построим искомое уравнения в стандартизованном масштабе

$$t_y = \beta_1 * t_{x1} + \beta_2 * t_{x2}.$$

Расчёт β-коэффициентов выполняется по формулам:

$$\beta_1 = \frac{r_{yx1} - r_{yx2} \cdot r_{x1x2}}{1 - r^2_{x1x2}}$$

$$\beta_2 = \frac{r_{yx2} - r_{yx1} \cdot r_{x1x2}}{1 - r^2_{x1x2}}. \qquad (12)$$

Уравнение в стандартизованном масштабе

$$t_y = 0{,}9343 \ . t_{x1} - 0{,}0265 \ . t_{x2}.$$

Для построения уравнения в естественной форме рассчитаем b_1 и b_2, используя формулы перехода;

$$\beta_i = b_i \frac{\sigma_{x_i}}{\sigma_y} \qquad b_i = \beta_i \frac{\sigma_y}{\sigma_{x_i}}$$

Значение a определить из соотношения

$$a = \bar{y} - b_1\bar{x}_1 - b_2\bar{x}_2 .$$

В уравнение множественной регрессии

$$\hat{y} = a + b_1\, x_1 + b_2\, x_2 .$$

подставить полученные коэффициенты:

Пример.

Линейное уравнение множественной регрессии ***y*** от $\boldsymbol{x_1}$ и $\boldsymbol{x_2}$ (Рисунок 12) имеет вид $\hat{y} = a + b_1\, x_1 + b_2\, x_2$.

	A	B	C	D
4	№ предпр.	y	x1	x2
5	1	7	3,9	10
6	2	7	3,9	14
7	3	7	3,7	15
8	4	7	4	16
9	5	7	3,8	17
10	6	7	4,8	19
11	7	8	5,4	19
12	8	8	4,4	20
13	9	8	5,3	20
14	10	10	6,8	20
15	Сумма	76	46	170
16	Среднее	7,6	4,6	17

Рисунок 12. Диапазон входных данных для множественной регрессии

Оценим показатели варьирования признаков.

Сводную таблицу основных статистических характеристик для нескольких массивов данных получим с помощью инструмента EXCEL анализа данных Описательная статистика.

Для этого выполним следующие шаги:

- Вводим исходные данные (Рисунок 12),
- Выполним команду меню *Данные, Анализ данных, Описательная статистика* (Рисунок 13).

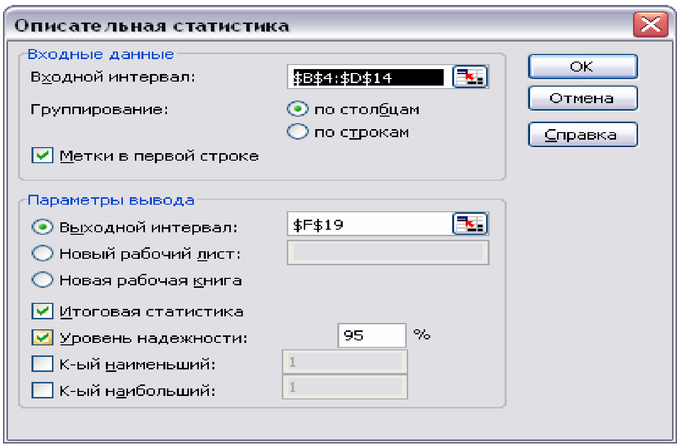


Рисунок 13. Окно для ввода исходных данных *y*, x_1 и x_2

Входной интервал включает диапазон y, x_1 и x_2.

Задаем уровень надежности среднего 95%, т.е. уровень значимости будет равен 0,05.

По результатам вычислений определяем уровень варьирования признаков, взяв из *Описательной статистики* необходимые данные:

$$\upsilon_y = \frac{\sigma_y}{\bar{y}}100\% \qquad \upsilon_{x1} = \frac{\sigma_{x1}}{\bar{x}_1}100\% \qquad \upsilon_y = \frac{\sigma_y}{\bar{y}}100\% .$$

где σ_{x1} – стандартное отклонение по x_1,

σ_{x2} – стандартное отклонение по x_2,

σ_y – стандартное отклонение по y,

$\bar{x}_1$, $\bar{x}_2$ – $\bar{y}$еднее арифметическое квадратов отклонений по x_1, по x_2, по y соответственно.

Имеем U_y=12,71%; U_{x1}=21,55; U_{x2}=19,41.

Приходим к выводу об умеренном уровне варьирования признаков, не превышающем 35% (т.е. совокупность данных по предприятиям однородна), и возможности применения метода наименьших квадратов (МНК) для их изучения.

Анализируем линейные коэффициенты парной корреляции.

Матрицу парных коэффициентов корреляции переменных получаем, используя инструмент анализа данных *Корреляция* (Рисунок 14**).**

Выполним команду *Меню, Данные, Анализ данных, Корреляция* и заполним диалоговое окно

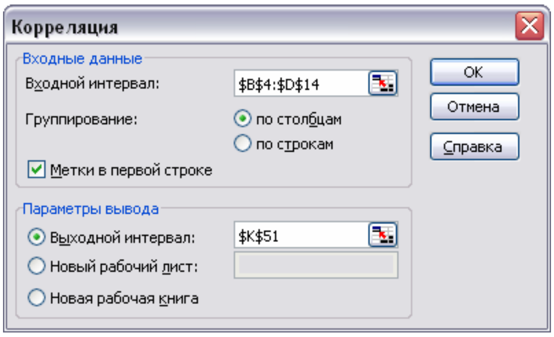


Рисунок 14. Окно для ввода исходных данных y, x_1 и x_2

2 Сервис-Анализ данных-Корреляция			
	y1 объясняемая зависимая	x1 объясняющая независимая переменная	x2 объясняющая независимая переменная
y1 объясняемая зависимая переменная	1		
x1 объясняющая независимая	0,9168	1	
x2 объясняющая независимая	0,5925	0,66251	1

Рисунок 15. Коэффициенты парной корреляции

Вывод. Из анализа коэффициентов ***парной*** корреляции (

Рисунок **15**) следует, что значение $r_{y1\,x1}$=0,9168 указывает на тесную связь между y_1 и x_1, а значение $r_{x2\,x1}$=0,6625 говорит о тесной связи между x_2 и x_1, при этом $r_{y1\,x2}$=0,5925<. $r_{x2\,x1}$, т.е. x_2 можно пренебречь.

Для расчёта коэффициентов частной корреляции в ППП EXCEL нет специального инструмента. Их можно рассчитать по рекуррентной формуле (11) через коэффициенты парной корреляции:

$r_{y\,x1/x2}$=0,869; $r_{yx2/x1}$=-0,05; $r_{x1\,x2/y}$=0,37

Вывод. Из анализа ***частных*** коэффициентов множественной корреляции следует, что значение $r_{y\,x1/x2}$=0,869 (x_2 фиксируем) указывает на тесную связь между y и x_1, а значение $r_{y\,x2/x1}$= – 0,05 (x_1 фиксируем) говорит о слабой связи между x_2 и y.

В связи с этим, для улучшения данной модели можно исключить из неё фактор x_2, как малоинформативный, недостаточно статистически надёжный.

Для вычисления методом стандартизации коэффициентов множественной регрессии построим искомое уравнения в стандартизованном масштабе:

$$t_y=\beta_1 \cdot t_{x1}+ \beta_2 \cdot t_{x2}.$$

Расчёт β-коэффициентов выполняем по формулам (12).

