ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
«САМАРСКИЙ ГОСУДАРСТВЕННЫЙ АЭРОКОСМИЧЕСКИЙ
УНИВЕРСИТЕТ имени академика С. П. КОРОЛЕВА»

ЭКОНОМЕТРИКА

С А М А Р А 2008

ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ

ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
«САМАРСКИЙ ГОСУДАРСТВЕННЫЙ АЭРОКОСМИЧЕСКИЙ
УНИВЕРСИТЕТ имени академика С. П. КОРОЛЕВА»

ЭКОНОМЕТРИКА

Утверждено Редакционно-издательским советом университета
в качестве методических указаний к лабораторному практикуму

САМАРА
Издательство СГАУ
2008

ББК У9(2) 21,0

Составители: ***С.А. Озерная, Т.В. Макаренко***

Рецензент В. Г. М и х а й л о в

Эконометрика: метод. указания к лабораторному практикуму / сост. *С.А. Озерная, Т.В. Макаренко.* – Самара: Изд-во Самар. гос. аэрокосм. ун-та, 2008. 59с.

Содержат методы решения эконометрических задач, примеры решённых задач. Составлены применительно к учебному плану по специальностям 080116 и 080507. В указаниях учтены требования государственного образовательного стандарта высшего профессионального образования по вышеуказанным специальностям и стандарта организации СТО СГАУ 02068410-003–2006.

Методические указания предназначены для очной, очно-заочной и вечерней форм обучения по специальностям «Математические методы в экономике», «Менеджмент организации», «Маркетинг». Подготовлены на кафедре математических методов в экономике.

ББК У9(2) 21,0

Лабораторная работа №1
ПАРНАЯ РЕГРЕССИЯ

ПАРНАЯ ЛИНЕЙНАЯ РЕГРЕССИЯ

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

Лабораторная работа №1 **а**

Парная линейная регрессия и корреляция

Параметры a и b **линейной регрессии y=a+b*x** рассчитываются в результате решения системы нормальных уравнений относительно a и b:

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

Расходы на покупку продовольственных товаров,% к общему объему расходов, **y**
Среднемесячная заработная плата одного работающего, тыс.р., **x**

Номер региона	x	y	x*y	x^2	y^2	$y_{лр}$=a+b*x	y-$y_{лр}$	abs((y-$y_{лр}$)/y)*100
1	4,5	68,8	309,6	20,25	4733,44	67,1	1,7	2,51
2	5,9	58,3	343,97	34,81	3398,89	59,3	-1,0	1,80
3	5,7	62,6	356,82	32,49	3918,76	60,5	2,1	3,43
4	7,2	52,1	375,12	51,84	2714,41	52,2	-0,1	0,14
5	6,2	54,5	337,9	38,44	2970,25	57,7	-3,2	5,86
6	6,0	57,1	342,6	36	3260,41	58,8	-1,7	2,97
7	7,8	51,0	397,8	60,84	2601	48,9	2,1	4,19
Σ	43,3	404,4	2463,8	274,7	23597,2	404,4	0,00	20,90
Среднее значение	6,186	57,771	351,973	39,239	3371,023			
Количество регионов	7							

Система нормальных уравнений составит:

$$\begin{cases} 7a + 43.3b = 404.4 \\ 43.3a + 274.67b = 2463.81 \end{cases}$$

Решаем систему методом определителей

$$\Delta = \begin{vmatrix} 7 & 43.3 \\ 43.3 & 274.67 \end{vmatrix}$$

$$\Delta\alpha = \begin{vmatrix} 404.4 & 43.3 \\ 2463.81 & 274.67 \end{vmatrix}$$

$$\Delta\beta = \begin{vmatrix} 7 & 404.4 \\ 43.3 & 2463.8 \end{vmatrix}$$

Δ= 47,80

Δα= 4393,575

Δβ= -263,85

α=Δα/Δ = 91,916

β=Δβ/Δ = – 5,5199

Получаем уравнение линейной регрессии $y_{лю}$

$y_{лр}$=91,916 - 5,5199*x

Коэффициенты a и b можно получить по другим формулам:

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

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

$$a = \overline{y} - b\overline{x}$$

$$\sigma_y^2 = \frac{\sum y^2}{n} - \overline{y}^2$$

b= -5,52

a= 91,916

$\overline{y}$ =-среднее значение у

$\overline{xy}$ =-среднее значение произведения х*у

$\overline{x}$ =-среднее значение **х**

*Экономическое содержание параметра **a** при линейной парной регрессии.* Если ***x***-фактор не может иметь нулевые значения, то трактовка ***a*** не имеет смысла. Если ***a*** < 0, то его экономическая трактовка может привести к абсурду. Интерпретировать можно лишь знак при ***a:*** если ***a*** > 0, то относительное изменение результата ***y*** происходит медленнее, чем изменение ***x***-фактора.

Экономический смысл коэффициента ***b*** – его величина показывает среднее изменение ***y***-фактора 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%.

Уравнение регрессии дополняется показателем тесноты связи. При использовании линейной регрессии в качестве такого показателя выступает линейный коэффициент корреляции r_{xy}. Линейный коэффициент корреляции находится в определенных пределах: (-1) ≤ r_{xy} ≤ (+1). При этом чем ближе r_{xy} к нулю, тем слабее корреляция, а при r_{xy} = 0 линия регрессии параллельна оси ***x ;*** чем ближе r_{xy} к (-1) или к (+1), тем сильнее корреляция, т.е. зависимость ***x*** и ***y*** близка к линейной.

Линейное уравнение регрессии дополняется расчетом линейного коэффициента корреляции:

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

$\boxed{\sigma_x =}$ 0,98768

$\boxed{\sigma_y =}$ 5,78661

$r_{yx} =$ –0,94215

Вывод. Значение r_{xy}= -0,94215 , т.е. близок к (-1) и существует сильная корреляция ***y*** и ***x,*** или иначе - зависимость ***y*** и ***x*** близка к линейной.

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

Проверить 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. Согласно основной идее дисперсионного анализа для парной регрессии число степеней свободы уравнения регрессии $k_1 = m - 1$, а число степеней свободы остаточной дисперсии $k_2 = n - m$, где m – число оцениваемых параметров уравнения регрессии (m = 2), n – число наблюдений (n = 7).

Коэффициент детерминации составит:

$r_{yx}^2 =$ 0,88765

Вывод: вариации Y на 88,8% объясняются вариацией X. На долю прочих факторов, не учитываемых в регрессии, приходится 11,2%.

F-критерий Фишера будет равен:

$$F = \frac{r_{yx}^2}{1 - r_{yx}^2}(n - 2) =$$ 39,5

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 регрессии статистически значимо.

Вывод. Коэффициент детерминации $0 \le r_{xy}^2 \le (+1)$, чем ближе к 1, тем регрессия аппроксимирует лучше эмпирические данные.

Ошибки аппроксимации для каждого наблюдения определяются как:

$$\left|\frac{y_i - y_{пр}}{y_i}\right| * 100$$

Средняя ошибка аппроксимации находится как средняя арифметическая простая из индивидуальных ошибок:

$$\bar{A} = \frac{1}{n}\sum\left|\frac{y_i - y_{пр}}{y_i}\right| * 100 =$$ 2,99%

Вывод: величина средней ошибки аппроксимации показывает хорошее соответствие

СТАНДАРТНАЯ ФУНКЦИЯ ЛИНЕЙН(y,x,1,1).