В результате получаем β-коэффициенты

$$\beta_1 = 0{,}9343; \ \beta_2 = -0{,}0265.$$

Уравнение в стандартизованном масштабе

$$t_y = 0{,}9343\ t_{x1} - 0{,}0265\ t_{x2}.$$

Для построения уравнения в естественной форме рассчитаем b_1 и b_2, используя формулы перехода;

$$\beta_i = b_i \frac{\sigma_{x_i}}{\sigma_y} \qquad b_i = \beta_i \frac{\sigma_y}{\sigma_{x_i}}$$

В результате получаем: b_1 =0,9108; b_2= – 0,007756.

Значение a = 3,5422923 определим из соотношения

$$a = \bar{y} - b_1\bar{x}_1 - b_2\bar{x}_2 \cdot$$

Уравнение множественной регрессии:

$$\hat{y} = 3{,}54 + 0{,}9108\ x_1 - 0{,}0078\ x_2$$

Вывод. В данном примере статистически значимыми являются a и b_1, а величина b_2 сформировалась под воздействием случайных причин, поэтому фактор x_2, силу влияния которого оценивает b_2, можно исключить как несущественно влияющий, неинформативный.

ЛАБОРАТОРНАЯ РАБОТА №6. ПОЛУЧЕНИЕ МНОЖЕСТВЕННОЙ РЕГРЕССИИ ЧЕРЕЗ ФУНКЦИИ РЕГРЕССИЯ И ПОИСК РЕШЕНИЯ

Второй способ получения оценок параметров линейного уравнения множественной регрессии – с помощью инструмента EXCEL *Регрессия*. Третий способ – вычисление параметров с помощью функции Поиск решения.

Цель работы.

1. Вычислить параметры с помощью функции *Регрессия* ***и проанализировать их взаимосвязь.***

2. Рассчитать параметры с помощью функции *Поиск решения* ***и сделать анализ.***

Линейное уравнение множественной регрессии y от x_1 и x_2 имеет вид:

$$\hat{y} = a + b_1\, x_1 + b_2\, x_2\,.$$

В качестве оценок параметров b_0 и b_i принимаются величины $\hat{b}_0$ и $\hat{b}_i$, минимизирующие сумму квадратов отклонений наблюдаемых значений результативного признака y_k от расчётных теоретических значений:

$$S(\hat{b}_0, \hat{b}_1, \hat{b}_2) = \sum (y_k - (\hat{b}_0 + \hat{b}_1 \cdot x_{1k} + \hat{b}_2 x_{2k}))^2 \to \min$$

где x_{ik} и y_k – данные наблюдения, $\hat{b}_0$ и $\hat{b}_i$ – переменные.

Чтобы найти минимум функции двух переменных, нужно вычислить частные производные по каждому параметру и приравнять их к нулю:

В результате получим систему линейных уравнений:

$$\begin{cases} \sum y_k = \hat{b}_0 \cdot n + \hat{b}_1 \cdot \sum x_{1k} + \hat{b}_2 \cdot \sum x_{2k} \\ \sum x_1 y_k = \hat{b}_0 \cdot \sum x_{1k} + \hat{b}_1 \cdot \sum x^2{}_{1k} + \hat{b}_2 \cdot \sum \cdot x_{1k} x_{2k} \\ \sum x_2 y_k = \hat{b}_0 \cdot \sum x_{2k} + \hat{b}_1 \cdot \sum x_{1k} \cdot x_{2k} + \hat{b}_2 \cdot \sum x_{2k}^2 \end{cases}. \tag{13}$$

3. Рассчитать средние частные коэффициенты эластичности и дать на их основе сравнительную оценку силы влияния факторов на результат.

Частные коэффициенты эластичности показывают, на сколько процентов в среднем изменяется признак-результат y с увеличением признака-фактора x_i на 1 % от своего среднего уровня при фиксированном положении других факторов модели. Частные коэффициенты эластичности рассчитываются по формуле;

$$Э_i = b_i \cdot \frac{\bar{x}_i}{\bar{y}}\,. \tag{14}$$

4. Вычисление параметров с помощью функции Поиск решения.

5. С помощью общего и частного F-критерия Фишера оценить статистическую надёжность уравнения регрессии и коэффициента корреляции множественной регрессии ($R^2_{yx1\,x2}$).

Сравнить значения скорректированного и нескорректированного линейных коэффициентов множественной детерминации

С помощью F-критерия Фишера оценить целесообразность включения в уравнение множественной регрессии фактора x_1 после x_2 и фактора x_2 после x_1.

6. Сравнение результатов трех способов.

Пример 1.

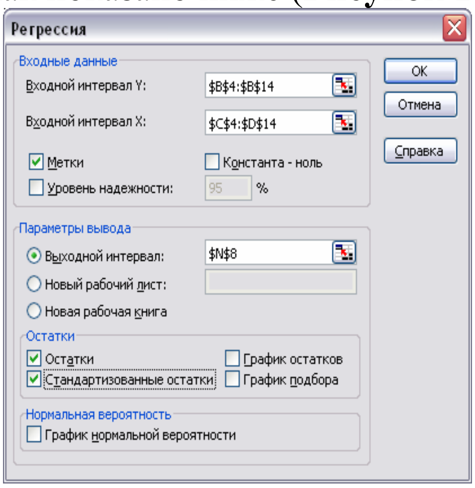
Получаем оценку параметров линейного уравнения множественной регрессии для y; x_1 и x_2, заданных в таблице (

Рисунок **16**), с помощью инструмента EXCEL *Регрессия*:

	A	B	C	D
4	№ предпр.	y	x1	x2
5	1	7	3,9	10
6	2	7	3,9	14
7	3	7	3,7	15
8	4	7	4	16
9	5	7	3,8	17
10	6	7	4,8	19
11	7	8	5,4	19
12	8	8	4,4	20
13	9	8	5,3	20
14	10	10	6,8	20
15	**Сумма**	76	46	170
16	**Среднее**	7,6	4,6	17

Рисунок 16. Исходный диапазон

Выполняем команду меню *Данные, Анализ данных, Регрессия*. Заполняем диалоговое окно, как показано ниже (**Рисунок 17**):



ВЫВОД ИТОГОВ	
Регрессионная статистика	
Множественный R	0,9169867
R-квадрат	0,8408646
Нормированный R-квадрат	0,7953973
Стандартная ошибка	0,4369926
Наблюдения	10

Рисунок 17. Результат отображения функции *Регрессия*

Дисперсионный анализ					
	df	*SS*	*MS*	*F*	*Значимость F*
Регрессия	2	7,063262493	3,531631247	18,49384685	0,001607622
Остаток	7	1,336737507	0,190962501		
Итого	9	8,4			

Рисунок 18. Критерий Фишера

	Коэффициенты a b_1 b_2	Стандартная ошибка	t-статистика (t -критерий Стьюдента)
Y-пересечение	3,5422923	0,799560587	4,430298743
x1 объясняющая независимая переменная	0,9107742	0,196217472	4,641656827
x2 объясняющая независимая переменная	-0,007756	0,058931944	-0,131610871

Рисунок 19. Коэффициенты множественной регрессии

Уравнение регрессии (Рисунок 19):

$$\hat{y} = 3{,}54 + 0{,}9108\, x_1 - 0{,}0078\, x_2 .$$

Результаты анализа:

- Значения случайных ошибок (Рисунок 19) параметров $\boldsymbol{a}$, $\boldsymbol{b_1}$ и $\boldsymbol{b_2}$ с учётом округления соответственно равны 0,7996 0,1962 и 0,0589 Они показывают, какое значение данной характеристики сформировалось под влиянием случайных факторов.
- Значения t-критерия Стьюдента (Рисунок 19) соответственно равны 4,4303 4,6417 и -0,1316. Если значение *t*-критерия больше 2-3, можно сделать вывод о существенности данного параметра, который формируется под воздействием неслучайных причин.

В данном примере статистически значимыми являются $\boldsymbol{a}$ и $\boldsymbol{b_1}$, а величина $\boldsymbol{b_2}$ сформировалась под воздействием случайных причин, поэтому фактор x_2 , силу влияния которого оценивает b_2 , можно исключить как несущественно влияющий, неинформативный.

Главным показателем качества модели множественной регрессии, как и для парной корреляции, является коэффициент множественной детерминации $\boldsymbol{R^2}$, который характеризует совместное влияние всех факторов на результат.

Расчёт линейного коэффициента множественной корреляции:

$$\boldsymbol{R}_{yx1x2} = \sqrt{\boldsymbol{r}_{yx1} \cdot \beta_1 + \boldsymbol{r}_{yx2} \cdot \beta_2}$$

Получаем $\boldsymbol{R}_{y\,x1\,x2}$ = 0,9170 (сравните с результатами функции *Регрессии* (Рисунок 17)). Зависимость $\boldsymbol{y}$ от $\boldsymbol{x_1}$ и $\boldsymbol{x_2}$ характеризуется как тесная.

При вычислении параметров с помощью функции Поиск решения необходимо подсчитать коэффициенты при $\hat{b}_0$ и $\hat{b}_i$ для системы уравнений (14).. Расчеты сведены в таблицу (

Рисунок 20).

	A	B	C	D	E	F	G	H	I
4	№ предпр.	y	x1	x2	y*x1	y*x2	x1^2	x2^2	x1*x2
5	1	7	3,9	10	27,3	70	15,21	100	39
6	2	7	3,9	14	27,3	98	15,21	196	54,6
7	3	7	3,7	15	25,9	105	13,69	225	55,5
8	4	7	4	16	28	112	16	256	64
9	5	7	3,8	17	26,6	119	14,44	289	64,6
10	6	7	4,8	19	33,6	133	23,04	361	91,2
11	7	8	5,4	19	43,2	152	29,16	361	102,6
12	8	8	4,4	20	35,2	160	19,36	400	88
13	9	8	5,3	20	42,4	160	28,09	400	106
14	10	10	6,8	20	68	200	46,24	400	136
15	Сумма	76	46	170	357,5	1309	220,44	2988	801,5
16	Среднее	7,6	4,6	17	35,75	130,9	22,044	298,8	80,15

Рисунок 20. Расчеты для системы уравнений (14).

Подставим известные значения и получим следующую систему линейных уравнений

$$\begin{cases} 76 = \hat{b}_0 \cdot 10 + \hat{b}_1 \cdot 46 + \hat{b}_2 \cdot 170 \\ 357{,}5 = \hat{b}_0 \cdot 46 + \hat{b}_1 \cdot 220{,}44 + \hat{b}_2 \cdot 801{,}5 \\ 1309 = \hat{b}_0 \cdot 170 + \hat{b}_1 \cdot 801{,}5 + \hat{b}_2 \cdot 2988 \end{cases}$$

Решаем систему, применяя инструмент ППП EXCEL *Поиск решения*, предварительно разместив данные в соответствии с (

Рисунок **21**):

	C	D	E	F	G
19	10	46	170	0	76
20	46	220,44	801,5	0	357,5
21	170	801,5	2988	0	1309
22	**b0**	**b1**	**b2**		
23	0,0000	0,0000	0,0000		

Рисунок 21. Размещение массива коэффициентов

В ячейки с ***F19*** по ***F21*** добавляем формулы (

Таблица **13**):

Таблица 13. Формулы в ячейках

Ячейки	Формулы
F19	СУММПРОИЗВ($C$23: $E$23;C19:E19),
F20	СУММПРОИЗВ($C$23: $E$23; C20:E20)
F21	СУММПРОИЗВ($C$23: $E$23; C21:E21)

Далее выполняем команду меню *Данные****,*** *Поиск решения* и заполняем, как показано на рисунке ниже (

Рисунок **22**).

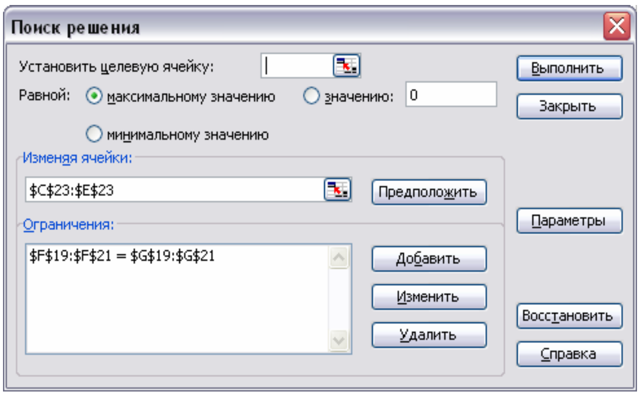


Рисунок 22. Окно *Поиска решения*

Результат выполнения

	C	D	E	F	G
19	10	46	170	76,000001	76
20	46	220,44	801,5	357,5	357,5
21	170	801,5	2988	1309	1309
22	**b0**	**b1**	**b2**		
23	3,5423	0,9108	-0,0078		

Рисунок 23. Результат *Поиска решения*

Таким образом, получаем уравнение множественной регрессии
$\hat{y} = 3{,}54 + 0{,}9108x_1 - 0{,}0078x_2$

Вывод. Значение коэффициента при второй объясняющей переменной очень мало, что указывает на очень малое влияние второй объясняющей переменной на результативный фактор, поэтому фактор x_2, силу влияния которого оценивает b_2, можно исключить как несущественно влияющий, неинформативный.

Частные коэффициенты эластичности показывают, на сколько процентов в среднем изменяется признак-результат $\boldsymbol{y}$ с увеличением признака-фактора $\boldsymbol{x_i}$ на 1 % от своего среднего уровня при фиксированном положении других факторов модели. Частные коэффициенты эластичности рассчитываем по формуле (13);

После расчёта получаем $Э_1$=0,5513; $Э_2$= – 0,0173.

Из вычислений методом стандартизации $\boldsymbol{\beta}$-коэффициенты равны: $\boldsymbol{\beta_1}$ =0,9343, $\boldsymbol{\beta_2}$= – 0,0265.

Вывод. В нашем случае $Э_1 > Э_2$, и $\beta_1 > \beta_2$, следовательно второй фактор имеет очень малое влияние на фактор-результат.

Расчёт общего и частного F-критерия Фишера.

Общий F-критерий проверяет гипотезу H_0 о статистической значимости уравнения регрессии и показателя тесноты связи ($R^2=0$)

$$F_{набл} = \frac{R^2_{yx1x2}}{1 - R^2_{yx1x2}} \cdot \frac{n - m - 1}{m}$$

где n – число наблюдений,

m – количество пар оцениваемых параметров в уравнении регрессии.

Получаем следующий результат $F_{набл}$=18,49 при n=10 и $m=2$.