Параметры линейного приближения по методу наименьших квадратов можно получить, используя стандартную функцию **ЛИНЕЙН(y,x,1,1).**

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

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

Параметры линейного приближения по методу наименьших квадратов

Функция ЛИНЕЙН(y,x,1,1)

Значение коэффициента b	-5,51987	91,91579	Значение коэффициента a
Среднеквадратическое отклонение b	0,878238	5,501343	Среднеквадратическое отклонение a
Коэффициент детерминации r^2	0,887649	2,29497	Среднеквадратическое отклонение y
F-статистика	39,50336	5	Число степеней свободы
Регрессионная сумма квадратов	208,0598	26,33445	Остаточная сумма квадратов

Параметры линейной регрессии y=a+b*x были получены в результате решения системы нормальных уравнений относительно a и b:

Решение системы методом определителей

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

ПАРНАЯ СТЕПЕННАЯ РЕГРЕССИЯ

Лабораторная работа № 1 б

Парная степенная регрессия и корреляция

Регрессия степенная в виде $y=ax^b$. Для оценки параметров модели линеаризуем модель путем логарифмирования: Ln y=Ln a+b*Ln x. Обозначим Ln y=Y, Ln a=A, Ln x=X. Тогда получим: Y=A+b*X. Применяем метод наименьших квадратов и получаем систему уравнений:

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

Расходы на покупку продовольтвенных товаров,% к общему объему расходов, **y**
Среднемесячная заработная плата одного работающего, тыс.р., **x**

Номер региона	x	y	X=Ln x	Y=Ln y	Y*X	X^2	Y^2	$y_{стр} = e^a * x^b$	$(y-y_{стр})^2$	abs((y-y_{стр})/y)*100
1	4,5	68,8	1,5	4,23	6,36406	2,262	17,903	68,521	0,08	0,41
2	5,9	58,3	1,77	4,07	7,21625	3,15	16,529	58,634	0,11	0,57
3	5,7	62,6	1,74	4,14	7,1999	3,029	17,113	59,809	7,79	4,46
4	7,2	52,1	1,97	3,95	7,80387	3,897	15,628	52,288	0,04	0,36
5	6,2	54,5	1,82	4	7,29491	3,329	15,986	56,985	6,18	4,56
6	6,0	57,1	1,79	4,04	7,24732	3,21	16,360	58,070	0,94	1,70
7	7,8	51,0	2,05	3,93	8,07646	4,219	15,459	49,936	1,13	2,09
Σ	43,3	404,4	12,7	28,4	51,2028	23,1	114,978	404,244	16,265	14,146
Среднее значение	6,186	57,771	1,809	4,052	7,315	3,300	16,425			
Количество регионов	7									

Система нормальных уравнений составит:

$$\begin{cases} 7A + 12.664\,b = 28.362 \\ 12.664A + 23.098\,b = 51.203 \end{cases}$$

Решаем систему методом определителей

$$\Delta = \begin{vmatrix} 7 & 12.664 \\ 12.664 & 23.098 \end{vmatrix}$$

$$\Delta A = \begin{vmatrix} 28.362 & 12.664 \\ 51.203 & 23.098 \end{vmatrix}$$

$$\Delta\beta = \begin{vmatrix} 7 & 28.362 \\ 12.664 & 51.203 \end{vmatrix}$$

Δ=	1,306971	1,30910
ΔA=	6,655523	6,6707
Δβ=	-0,75181	-0,7554
α=ΔA/Δ	5,09233	
β=Δβ/Δ	-0,57523	

Получаем уравнение степенной регрессии $y_{стр}$

Ln $y_{стр}$=5,09233 - 0,57523*Ln x

Выполнив потенцирование, получим: $y = e^{5,09233} * x^{-0,57523}$ или

$y = 162,76791 * x^{-0,57523}$

162,768 $= e^{5,09233}$

Степенное уравнение регрессии дополняется расчетом индекса корреляции: R

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

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

R= 0,965

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

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

$r_{yx} =$ 1,708831

$\Bigg\}$ с листа линейной регрессии

Степенное уравнение регрессии дополняется расчетом коэффициента корреляции между Ln y и Ln x:

$$\boxed{r_{Lny\ Lnx} = \frac{\overline{LnxLny} - \overline{Lnx} * \overline{Lny}}{\sigma_{Lnx} * \sigma_{Lny}} =}$$ -7,9E-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:

$\boxed{R^2 =}$ 0,930609

Вывод: вариации Y на 93,1% объясняются вариацией X. На долю прочих факторов, не учитываемых в регрессии, приходится 6,9%.

F-критерий Фишера будет равен:

$$\boxed{F = \frac{R^2}{1 - R^2}(n-2) =}$$ 67,1

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 регрессии статистически значимо.

Ошибки аппроксимации для каждого наблюдения определяются как:

$$\left|\frac{y_i - y_{пр}}{y_i}\right| * 100$$

Средняя ошибка аппроксимации находится как средняя арифметическая простая из индивидуальных ошибок:

$$\overline{A} = \frac{1}{n}\sum\left|\frac{y_i - y_{пр}}{y_i}\right| * 100 =$$ 2,02%

Вывод: величина средней ошибки аппроксимации показывает хорошее соответствие расчетных и фактических данных.

ПАРНАЯ ЭКСПОНЕНЦИАЛЬНАЯ РЕГРЕССИЯ

Лабораторная работа №1 **в**

Парная экспоненциальная регрессия и корреляция

Регрессия экспоненциальная в виде y=ae^{bx}. Для оценки параметров модели линеаризуем модель путем логарифмирования: Ln y=Ln a+b*x. Обозначим Ln a=A. Применяем метод наименьших квадратов и получаем систему уравнений:

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

Расходы на покупку продовольственных товаров, % к общему объему расходов, **y**.

Среднемесячная заработная плата одного работающего, тыс. р., **x**.

Номер региона	x	y	Y=1/y	Y*x	x^2	ln(y)	x*ln(y)	$y_{экср}=a\ e^{bx}$	$(y-y_{экср})^2$	abs((y-$y_{экср}$)/y)*100
1	4,5	68,8	0,01453	0,0654	20,25	4,2312	19,040	67,346	2,11	2,11
2	5,9	58,3	0,01715	0,1012	34,81	4,0656	23,987	59,055	0,57	1,29
3	5,7	62,6	0,01597	0,0911	32,49	4,1368	23,580	60,174	5,89	3,88
4	7,2	52,1	0,01919	0,1382	51,84	3,9532	28,463	52,272	0,03	0,33
5	6,2	54,5	0,01835	0,1138	38,44	3,9982	24,789	57,415	8,50	5,35
6	6,0	57,1	0,01751	0,1051	36,00	4,0448	24,269	58,503	1,97	2,46
7	7,8	51,0	0,01961	0,1529	60,84	3,9318	30,668	49,410	2,53	3,12
Σ	43,3	404,4	0,12233	0,7676	274,67	28,362	174,80	404,176	21,595	18,538
Количество регионов	7									