По таблицам распределения находим критическое значение $\boldsymbol{F}$-критерия в зависимости от уровня значимости $\boldsymbol{\alpha}$ (обычно его берут равным 0,05) и двух чисел степеней свободы $k1=m-1$ и $k2=n-m,$ где m –количество пар оцениваемых параметров в уравнении регрессии, а n – число наблюдений $F_{табл}$ =5,32.

Вывод. Так как $F_{табл} < F_{набл}$, то с вероятностью 0,95 делаем заключение о статистической значимости уравнения в целом и показателя тесноты связей, которые сформировались под неслучайным воздействием факторов $\boldsymbol{x_1}$ и $\boldsymbol{x_2}$.

Частные $\boldsymbol{F}$-критерии F_{x1} и F_{x2} оценивают статистическую значимость присутствия факторов $\boldsymbol{x_1}$ и $\boldsymbol{x_2}$ в уравнении множественной регрессии, оценивают целесообразность включения в уравнение фактора $\boldsymbol{x_1}$ после того, как в него был включен фактор $\boldsymbol{x_2}$. Соответственно, F_{x2} указывает на целесообразность включения в уравнение фактора $\boldsymbol{x_2}$ после того, как в него был включен фактор $\boldsymbol{x_1}$:

$$F_{x1факт} = \frac{R^2_{yx1x2} - r^2_{yx2}}{1 - R^2_{yx1x2}} \cdot \frac{n - m - 1}{m}$$

Вывод. После расчётов получаем $F_{x1факт}$ = 10,7725. Сравниваем с $F_{табл}$ = 5,32. Видим, что $F_{табл} < F_{x1факт}$, приходим к выводу о целесообразности включения в модель фактора x_1 после фактора x_2.

Рассчитываем

$$F_{x2факт} = \frac{R^2_{yx1x2} - r^2_{yx1}}{1 - R^2_{yx1x2}} \cdot \frac{n - m - 1}{m}$$

После расчётов получаем $F_{x2факт}$ = -0,00866.

Вывод. Низкое значение $F_{x2факт}$ свидетельствует о статистической незначимости прироста парного коэффициента корреляции r_{yx1}^2 за счёт включения в модель фактора x_2 после фактора x_1.

Следовательно, подтверждается нулевая гипотеза H_0 о нецелесообразности включения в модель фактора x_2.

Сравнение результатов разных методов.

Вычисление ***методом стандартизации*** переменных дает следующее линейное уравнение множественной регрессии:

$$\hat{y} = 3{,}54 + 0{,}9108\, x_1 - 0{,}0078\, x_2 .$$

Вывод. В данном примере статистически значимыми являются a и b_1, а величина b_2 сформировалась под воздействием случайных причин, поэтому фактор x_2, силу влияния которого оценивает b_2, можно исключить как несущественно влияющий, неинформативный.

При вычислении параметров с помощью ***функции Регрессия*** уравнение регрессии:

$$\hat{y} = 3{,}54 + 0{,}9108\, x_1 - 0{,}0078\, x_2 .$$

позволяет сделать следующий анализ:

- величины случайных ошибок параметров a, b_1 и b_2, равные 0,7996 ; 0,1962 и 0,0589, показывают, что характеристика b_2 сформировалась под влиянием случайных факторов;
- значения t-критерия Стьюдента, равны 4,4303; 4,6417 и -0,1316, что статистически значимыми являются a и b_1, поэтому фактор x_2, можно исключить как несущественно влияющий, неинформативный.

Применяя инструмент ***Поиск решения,*** получаем уравнение множественной регрессии $\hat{y} = 3{,}54 + 0{,}9108\, x_1 - 0{,}0078\, x_2 .$

Вывод. Значение коэффициента при второй объясняющей переменной очень мало, что указывает на очень малое влияние второй объясняющей переменной на результативный фактор, поэтому фактор x_2, силу влияния которого оценивает b_2, можно исключить как несущественно влияющий, неинформативный.

ЛАБОРАТОРНАЯ РАБОТА №7.
ВРЕМЕННЫЕ РЯДЫ В ЭКОНОМЕТРИЧЕСКИХ ИССЛЕДОВАНИЯХ

Одна из важнейших задач (этапов) анализа временного (динамического) ряда состоит в прогнозировании на его основе развития изучаемого процесса.

Цель работы. Подобрать тренды и сделать прогноз.

Порядок выполнения работы.

1. Провести сглаживание исходного массива у методом скользящих средних.

Используя простую среднюю арифметическую с интервалом сглаживания m=3года, скользящие средние находим по формуле:

$$y_{i\ усред} = \frac{\sum_{i=t-p}^{t+p} y_i}{m},$$

где t – номер года, m=(2p-1) – нечетное число, при m =3 p=1.

Получаемый ряд скользящих средних ведет себя более гладко, чем исходный, из-за усреднения отклонений ряда.

2. Провести расчет параметров линейного тренда методом наименьших квадратов с помощью статистической функции ЛИНЕЙН(y,x,1,1).

3. Провести расчет параметров логарифмического тренда методом наименьших квадратов с помощью статистической функции ЛГРФПРИБЛ(y,x,1,1).

4. Подобрать различные виды трендов, построенных графически.

5. Выбрать наилучший среди всех выше названных трендов по значению коэффициента детерминации R^2.

6. Сделать прогноз на несколько периодов вперед на наилучшем тренде.

Пример.

Рассмотрим в качестве примера динамику расходов (Рисунок 24) на покупку продовольственных товаров по годам, где ***x*** – среднемесячная заработная плата одного работающего, тыс.р., ***y*** – расходы на покупку продовольственных товаров, % к общему объему расходов.

№	Год	y	$y_{усред}$
1	1997	35	41,04
2	1998	44,375	44,58
3	1999	43,75	45,21
4	2000	45,625	42,92
5	2001	46,25	43,54
6	2002	36,875	40,83
7	2003	47,5	40,83
8	2004	38,125	36,04
9	2005	36,875	36,04
10	2006	33,125	35,83
11	2007	38,125	35,42
12	2008	36,25	39,17
13	2009	31,875	
14	2010	49,375	

Рисунок 24. *Динамика расходов*

Проведем сглаживание исходного массива у методом скользящих средних (Рисунок 24**).**

Расчет линейного тренда проведем методом наименьших квадратов с помощью статистической функции *ЛИНЕЙН(y,x,1,1)*.

Таблица 14. Функция ЛИНЕЙН(y,x,1,1)

Значение коэффициента *b*	0,39	Значение коэффициента *a*	21,20
Среднеквад. отклонение *b*	0,12	Среднеква. отклонение *a*	6,16
Коэффициент детерминации r^2	0,49	Среднеквад. отклонение *y*	2,69
F-статистика	9,58	Число степен. свободы	10,00
Регрессионная сумма квадратов	69,43	Остаточная сумма квадратов	72,45

Записываем уравнение линейного тренда, полученного методом наименьших квадратов с помощью статистической функции *ЛИНЕЙН(y,x,1,1)*.

$\boldsymbol{y}_{\text{лин тр}}$=0,39+21,20$\boldsymbol{x}$.

Коэффициент детерминации r^2=0,49.

Расчет логарифмического тренда проведем методом наименьших квадратов с помощью статистической функции *ЛГРФПРИБЛ(y,x,1,1)*.

Таблица 15 Функция ЛГРФПРИБЛ(y,x,1,1)

Значение коэффициента *b*	1,01	Значение коэффициента *a*	24,83
Среднеквадратическое отклонение *b*	0,00	Среднеквадратическое отклонение *a*	0,15
Коэффициент детерминации r^2	0,49	Среднеквадратическое отклонение *y*	0,07
F-статистика	9,66	Число степеней свободы	10,00
Регрессионная сумма квадратов	0,04	Остаточная сумма квадратов	0,05

Записываем уравнение логарифмического тренда, полученного методом наименьших квадратов с помощью статистической функции *ЛГРФПРИБЛ(y,x,1,1)*

$$y_{\text{логар тр}}=24{,}83\cdot 1{,}01^x.$$

Коэффициент детерминации r^2= 0,49.

Подбор трендов, построенных графически

Для получения линий трендов необходимо построить с помощью *Мастера диаграмм* сначала график расходов на покупку продовольственных товаров по годам, а затем подобрать линии трендов, задав соответствующие параметры.

Для полиномиального тренда нужно задать степень аппроксимирующего полинома. В качестве дополнительной информации на диаграмме можно отобразить уравнение регрессии и коэффициент детерминации.

На графике тренда (Рисунок 25) по величине коэффициента детерминации можно судить о качестве аппроксимации.

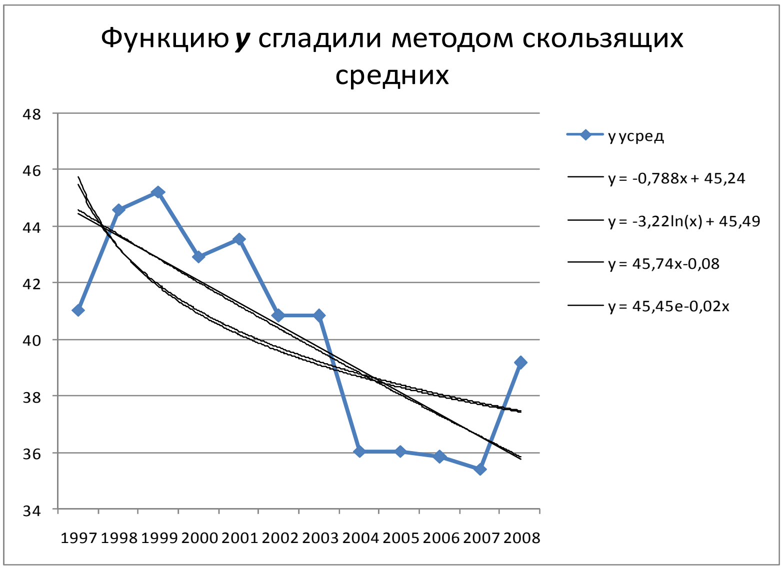


Рисунок 25. Графики линейного, логарифмического, степенного и экспоненциального трендов

Ниже приведены графики полиномиальных (второй, третьей и четвертой степеней) трендов:

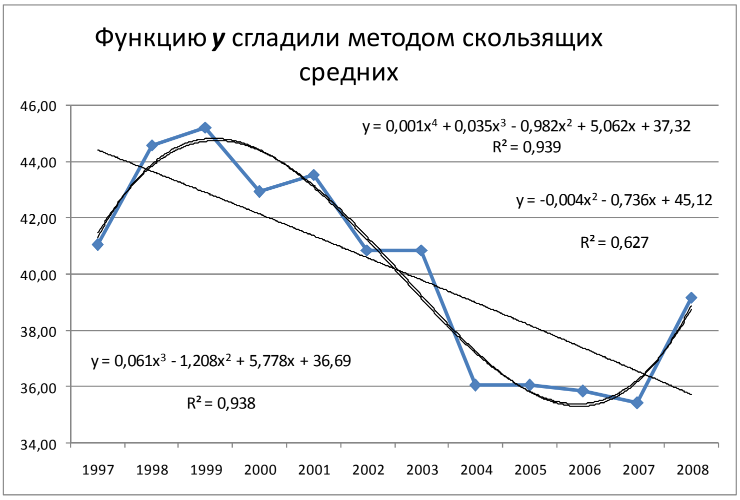


Рисунок 26. Графики полиномиальных трендов

Выбор наилучшего тренда

№	Тренд	Уравнение регрессии	Величина достоверности аппроксимации r^2
1	Линейный	$y = -0,788x + 45,24$	0,627
2	ЛИНЕЙН(y,x,1,1)	$y_{лин\ тр}=0,39+21,20\boldsymbol{x}$	0,49
3	Показательный	$y = -3,22\ln(x) + 45,49$	0,46
4	ЛГРФПРИБЛ(y,x,1,1)	$y_{логар\ тр}=24,83 \cdot 1,01^{x}$	0,49
5	Степенной	$y = 45,74x^{-0,08}$	0,460
6	Экспоненциальный	$y = 45,45e^{-0,02x}$	0,624
7	Полиномиальный, второй степени	$y = -0,004x^2 – 0,736x + 45,12$	0,627
8	Полиномиальный, третьей степени	$y = 0,061x^3 – 1,208x^2 + 5,778x + 36,69$	0,938
9	Полиномиальный, четвертой степени	$y = 0,001x^4 + 0,035x^3 – 0,982x^2 + +5,062x + 37,32$	0,939

Вывод. Среди трендов полиномиальный четвертой степени имеет максимум $r^{2}=$ 0,939.

Прогноз нескольких периодов вперед.

Необходимо сделать прогноз расходов на покупку продовольственных товаров нескольких периодов вперед на наилучшем тренде

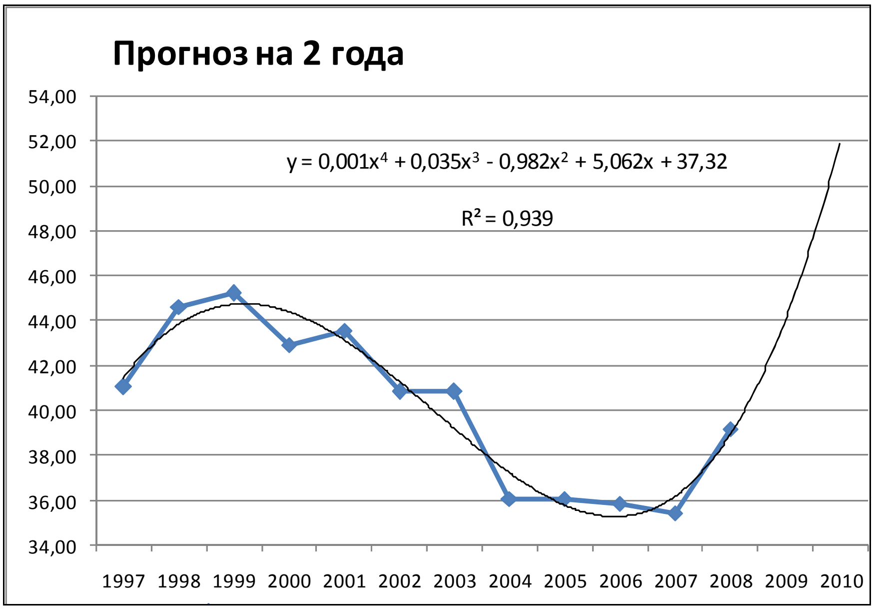


Рисунок 27. Прогноз на тренде полиномиальном четвертой степени

Вывод; Наблюдается резкий подъем расходов на покупку продовольственных товаров по отношению к общему объему расходов.