Система нормальных уравнений составит:

$$\begin{cases} 7\ln a + 43.3\,b = 28.362 \\ 43.3\ln a + 274.77\,b = 174.8 \end{cases}$$

Решаем систему методом определителей

$$\Delta = \begin{vmatrix} 7 & 43.3 \\ 43.3 & 274.77 \end{vmatrix}$$

$$\Delta A = \begin{vmatrix} 28.362 & 43.3 \\ 174.8 & 274.77 \end{vmatrix}$$

$$\Delta\beta = \begin{vmatrix} 7 & 28.362 \\ 43.3 & 174.8 \end{vmatrix}$$

Δ= 47,80000

ΔA= 221,41646

Δβ= -4,48573

α=ΔA/Δ 4,63214

β=Δβ/Δ -0,09384

Выполнив потенцирование, получим: $y = e^{4,63214} * x^{-0,09384}$ **или**

$$y = 102,73403 * e^{-0,09384*x}$$

102,734 $= e^{4,63214}$

Экспоненциальное уравнение регрессии дополняется расчетом индекса корреляции: R

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

$\sum (y-\bar{y})^2 = n * \sigma_y^2 =$ **234,39429**

R= 0,9528211

$\sigma_x =$ 0,987679

$\sigma_y =$ 5,786614

с листа линейной регрессии

Коэффициент детерминации составит:

$R^2 =$ 0,907868

Вывод: вариации Y на 90,1% объясняются вариацией X. На долю прочих факторов, не учитываемых в регрессии, приходится 9,9%.

F-критерий Фишера будет равен:

$$F = \frac{R^2}{1 - R^2}(n-2) = 49{,}3$$

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 регрессии статистически значимо.

Ошибки аппроксимации для каждого наблюдения определяются как:

$$\left| \frac{y_i - y_{\text{пр}}}{y_i} \right| * 100$$

Средняя ошибка аппроксимации находится как средняя арифметическая простая из индивидуальных ошибок:

$$\overline{A} = \frac{1}{n}\sum\left|\frac{y_i - y_{пр}}{y_i}\right| *100 = \text{2,65\%}$$

Вывод: величина средней ошибки аппроксимации показывает хорошее соответствие расчетных и фактических данных.

ПАРНАЯ ПОКАЗАТЕЛЬНАЯ РЕГРЕССИЯ

Лабораторная работа №1 **г**

Парная показательная регрессия и корреляция

Регрессия показательная в виде y=ab^x^

. Для оценки параметров модели линеаризуем модель путем логарифмирования: Ln y=Ln a+x*Ln b. Обозначим Ln a=A,Ln b=B. Применяем метод наименьших квадратов и получаем систему уравнений:

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

Расходы на покупку продовольственных товаров, % к общему объему расходов, **y**.

Среднемесячная заработная плата одного работающего, тыс. р., **x**.

Номер региона	x	y	Y=1/y	Y*x	x^2	ln(y)	x*ln(y)	$y_{покр}=a*b^x$	$(y-y_{покр})^2$	abs((y-$y_{покр}$)/y)*100
1	4,5	68,8	0,01453	0,0654	20,25	4,2312	19,040	67,346	2,11	2,11
2	5,9	58,3	0,01715	0,1012	34,81	4,0656	23,987	59,055	0,57	1,29
3	5,7	62,6	0,01597	0,0911	32,49	4,1368	23,580	60,174	5,89	3,88
4	7,2	52,1	0,01919	0,1382	51,84	3,9532	28,463	52,272	0,03	0,33
5	6,2	54,5	0,01835	0,1138	38,44	3,9982	24,789	57,415	8,50	5,35
6	6,0	57,1	0,01751	0,1051	36,00	4,0448	24,269	58,503	1,97	2,46
7	7,8	51,0	0,01961	0,1529	60,84	3,9318	30,668	49,410	2,53	3,12
Σ	43,3	404,4	0,12233	0,7676	274,7	28,362	174,80	404,176	21,595	18,538
Количество регионов	7									

Система нормальных уравнений составит:

$$\begin{cases} 7\ln a + 43.3\ln b = 28.362 \\ 43.3\ln a + 274.77\ln b = 174.8 \end{cases}$$

Решаем систему методом определителей

$$\Delta = \begin{vmatrix} 7 & 43.3 \\ 43.3 & 274.77 \end{vmatrix}$$

$$\Delta A = \begin{vmatrix} 28.362 & 43.3 \\ 174.8 & 274.77 \end{vmatrix}$$

$$\Delta B = \begin{vmatrix} 7 & 28.362 \\ 43.3 & 174.8 \end{vmatrix}$$

Δ=	47,80000
ΔA=	221,41646
ΔB=	-4,48573
A=ΔA/Δ	4,63214
B=ΔB/Δ	-0,09384

Выполнив потенцирование, получим: **a = $e^{4,63214}$; b=$e^{-0,09384}$**

$$y_{покр} = 102{,}73403 * 0{,}91042^x$$

102,734 =$e^{4,63214}$ =a **b=$e^{-0,09384}$= 0,91**

Показательное уравнение регрессии дополняется расчетом индекса корреляции: R

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

$\sum (y-\bar{y})^2 = n * \sigma_y^2 =$ **234,394**

R= 0,95282

$\sigma_x =$ 0,9876792

$\sigma_y =$ 5,7866137

$r_{yx} =$ 1,587121

} с листа линейной регрессии

Коэффициент детерминации составит:

$R^2 =$ 0,9078681

Вывод: вариации Y на 90,8% объясняются вариацией X. На долю прочих факторов, не учитываемых в регрессии, приходится 9,2%.

F-критерий Фишера будет равен:

$$F = \frac{R^2}{1 - R^2}(n-2) =$$ 49,3

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 регрессии статистически значимо.

Ошибки аппроксимации для каждого наблюдения определяются как:

$$\left|\frac{y_i - y_{\text{пр}}}{y_i}\right| * 100$$

Средняя ошибка аппроксимации находится как средняя арифметическая простая из индивидуальных ошибок:

$$\overline{A} = \frac{1}{n}\sum\left|\frac{y_i - y_{\text{пр}}}{y_i}\right| *100 =$$ 2,65%

Вывод: величина средней ошибки аппроксимации показывает хорошее соответствие расчетных и фактических данных.

СТАНДАРТНАЯ ФУНКЦИЯ 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 диапазон *Известные_значения_y*, 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 **=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:

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

	Функция ЛГРФПРИБЛ(y,x,1,1)		
Значение коэффициента b	0,910425	102,734	Значение коэффициента a
Среднеквадратическое отклонение b	0,013801	0,086448	Среднеквадратическое отклонение a
Коэффициент детерминации r^2	0,902419	0,036063	Среднеквадратическое отклонение y
F-статистика	46,2395	5	Число степеней свободы
Регрессионная сумма квадратов	0,060137	0,006503	Остаточная сумма квадратов

Решение системы методом определителей

Значение коэффициента b	0,910425	102,734	Значение коэффициента a
Среднеквадратическое отклонение b			Среднеквадратическое отклонение a
Коэффициент детерминации r^2	0,907868		Среднеквадратическое отклонение y
F-статистика	49,3		Число степеней свободы
Регрессионная сумма квадратов			Остаточная сумма квадратов

ПАРНАЯ РЕГРЕССИЯ В ВИДЕ РАВНОСТОРОННЕЙ ГИПЕРБОЛЫ

Лабораторная работа №1 **д**

Парная регрессия в виде равносторонней гиперболы и корреляция

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

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

Расходы на покупку продовольственных товаров, % к общему объему расходов, **y**.

Среднемесячная заработная плата одного работающего, тыс. р., **x**.

Номер региона	x	y	z=1/x	y*z	z^2	ln(y)	x*ln(y)	$y_{гипр}$=a+b/x	$(y-y_{гипр})^2$	abs((y-$y_{гипр}$)/y)*100
1	4,5	68,8	0,22222	15,2889	0,049	4,2312	19,040	69,035	0,06	0,34
2	5,9	58,3	0,16949	9,8814	0,029	4,0656	23,987	58,465	0,03	0,28
3	5,7	62,6	0,17544	10,9825	0,031	4,1368	23,580	59,658	8,66	4,70
4	7,2	52,1	0,13889	7,2361	0,019	3,9532	28,463	52,331	0,05	0,44
5	6,2	54,5	0,16129	8,7903	0,026	3,9982	24,789	56,822	5,39	4,26
6	6,0	57,1	0,16667	9,5167	0,028	4,0448	24,269	57,899	0,64	1,40
7	7,8	51,0	0,12821	6,5385	0,016	3,9318	30,668	50,190	0,66	1,59
Σ	43,3	404,4	1,1622	68,2343	0,198	28,362	174,80	404,400	15,479	13,018
Количество регионов	7									

Система нормальных уравнений составит:

$$\begin{cases} 7a + 1.1622\,b = 404.4 \\ 1.1622a + 0.19841b = 68.234 \end{cases}$$

Решаем систему методом определителей

$$\Delta = \begin{vmatrix} 7 & 1.1622 \\ 1.1622 & 0.198411 \end{vmatrix}$$

$$\Delta A = \begin{vmatrix} 404.4 & 1.1622 \\ 68.234 & 0.19841 \end{vmatrix}$$

$$\Delta B = \begin{vmatrix} 7 & 404.4 \\ 1.1622 & 68.234 \end{vmatrix}$$

Δ=	0,03814
Δα=	0,93403
Δβ=	7,64481
α=Δα/Δ	24,49081
β=Δβ/Δ	200,45059

Получаем уравнение регрессии в виде равносторонней гиперболы $y_{гипр}$

$y_{гипр}$=24,49081+200,45059/x

Уравнение регрессии в виде равносторонней гиперболы дополняется расчетом индекса корреляции: R

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

$\sum (y-\bar{y})^2 = n * \sigma_y^2 =$ **234,394**

R= 0,96642

$\sigma_x =$ 0,987679

$\sigma_y =$ 5,786614

$r_{yx} =$ 1,782744

с листа линейной регрессии

Коэффициент детерминации составит:

$R^2 =$ 0,933961

Вывод: вариации Y на 93,4% объясняются вариацией X. На долю прочих факторов, не учитываемых в регрессии, приходится 6,6%.

F-критерий Фишера будет равен:

$$F = \frac{R^2}{1 - R^2}(n-2) =$$ 70,7

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 регрессии статистически значимо.

Ошибки аппроксимации для каждого наблюдения определяются как:

$$\left|\frac{y_i - y_{пр}}{y_i}\right| * 100$$

Средняя ошибка аппроксимации находится как средняя арифметическая простая из индивидуальных ошибок:

$$\overline{A} = \frac{1}{n}\sum\left|\frac{y_i - y_{пр}}{y_i}\right| * 100 = 1{,}86\%$$

Вывод: величина средней ошибки аппроксимации показывает хорошее соответствие расчетных и фактических данных.

ПАРНАЯ РЕГРЕССИЯ В ВИДЕ ОБРАТНОЙ ФУНКЦИИ

Лабораторная работа №1 **е**

Парная регрессия в виде обратной функции и корреляция

Регрессия в виде обратной функции y=1/(a+b*x)

. Для оценки параметров модели линеаризуем модель путем замены: Y=1/y, тогда Y=a+b*x. Применяем метод наименьших квадратов и получаем систему уравнений:

$$\begin{cases} n\,a + b\sum z = \sum Y \\ a\sum x + b\sum x^2 = \sum Yx \end{cases}$$

Расходы на покупку продовольственных товаров, % к общему объему расходов, **у**.

Среднемесячная заработная плата одного работающего, тыс. р., **х**.

Номер региона	x	y	Y=1/y	Y*x	x^2	$y_{обрр}$=1/(a+b*x)	$(y-y_{обрр})^2$	abs((y-$y_{обрр}$)/y)*100
1	4,5	68,8	0,01453	0,0654	20,25	67,717	1,17	1,57
2	5,9	58,3	0,01715	0,1012	34,81	58,768	0,22	0,80
3	5,7	62,6	0,01597	0,0911	32,49	59,899	7,30	4,32
4	7,2	52,1	0,01919	0,1382	51,84	52,344	0,06	0,47
5	6,2	54,5	0,01835	0,1138	38,44	57,149	7,02	4,86
6	6,0	57,1	0,01751	0,1051	36,00	58,218	1,25	1,96
7	7,8	51,0	0,01961	0,1529	60,84	49,830	1,37	2,29
Σ	43,3	404,4	0,12233	0,7676	274,7	403,926	18,385	16,273
Количество регионов	7							

Система нормальных уравнений составит:

$$\begin{cases} 7a + 43.3b = 0.12233 \\ 43.3a + 274.67\,b = 0.7676 \end{cases}$$

Решаем систему методом определителей

$$\Delta = \begin{vmatrix} 7 & 43.3 \\ 43.3 & 274.67 \end{vmatrix}$$

$$\Delta A = \begin{vmatrix} 0.12233 & 43.3 \\ 0.7676 & 274.67 \end{vmatrix}$$

$$\Delta B = \begin{vmatrix} 7 & 0.12233 \\ 43.3 & 0.7676 \end{vmatrix}$$

Δ=	47,80000
Δα=	0,36035
Δβ=	0,07678
α=Δα/Δ	0,00754
β=Δβ/Δ	0,00161

Получаем уравнение регрессии в виде обратной функции $y_{обрр}$

$y_{обрр}$=1/(0,00754+0,001606x)

Уравнение регрессии в виде обратной функции дополняется расчетом индекса корреляции: R

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

$$\sum (y-\bar{y})^2 = n*\sigma_y^2 = 234,3943$$

R= 0,959981

$\sigma_x =$ 0,987679

$\sigma_y =$ 5,786614

с листа линейной регрессии

Коэффициент детерминации составит:

$R^2 =$ 0,921563

Вывод: вариации Y на 92,2% объясняются вариацией X. На долю прочих факторов, не учитываемых в регрессии, приходится 7,8%.

F-критерий Фишера будет равен:

$$F = \frac{R^2}{1 - R^2}(n - 2) = 58{,}7$$

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 регрессии статистически значимо.