ЛАБОРАТОРНАЯ РАБОТА №8.
СИСТЕМА ЭКОНОМЕТРИЧЕСКИХ УРАВНЕНИЙ

При использовании в экономических расчётах отдельных уравнений регрессии, в большинстве случаев предполагается, что факторы можно изменять независимо друг от друга. На самом деле изменение одного фактора, как правило, не может происходить при абсолютной неизменности других факторов. Следовательно, отдельно взятое уравнение множественной регрессии не может характеризовать истинные влияния отдельных признаков на вариацию результирующей переменной.

Цель работы.

Построить систему эконометрических уравнений, предварительно провести идентификацию модели. Оценить параметры системы с точки зрения экономики.

Порядок выполнения работы.

1. По имеющимся параметрам построить систему эконометрических уравнений.

Структурная форма модели описывает реальное экономическое явление или процесс.

Простейшая ***структурная форма*** модели имеет вид:

$$\begin{cases} y_1 = b_{12}\ y_2 + a_{11}\ x_1 + \varepsilon_1 \\ y_2 = b_{21}\ y_1 + a_{22}\ x_2 + \varepsilon_2 \end{cases} \qquad (14)$$

где $\boldsymbol{y}$ – эндогенные переменные, $\boldsymbol{x}$ – экзогенные переменные.

Использование метода наименьших квадратов (МНК) для оценивания структурных коэффициентов модели даёт смещённые и несостоятельные оценки. Поэтому обычно структурная форма модели преобразуется в приведённую форму.

Приведённая форма строится для того, чтобы по МНК – оценкам её параметров – определить оценки структурных коэффициентов. Зная оценки приведённых коэффициентов, можно определить параметры структурной формы модели, но только тогда, когда модель является идентифицированной.

2. Провести идентификацию модели

Правила идентификации модели:

Пусть ***M*** – число предопределённых переменных в модели

m – число предопределённых переменных в уравнении

K – число эндогенных переменных в модели

K **–** число эндогенных переменных в данном уравнении

A **–** матрица коэффициентов при переменных, не входящих в данное уравнение.

Необходимое и достаточное условия идентификации уравнения модели.

1. Если ***M-m>k+1*** и ранг матрицы ***A*** равен ***K-1***, то уравнение сверхидентифицированное.
2. Если ***M-m=k-1*** и ранг матрицы ***A*** равен ***K-1***, то уравнение точно идентифицированное.
3. Если ***M-m ≥ k-1*** и ранг матрицы ***A*** меньше ***K-1***, то уравнение неидентифицированное.
4. Если ***M-m<k-1*** , то уравнение неидентифицированное.

3. Оценить параметры системы

Структурная форма модели преобразуется в ***приведённую форму***, в которой коэффициенты при $\boldsymbol{x}$ определяются методом наименьших квадратов

$$\begin{cases} y_1 = \delta_{11}\ x_1 + \delta_{12}\ x_2 \\ y_2 = \delta_{21}\ x_1 + \delta_{22}\ x_2 \end{cases} \qquad (15)$$

Для нахождения значений $\boldsymbol{\delta_{11}}$ и $\boldsymbol{\delta_{12}}$ система нормальных уравнений имеет вид:

$$\begin{cases} \sum y_1 x_1 = \delta_{11} \sum x_1^{\,2} + \delta_{12} \sum x_1 x_2 \\ \sum y_1 x_2 = \delta_{11} \sum x_1 x_2 + \delta_{12} \sum x_2^{\,2} \end{cases} \qquad (16)$$

Аналогично строим систему нормальных уравнений для определения коэффициентов $\boldsymbol{\delta_{21}}$ и $\boldsymbol{\delta_{22}}$

$$\begin{cases} \sum y_2 x_1 = \delta_{21} \sum {x_1}^2 + \delta_{22} \sum x_1 x_2 \\ \sum y_2 x_2 = \delta_{21} \sum x_1 x_2 + \delta_{22} \sum {x_2}^2 \end{cases} \qquad (17)$$

Решаем системs в EXCEL с помощью инструмента *Поиск решения.*

4. Рассчитать коэффициенты структурной модели.

Пример.

Проанализируем приведенные (

Рисунок *28*) данные за пять лет.

Год	Годовое потребление продукта, y_1	Оптовая цена за кг., y_2	Доход на душу населения, x_1	Расходы по обработке продукта, x_2
1	60	5	1300	60
2	62	4	1300	56
3	65	4,2	1500	56
4	62	5	1600	63
5	66	3,8	1800	50
среднее	63,0	4,4	1500,0	57,0

Рисунок 28. Исходные данные

В этой таблице применены следующие обозначения;

y_1 – годовое потребление продукта,

y_2 – оптовая цена за кг.,

x_1 – доход на душу населения,

x_2 – расходы по обработке продукта.

Идентификация модели.

Система одновременных уравнений с двумя эндогенными и двумя экзогенными переменными имеет вид (14):

$$\begin{cases} y_1 = b_{12}\, y_2 + a_{11}\, x_1 + \varepsilon_1 \\ y_2 = b_{21}\, y_1 + a_{22}\, x_2 + \varepsilon_2 \end{cases}$$

Модель имеет две эндогенные (y_1, y_2) и две экзогенные (x_1, x_2) переменные.

Проверим каждое уравнение системы на необходимое (**Н**) и достаточное (**Д**) условие идентификации.

Первое уравнение:

Н: эндогенных переменных – 2 (y_1, y_2),

отсутствующих экзогенных – 1 (x_2).

Выполняется необходимое равенство 2=1+1, следовательно, Первое уравнение полностью идентифицированное.

Д: в первом уравнении отсутствует x_2 . ***M=2, m=1, K=2, k= 2.*** Проверяем : ***M-m=1***, ***k-1=1***, т.е. выполняется правило 2. Ранг матрицы $A=a_{22}$, определитель матрицы не равен нулю, ранг матрицы равен ***К-1=1***, следовательно первое уравнение точно идентифицированное.

Второе уравнение.

Н: эндогенных переменных – 2 (y_1, y_2)

отсутствующих экзогенных – 1 (x_1)
Выполняется необходимое равенство 2=1+1, следовательно, Второе уравнение полностью идентифицированное.
Д: во втором уравнении отсутствует x_1 . $M=2, m=1, K=2, k= 2$. Проверяем : $M\text{-}m=1$, $k\text{-}1=1$, т.е. выполняется правило 2. Ранг матрицы $A= a_{11}$, определитель матрицы не равен нулю, ранг матрицы равен $K\text{-}1=1$, следовательно второе уравнение точно идентифицированное.

Оценка параметров системы.

При решении системы (16)

$$\begin{cases}\sum y_1x_1 = \delta_{11}\sum x_1^{\,2} + \delta_{12}\sum x_1x_2\\ \sum y_1x_2 = \delta_{11}\sum x_1x_2 + \delta_{12}\sum x_2^{\,2}\end{cases}.$$

выразим x и y через отклонения от средних уровней.