Ошибки аппроксимации для каждого наблюдения определяются как:

$$\left|\frac{y_i - y_{\text{пр}}}{y_i}\right| * 100$$

Средняя ошибка аппроксимации находится как средняя арифметическая простая из индивидуальных ошибок:

$$\overline{A} = \frac{1}{n}\sum\left|\frac{y_i - y_{\text{пр}}}{y_i}\right| * 100 = 2{,}32\%$$

Вывод: величина средней ошибки аппроксимации показывает хорошее соответствие расчетных и фактических данных.

ВЫБОР НАИЛУЧШЕЙ МОДЕЛИ

Парная регрессия и корреляция						
№	Название регрессии	Уравнение регрессии	Коэффициент детерминации R2	F-критерий Фишера	Статистическая значимость уравнения	Средняя ошибка аппроксимации, %
1 а	линейная	y=a+b*x	0,8876	39,50	Статистически значимо	2,99%
1 б	степенная	y=axb	0,9306	67,06	Статистически значимо	2,02%
1 в	экспоненциальная	y=aebx	0,9079	49,27	Статистически значимо	2,65%
1 г	показательная	y=abx	0,9079	49,27	Статистически значимо	2,65%
1 д	в виде равносторонней гиперболы	y=a+b/x	0,9340	70,71	Статистически значимо	1,86%
1 е	в виде обратной функции	y=1/(a+b*x)	0,9216	58,75	Статистически значимо	2,32%
	ВЫВОДЫ					
Регрессия	в виде равносторонней гиперболы	Имеет максим R2	0,9340			
	=ИНДЕКС(B4:B9;ПОИСКПОЗ(D11;D4:D9;0))		=МАКС(D4:D9)			
Регрессия	в виде равносторонней гиперболы	Имеет миним ошибку аппроксимации				1,86%
	=ИНДЕКС(B4:B9;ПОИСКПОЗ(F13;F4:F9;0))					=МИН(F4:F9)

Лабораторная работа №2

МНОЖЕСТВЕННАЯ РЕГРЕССИЯ И КОРРЕЛЯЦИЯ

Множественная регрессия широко используется в решении проблем спроса, доходности акций, при изучении функции издержек производства, в макроэкономических расчётах и целом ряде 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.

Линейное уравнение множественной регрессии $\boldsymbol{y}$ от $\boldsymbol{x_1}$ и $\boldsymbol{x_2}$ имеет вид

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

Поскольку одним из условий построения уравнения множественной регрессии является независимость действия факторов, то сначала необходимо:

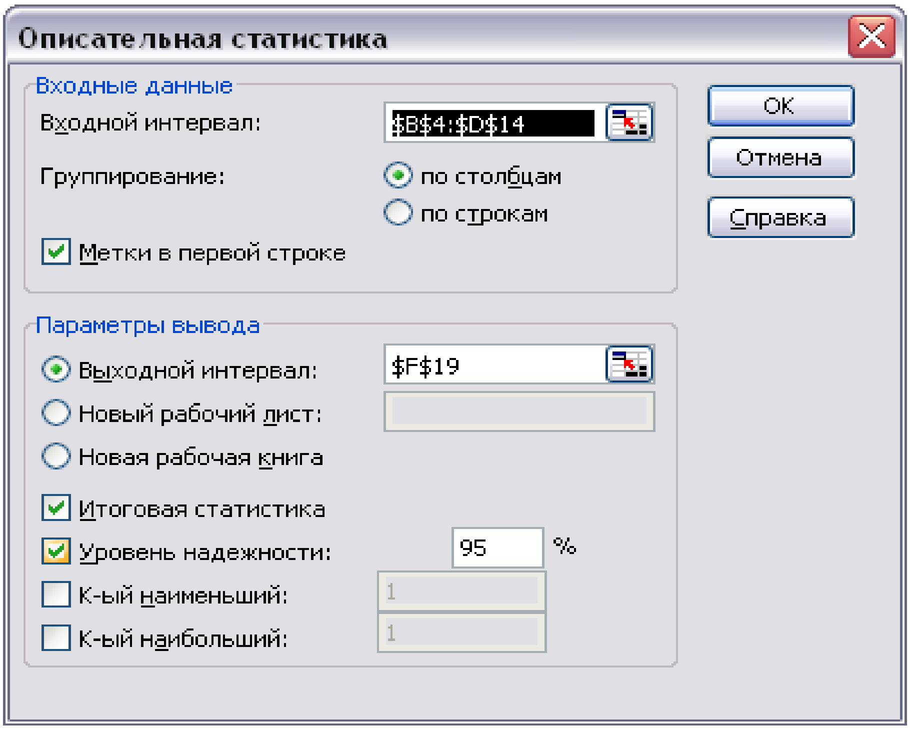
1. Оценить показатели вариации каждого признака и сделать вывод о возможностях применения метода наименьших квадратов (МНК) для их изучения.
2. Проанализировать линейные коэффициенты парной и частной корреляции.
3. Составить уравнение множественной регрессии, оценить значимость его параметров, пояснить их экономический смысл.
4. С помощью $\boldsymbol{F}$-критерия Фишера оценить статистическую надёжность уравнения регрессии и коэффициента корреляции множественной регрессии ($\boldsymbol{R^2_{yx1x2}}$). Сравнить значения скорректированного и нескорректированного линейных коэффициентов множественной детерминации.
5. С помощью $\boldsymbol{F}$-критерия Фишера оценить целесообразность включения в уравнение множественной регрессии фактора $\boldsymbol{x_1}$ после $\boldsymbol{x_2}$ и фактора $\boldsymbol{x_2}$ после $\boldsymbol{x_1}$.
6. Рассчитать средние частные коэффициенты эластичности и дать на их основе сравнительную оценку силы влияния факторов на результат.

ОЦЕНКА ПОКАЗАТЕЛЕЙ ВАРЬИРОВАНИЯ ПРИЗНАКОВ

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

Для этого выполните следующие шаги:

- Введите исходные данные.
- Выполните команду меню *Сервис, Анализ данных, Описательная статистика.*



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

Исходный диапазон

	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

Результаты вычисления

	Сервис-Анализ данных-Описательная статистика				
y1 объясняемая зависимая переменная		*x1 объясняющая независимая переменная*		*x2 объясняющая независимая переменная*	
Среднее	7,6	Среднее	4,6	Среднее	17
Стандартная ошибка	0,305505	Стандартная ошибка	0,313404	Стандартная ошибка	1,043498
Медиана	7	Медиана	4,2	Медиана	18
Мода	7	Мода	3,9	Мода	20
Стандартное отклонение	0,9660918	Стандартное отклонение	0,991071	Стандартное отклонение	3,29983
Дисперсия выборки	0,9333333	Дисперсия выборки	0,982222	Дисперсия выборки	10,88889
Эксцесс	4,1873178	Эксцесс	1,48694	Эксцесс	0,76696
Асимметричность	1,9592933	Асимметричность	1,335453	Асимметричность	-1,09004
Интервал	3	Интервал	3,1	Интервал	10
Минимум	7	Минимум	3,7	Минимум	10
Максимум	10	Максимум	6,8	Максимум	20
Сумма	76	Сумма	46	Сумма	170
Счет	10	Счет	10	Счет	10
Уровень надежности(95,0%)	0,6911004	Уровень надежности(95,0%)	0,70897	Уровень надежности(95,0%)	2,36056

Определяем уровень варьирования признаков:

$\upsilon_y = \dfrac{\sigma_y}{\bar{y}} 100\%$	$\upsilon_y =$ 12,71%
$\upsilon_{x1} = \dfrac{\sigma_{x1}}{\bar{x}_1} 100\%$	$\upsilon_{x1} =$ 21,55%
$\upsilon_{x2} = \dfrac{\sigma_{x2}}{\bar{x}_2} 100\%$	$\upsilon_{x2} =$ 19,41%

где $\boldsymbol{\sigma_{x1}}$ – стандартное отклонение по $\boldsymbol{x_1}$;

$\boldsymbol{\sigma_{x2}}$ – стандартное отклонение по $\boldsymbol{x_2}$;

$\boldsymbol{\sigma_y}$ – стандартное отклонение по $\boldsymbol{y}$;

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

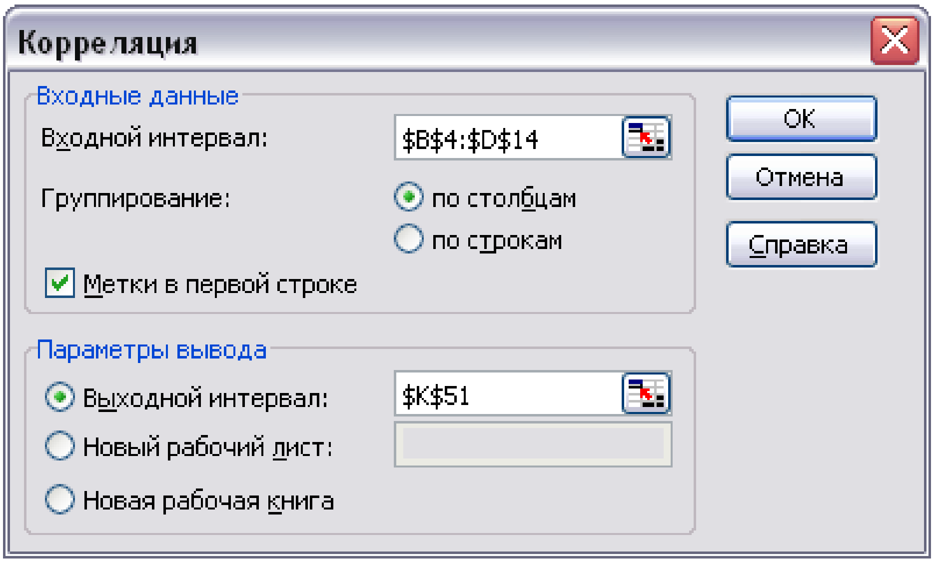
Приходим к выводу об умеренном уровне варьирования признаков, не превышающем 35% (т.е. совокупность данных по регионам однородна), и возможности применения метода наименьших квадратов (МНК) для их изучения.

АНАЛИЗ ЛИНЕЙНЫХ КОЭФФИЦИЕНТОВ 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.**

Выполните команду меню **Сервис, Анализ данных, Корреляция** и заполните диалоговое окно



Сервис-Анализ данных-Корреляция			
	y1 объясняе - мая 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
	это коэффициенты парной корреляции		

РАСЧЁТ КОЭФФИЦИЕНТОВ ЧАСТНОЙ КОРРЕЛЯЦИИ

В ППП EXCEL нет специального инструмента для расчёта линейных коэффициентов частной корреляции. Их можно рассчитать по рекуррентной формуле через коэффициенты парной корреляции.

$$r_{yx1/x2} = \frac{r_{yx1} - r_{yx2} \cdot r_{x1x2}}{\sqrt{(1 - r_{yx2}^2) \cdot (1 - r_{x1x2}^2)}} = 0{,}86877671,$$

$$r_{yx2/x1} = \frac{r_{yx2} - r_{yx1} \cdot r_{x1x2}}{\sqrt{(1 - r_{yx1}^2) \cdot (1 - r_{x1x2}^2)}} = -\ 0{,}04968,$$

$$r_{x1x2/y} = \frac{r_{x2x1} - r_{yx2} \cdot r_{yx1}}{\sqrt{(1 - r_{yx2}^2) \cdot (1 - r_{yx1}^2)}} = 0{,}37083798.$$

Пояснения. $r_{y\,x2/x1}$=т.е. x1 фиксируем

Вывод:

Из анализа коэффициентов парной корреляции следует, что значение r_{yx1} = 0,9168 указывает на тесную связь между ***y*** и x_1, а значение r_{x2x1} = 0,6625 говорит о тесной связи между x_2 и x_1, при этом r_{yx2} = 0,0,5925 < r_{x2x1}, т.е. x_2 можно пренебречь.

Из анализа частных коэффициентов множественной корреляции следует, что значение $r_{yx1/x2}$ = 0,8688 (x_2 фиксируем) указывает на тесную связь между ***y*** и x_1, а значение $r_{yx2/x1}$ = - 0,04968 (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.

ВЫЧИСЛЕНИЕ ПАРАМЕТРОВ ЛИНЕЙНОГО УРАВНЕНИЯ МНОЖЕСТВЕННОЙ РЕГРЕССИИ МЕТОДОМ СТАНДАРТИЗАЦИИ ПЕРЕМЕННЫХ

Стандартизованные частные коэффициенты регрессии – бета-коэффициенты показывают, на какую часть своего среднеквадратического отклонения изменится признак результат ***y*** с увеличением соответствующего фактора $\boldsymbol{x_i}$ на величину своего среднеквадратического отклонения при неизменном влиянии прочих факторов модели.

Для вычисления коэффициентов множественной регрессии применим метод стандартизации переменных и построим искомое уравнения в стандартизованном масштабе:

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

Расчёт **β**-коэффициентов выполняется по формулам:

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

$$\beta_2 = \frac{r_{yx2} - r_{yx1} \cdot r_{x1x2}}{1 - r_{x1x2}^2}$$

В результате получаем **β**-коэффициенты: $\boldsymbol{\beta_1}$ **=** 0,9343, $\boldsymbol{\beta_2}$= - 0,0265.

Уравнение в стандартизованном масштабе:

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

Для построения уравнения в естественной форме рассчитаем $\mathbf{b_1}$ и $\mathbf{b_2}$, 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: $\mathbf{b_1}$ **=** 0,9108, $\mathbf{b_2}$ = - 0,007756.