И тогда матрица исходных данных будет иметь вид:

среднее	63	4,4	1500	57
Год	y_1	y_2	x_1	x_2
1	-3,0	0,6	-200	3,0
2	-1,0	-0,4	-200	-1,0
3	2,0	-0,2	0	-1,0
4	-1,0	0,6	100	6,0
5	3,0	-0,6	300	-7,0

Рисунок 29 Матрица через отклонения от средних уровней

Соответственно, дальнейшие вычисления дают значения суммы:

Год	y_1*x_1	y_1*x_2	x_1^2	x_2^2	x_1*x_2
1	600	-9	40000	9	-600
2	200	1	40000	1	200
3	0	-2	0	1	0
4	-100	-6	10000	36	600
5	900	-21	90000	49	-2100
сумма	1600	-37	180000	96	-1900

Рисунок 30. Расчет коэффициентов при δ_{11} ,δ_{12}

Тогда система нормальных уравнений составит

$$\begin{cases}1600 = 180000\cdot\delta_{11} - 1900\cdot\delta_{12}\\ -37 = -1900\cdot\delta_{11} + 96\cdot\delta_{12}\end{cases}$$

Решаем систему в EXCEL с помощью инструмента *Поиск решения,* предварительно заносим информацию в ячейки:

	B	C	D	E
19				
25				
26	180000	-1900	=СУММПРОИЗВ(B26:C26;$B$29:$C$29)	1600
27	-1900	96	=СУММПРОИЗВ(B27:C27;$B$29:$C$29)	-37
28	Неизвестные			
29	0	0		

Рисунок 31. Размещение матрицы коэффициентов

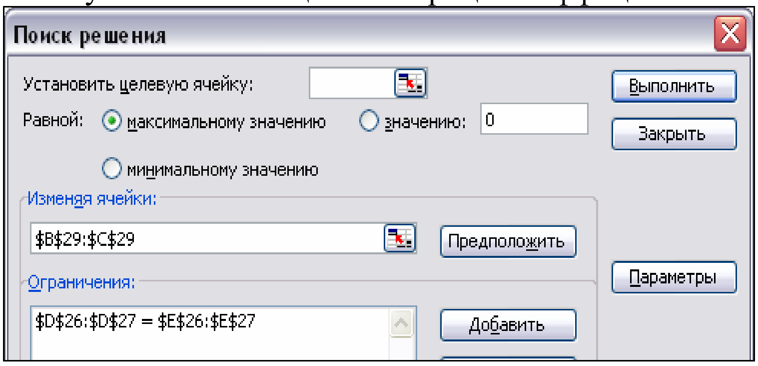


Рисунок 32. Параметры в окне *Поиск решения*

Получаем $\boldsymbol{\delta_{11}}$ = 0,00609 $\boldsymbol{\delta_{12}}$ =-0,26481, тогда

$$y_1 = 0{,}00609\,x_1 - 0{,}26481\,x_2 \tag{18}$$

Аналогично строим систему нормальных уравнений для определения коэффициентов $\boldsymbol{\delta_{21}}$ и $\boldsymbol{\delta_{22}}$ (17)

$$\begin{cases} \sum y_2 x_1 = \delta_{21} \sum x_1^{\,2} + \delta_{22} \sum x_1 x_2 \\ \sum y_2 x_2 = \delta_{21} \sum x_1 x_2 + \delta_{22} \sum x_2^{\,2} \end{cases}.$$

Получаем

$$\begin{cases} -160 = 180000 \cdot \delta_{21} - 1900 \cdot \delta_{22} \\ 10{,}2 = -1900 \cdot \delta_{21} + 96 \cdot \delta_{22} \end{cases}$$

Аналогично находим решение этой системы уравнений.

Следовательно, $\boldsymbol{\delta_{21}}$=0,00029 и $\boldsymbol{\delta_{22}}$=0,11207. Тогда второе уравнение примет вид:

$$y_2 = 0{,}00029\,x_1 + 0{,}11207\,x_2 \tag{19}$$

Приведённая форма модели с учетом полученных уравнений (18) и (19) примет вид:

$$\begin{cases} y_1 = 0{,}00609\,x_1 - 0{,}26481\,x_2 \\ y_2 = 0{,}00029\,x_1 + 0{,}11207\,x_2 \end{cases} \tag{20}$$

Из чего определяем коэффициенты структурной модели:

$$\begin{cases} \boldsymbol{y}_1 = 0{,}00609 \cdot \boldsymbol{x}_1 - 0{,}26481 \cdot \boldsymbol{x}_2 \\ \boldsymbol{x}_2 = \dfrac{\boldsymbol{y}_2 - 0{,}00029\boldsymbol{x}_1}{0{,}11207} \end{cases}$$

$$y_1 = 0{,}00609 \cdot x_1 - 0{,}26481 \cdot \frac{y_2 - 0{,}00029\ x_1}{0{,}11207} =$$
$$-2{,}36290 \cdot y_2 + 0{,}00678 \cdot x_1$$

Аналогично

$$\begin{cases} \boldsymbol{y}_2 = 0{,}00029 \cdot \boldsymbol{x}_1 + 0{,}11207 \cdot \boldsymbol{x}_2 \\ \boldsymbol{x}_1 = \dfrac{\boldsymbol{y}_1 + 0{,}26481 \cdot \boldsymbol{x}_2}{0{,}00609} \end{cases}$$

Тогда

$$y_2 = 0{,}00029 \cdot \frac{y_1 + 0{,}26481 \cdot x_2}{0{,}00609} + 0{,}11207 \cdot x_2 =$$
$$0{,}04762 \cdot y_1 + 0{,}12468 \cdot x_2$$

Итак, ***структурная форма модели*** имеет вид:

$$\begin{cases} \boldsymbol{y}_1 = -2{,}36290 \cdot \boldsymbol{y}_2 + 0{,}00678 \cdot \boldsymbol{x}_1 + \varepsilon_1 \\ \boldsymbol{y}_2 = 0{,}04762 \cdot \boldsymbol{y}_1 + 0{,}12468 \cdot \boldsymbol{x}_2 + \varepsilon_2 \end{cases}$$

В этой системе уравнений применены следующие обозначения;

$\boldsymbol{y_1}$ – годовое потребление продукта,

$\boldsymbol{y_2}$ – оптовая цена за кг.,

$\boldsymbol{x_1}$ – доход на душу населения,

$\boldsymbol{x_2}$ – расходы по обработке продукта.

Вывод. На годовое потребление продукта $\boldsymbol{y_1}$, причем уменьшение, большее влияние оказывает оптовая цена $\boldsymbol{y_2}$, чем доход на душу населения $\boldsymbol{x_1}$.

Вывод. На оптовую цену $\boldsymbol{y_2}$, причем увеличение, большее влияние оказывают расходы по обработке продукта $\boldsymbol{x_2}$.

СПИСОК ЛИТЕРАТУРЫ

1. Практикум по эконометрике / под ред. И.И. Елисеевой. – М.: Финансы и статистика, 2006. – 344 с.
2. Кремер, Н.Ш. Эконометрика / Н.Ш. Кремер, Б.А. Путко; под ред. Н.Ш. Кремера. – М.:ЮНИТИ, 2007. – 311 с.
3. Баклушина, О.А. Краткий курс по эконометрике / О.А. Баклушина – М.: О’кей-книга, 2007. – 127 с.
4. Эконометрика: учеб. / под ред. д-ра экон. наук, проф. В. С. Мхитаряна. – М.: Проспект, 2008. – 512 с.
5. Эконометрика: метод. указания к лабораторному практикуму / сост. С.А. Озерная, Т. В. Макаренко. – Самара: Изд-во Самар. гос. аэрокосм. ун-та, 2008. – 59 с.
6. Эконометрика: учеб./ под ред. И. И. Елисеевой. – М.: Проспект, 2009. – 451 с.