Значение **a** определим из соотношения

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

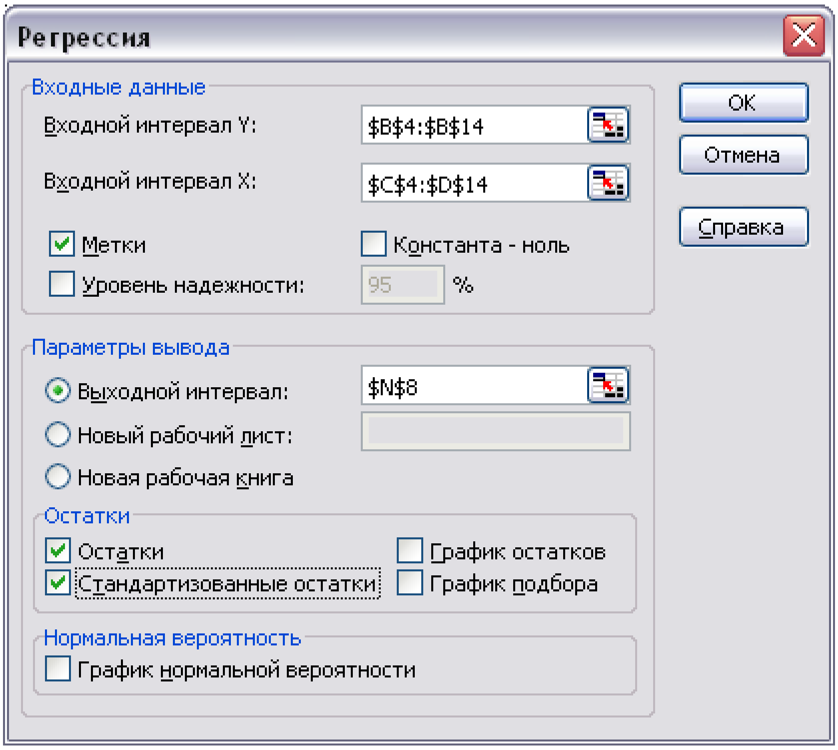
$$a = 3{,}5422923.$$

Уравнение множественной регрессии:

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

Второй способ получения оценок параметров уравнения множественной регрессии – с помощью инструмента EXCEL **Регрессия.**

Выполните команду меню **Сервис, Анализ данных, Регрессия**. Заполните диалоговое окно как показано на рисунке:



ВЫВОД ИТОГОВ

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

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			

	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

Уравнение регрессии:

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

Результаты анализа:

- Значения случайных ошибок параметров **a, b_1 и 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 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 **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.

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

Расчёт линейного коэффициента множественной корреляции:

$$R_{yx1x2} = \sqrt{r_{yx1} \cdot \beta_1 + r_{yx2} \cdot \beta_2}\ .$$

Получаем R_{yx1x2}=0,9170 (сравните с результатами функции **Регрессии**). Зависимость **y** от x_1 и x_2 характеризуется как тесная.

ОЦЕНКА КОЭФФИЦИЕНТОВ УРАВНЕНИЯ МНОЖЕСТВЕННОЙ ЛИНЕЙНОЙ РЕГРЕССИИ МЕТОДОМ НАИМЕНЬШИХ КВАДРАТОВ

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

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

Значения x_{ik} и y_k известны – это данные наблюдения. Переменными данной функции являются оценки параметров $\widehat{b}_0$ и $\widehat{b}_i$.

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

$$\frac{dS}{db_0} = 0,$$

$$\frac{dS}{db_1} = 0,$$

$$\frac{dS}{db_2} = 0.$$

В результате получим систему линейных уравнений

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

Подставим известные значения и получим следующую систему линейных уравнений:

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

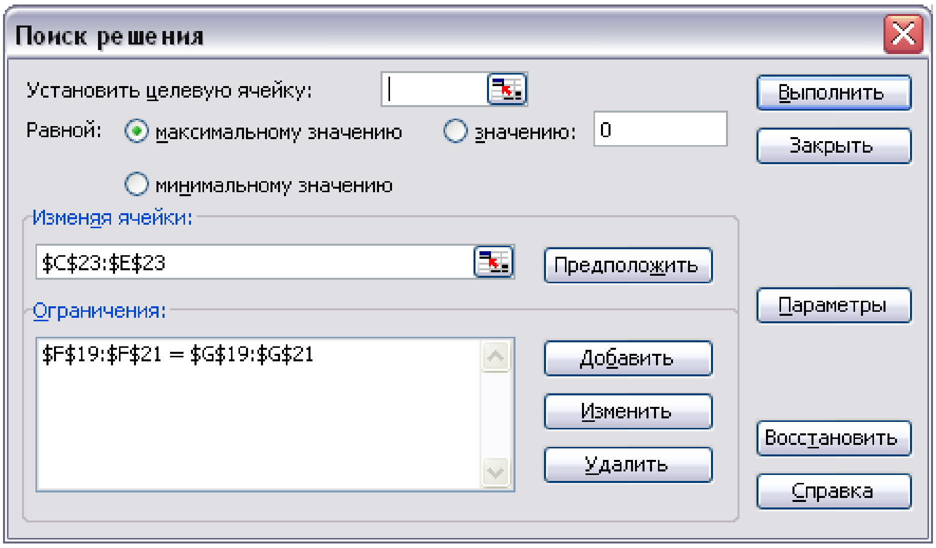
Решаем систему, применяя инструмент ППП EXCEL **Поиск решения.**

	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		

В ячейки с F19 по F21 добавить формулы:

В ячейках	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 выполнить команду ***Сервис, Поиск 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		

Таким образом, получаем уравнение множественной регрессии:

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

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

РАСЧЁТ ЧАСТНЫХ КОЭФФИЦИЕНТОВ ЭЛАСТИЧНОСТИ

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

$$Э_i = b_i \cdot \frac{\overline{x}_i}{\overline{y}}.$$

После расчёта в ППП EXCEL получаем **$Э_1$** = 0,5513, **$Э_1$** = **-** 0,0173.

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 ***F***-критерия в зависимости от уровня значимости ***α*** (обычно его берут равным 0,05) и двух чисел степеней свободы $k_1 = m - 1$ и $k_2 = n - m$, где ***m*** – количество пар оцениваемых параметров в уравнении регрессии, а ***n*** – число наблюдений. $F_{табл} = 5{,}32$.

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

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

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

После расчётов в ППП EXCEL 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 в EXCEL получаем $F_{x2факт}$ = -0,00866. Низкое значение $F_{x2факт}$ свидетельствует о статистической незначимости прироста парного коэффициента корреляции r^2_{yx1} 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 .

Лабораторная работа № 3
ВРЕМЕННЫЕ РЯДЫ В ЭКОНОМЕТРИЧЕСКИХ ИССЛЕДОВАНИЯХ

РАСЧЕТ ЛИНЕЙНОГО ТРЕНДА

Рассмотрим в качестве примера динамику расходов на покупку продовольственных товаров по годам.

Год	x	y
2000	4,5	68,8
2001	5,9	58,3
2002	5,7	62,6
2003	7,2	52,1
2004	6,2	54,5
2005	6,0	57,1
2006	7,8	51,0

Где ***x*** – среднемесячная заработная плата одного работающего, тыс.р.,

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

Задание:

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

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

3. Подобрать различные виды трендов, построенных графически.

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

5. Выбрать наилучший среди трендов, построенных графически, по значению коэффициента детерминации $\boldsymbol{R^2}$.