Приложение 1
Распределение Фишера (*F*-распределение)

k2	α	k1(число степеней свободы)											
		1	2	3	4	5	6	7	8	9	10	11	12
1	0,1	39,9	49,5	53,6	55,8	57,2	58,2	58,9	59,4	59,9	60,2	60,5	60,7
	0,05	161	200	216	225	230	234	237	239	241	242	243	244
2	0,10	8,53	9,00	9,16	9,24	9,29	9,33	9,35	9,37	9,38	9,39	9,40	9,41
	0,05	18,5	19,0	19,2	19,2	19,3	19,3	19,4	19,4	19,4	19,4	19,4	19,4
	0,01	98,5	99,2	99,2	99,2	99,3	99,3	99,4	99,4	99,4	99,4	99,4	99,4
3	0,10	5,54	5,46	5,39	5,34	5,31	5,28	5,27	5,25	5,24	5,23	5,22	5,22
	0,05	10,01	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,76	8,74
	0,01	34,1	30,3	29,5	28,7	28,2	27,9	27,7	27,5	27,3	27,2	27,1	27,1
4	0,10	4,54	4,32	4,19	4,11	4,05	4,01	3,98	3,95	3,94	3,92	3,91	3,90
	0,05	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,94	5,91
	0,01	21,2	18,0	16,7	16,0	15,5	15,2	15,0	14,8	14,7	14,5	14,4	14,4
5	0,10	4,06	3,78	3,62	3,52	3,45	3,40	3,37	3,34	3,32	3,30	3,28	3,24
	0,05	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,71	4,68
	0,01	16,3	13,3	12,1	11,4	11,0	10,7	10,5	10,3	10,2	10,1	9,96	9,89
6	0,10	3,78	3,46	3,29	3,18	3,11	3,05	3,01	2,98	2,96	2,94	2,92	2,90
	0,05	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,03	4,00
	0,01	13,7	10,9	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,79	7,72
7	0,10	3,59	3,26	3,07	2,96	2,88	2,83	2,78	2,75	2,72	2,70	2,68	2,67
	0,05	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,60	3,57
	0,01	12,2	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72	6,62	6,54	6,47
8	0,10	3,46	3,11	2,92	2,81	2,73	2,67	2,62	2,59	2,56	2,54	2,52	2,50
	0,05	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,31	3,28
	0,01	11,3	8,65	7,59	7,01	6,63	6,37	3,18	6,03	5,91	5,81	5,73	5,67
9	0,10	3,36	3,01	2,81	2,69	2,61	2,55	2,51	2,47	2,44	2,42	2,40	2,38
	0,05	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,10	3,07
	0,01	10,5	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35	5,26	5,18	5,11
10	0,10	3,29	2,92	2,73	2,61	2,52	2,46	2,41	2,38	2,35	2,32	2,30	2,28
	0,05	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,94	2,91
	0,01	10,0	7,56	6,55	5,99	5,54	5,39	5,20	5,06	4,94	4,85	4,77	4,71
11	0,10	3,23	2,86	2,66	2,54	2,45	2,39	2,34	2,30	2,27	2,25	2,23	2,21
	0,05	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,82	2,79
	0,01	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63	4,54	4,46	4,40

k2	α	k1(число степеней свободы)									
		1	2	3	4	5	6	7	8	9	10
12	0,05	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75
13	0,05	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67
14	0,05	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60
15	0,05	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54
16	0,05	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49
17	0,05	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45
18	0,05	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41
19	0,05	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38
20	0,05	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35

Приложение 2.
Распределение Стьюдента (Т-распределение)

	α – уровень значимости								
k	0,4	0,25	0,10	0,05	0,025	0,01	0,005	0,001	0,0005
1	0,325	1,000	3,078	6,341	12,706	31,821	63,657	31,831	6,366
2	0,289	0,816	1,886	2,920	4,303	6,965	9,925	22,327	31,6
3	0,277	0,765	1,638	2,353	3,182	4,541	5,841	10,214	12,94
4	0,271	0,741	1,533	2,132	2,776	3,747	4,604	7,173	8,610
5	0,267	0,727	1,476	2,015	2,571	3,365	4,032	5,893	6,859
6	0,265	0,718	1,440	1,943	2,447	3,143	3,707	5,208	5,959
7	0,263	0,711	1,415	1,895	2,365	2,998	3,499	4,785	5,405
8	0,262	0,706	1,397	1,860	2,306	2,896	3,355	4,501	5,041
9	0,261	0,703	1,383	1,833	2,262	2,821	3,250	4,297	4,781
10	0,260	0,700	1,372	1,812	2,228	2,764	3,169	4,144	4,587

	α – уровень значимости								
k	0,4	0,25	0,10	0,05	0,025	0,01	0,005	0,001	0,0005
11	0,259	0,697	1,363	1,796	2,201	2,718	3,106	4,025	4,437
12	0,259	0,695	1,356	1,782	2,179	2,681	3,055	3,930	4,318
13	0,258	0,694	1,350	1,771	2,160	2,650	3,012	3,852	4,221
14	0,258	0,692	1,345	1,761	2,145	2,624	2,977	3,787	4,140
15	0,258	0,691	1,341	1,753	2,131	2,602	2,947	3,733	4,073
16	0,258	0,690	1,337	1,746	2,120	2,583	2,921	3,686	4,015
17	0,257	0,689	1,333	1,740	2,110	2,567	2,898	3,646	3,965
18	0,257	0,688	1,330	1,734	2,101	2,552	2,878	3,610	3,922
19	0,257	0,688	1,328	1,729	2,093	2,539	2,861	3,579	3,883
20	0,257	0,687	1,325	1,725	2,086	2,528	2,845	3,552	3,850

Мультимедийное электронное издание

ЭКОНОМЕТРИКА

Мультимедийный электронный лабораторный практикум
для бакалавров в системе дистанционного обучения «Moodle»

Составитель
Озерная Светлана Алексеевна.

Редактор И.И. Спиридонова
Довёрстка И.И. Спиридонова

Электронный ресурс

Арт. Э24 / 2013.

Самарский государственный аэрокосмический университет.
443086, Самара, Московское шоссе, 34.

Изд-во Самарского государственного аэрокосмического университета.
443086, Самара, Московское шоссе, 34.

ВОПРОСЫ КОНТРОЛЯ
ВЫПОЛНЕНИЯ ЛАБОРАТОРНОГО ПРАКТИКУМА.

1. Каковы предпосылки «классического» метода наименьших квадратов (МНК)?
2. В чем суть МНК?
3. Приведите формулы расчета оценок коэффициентов линейной модели по МНК?
4. Какими свойствами обладают МНК-оценки классической линейной эконометрической модели?
5. Перечислите свойства фактической ошибки эконометрической модели.
6. Каким образом тестируется условие постоянства дисперсии ошибки модели?
7. Каким образом проверяется наличие автокорреляции ошибок модели?
8. Как оценивается дисперсия истинной ошибки модели?
9. Каковы последствия неправильного выбора состава независимых переменных модели?
10. Каковы особенности оценивания параметров с учетом наложенных ограничений?
11. Перечислите предпосылки метода максимального правдоподобия (ММП).
12. Опишите процедуру получения оценок параметров эконометрической модели с помощью ММП.
13. Какими свойствами обладают ММП-оценки параметров?
14. Каким образом оценивается дисперсия истинной ошибки модели?