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

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

Параметры линейного приближения по методу наименьших квадратов

Функция ЛИНЕЙН(y,x,1,1) y=a+b*x

Значение коэффициента b	**-5,51987448**	**91,915795**	**Значение коэффициента a**
Среднеквадратическое отклонение b	0,87823787	5,50134338	Среднеквадратическое отклонение a
Коэффициент детерминации r^2	**0,88764894**	2,2949704	Среднеквадратическое отклонение y
F-статистика	39,5033644	5	Число степеней свободы
Регрессионная сумма квадратов	208,05984	26,3344456	Остаточная сумма квадратов

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

Коэффициент детерминации r^2 = 0,88765,

$$\mathbf{y_{лин.\,тр} = 91,9158 - 5,51987*x.}$$

СТАНДАРТНАЯ ФУНКЦИЯ ЛИНЕЙН(y,x,1,1).

Параметры линейного приближения по методу наименьших квадратов можно получить, используя стандартную функцию **ЛИНЕЙН(y,x,1,1).**

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

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

РАСЧЕТ ЛОГАРИФМИЧЕСКОГО ТРЕНДА

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

Параметры приближения в виде показательной функции по методу наименьших квадратов

Функция ЛГРФПРИБЛ(y,x,1,1) **$y=ab^x$**			
Значение коэффициента b	**0,91042494**	**102,734033**	**Значение коэффициента a**
Среднеквадратическое отклонение b	0,01380064	0,08644816	Среднеквадратическое отклонение a
Коэффициент детерминации r^2	**0,90241903**	0,03606318	Среднеквадратическое отклонение y
F-статистика	46,2394973	5	Число степеней свободы
Регрессионная сумма квадратов	0,06013693	0,00650277	Остаточная сумма квадратов

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

$$\mathbf{y_{\text{логар тр}} = 102{,}73403 * 0{,}91042x}$$

Коэффициент детерминации $r^2 = 0{,}902419$

СТАНДАРТНАЯ ФУНКЦИЯ 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 диапазон *Известные_значения_у*, содержащий числовые значения массива объясняемой (зависимой) переменной у.

Известные_значения_х. - диапазон, содержащий числовые значения массива объясняющей (независимой) переменной х.

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.

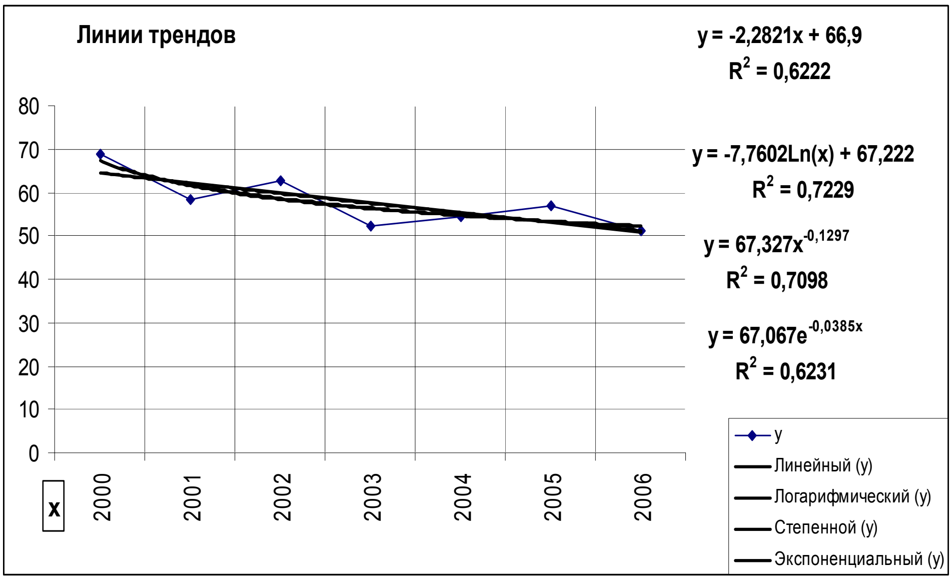
Выделить группу ячеек размером 5 строк и 2 столбца с ячейкой в верхнем левом углу, содержащей формулу **=ЛГРФПРИБЛ(y,x,1,1)**, затем сначала нажать на клавиатуре клавишу F2, потом – комбинацию клавиш <CTRL>+<SHIFT>+<ENTER> для раскрытия всей таблицы дополнительной информации по регрессионному анализу:

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

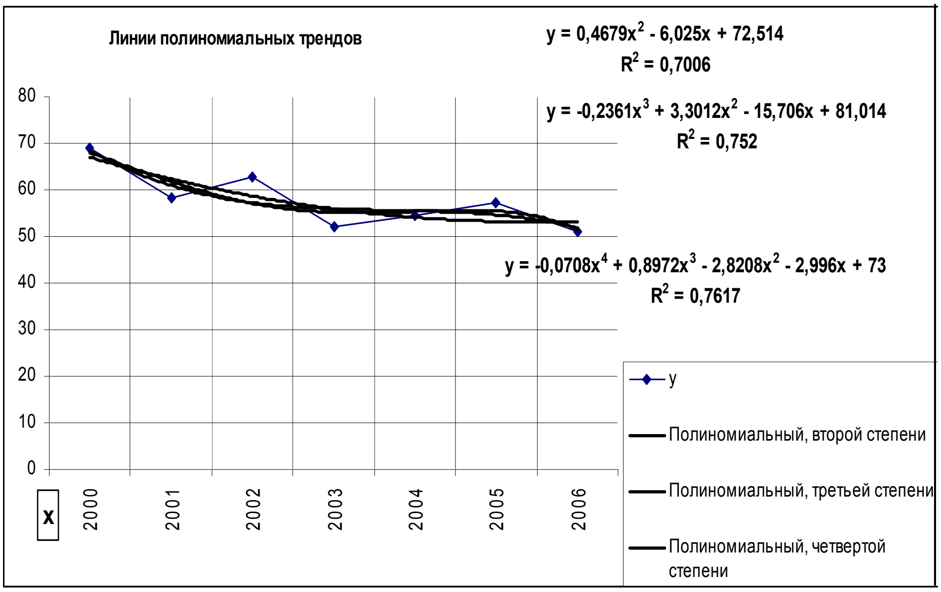
ПОДБОР ТРЕНДОВ, ПОСТРОЕННЫХ ГРАФИЧЕСКИ

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.

Ниже представлены графики линий трендов:



Графики линейного, логарифмического, степенного и экспоненциального трендов.



Графики полиномиальных (второй, третьей и четвертой степеней) трендов.

ВЫБОР НАИЛУЧШЕГО ТРЕНДА

Результаты построения трендов

№	Тренд	Уравнение регрессии	Коэффициент детерминации (Величина достоверности аппроксимации R^2)
1	2	3	4
1	Линейный	y = -2,2821x + 66,9	0,6222
2	ЛИНЕЙН(y,x,1,1)	$y_{лин\ тр}$=91,9158 - 5,51987*x	0,8876
3	Логарифмический (показательный)	y = -7,7602Ln(x) + 67,222	0,7229

1	2	3	4
4	ЛГРФПРИБЛ(y,x,1,1)	$y_{логар\ тр}$=102,73403 * $0,91042^x$	0,9024
5	Степенной	$y = 67{,}327x^{-0{,}1297}$	0,7098
6	Экспоненциальный	$y = 67{,}067e^{-0{,}0385x}$	0,6231

Результаты построения полиномиальных трендов

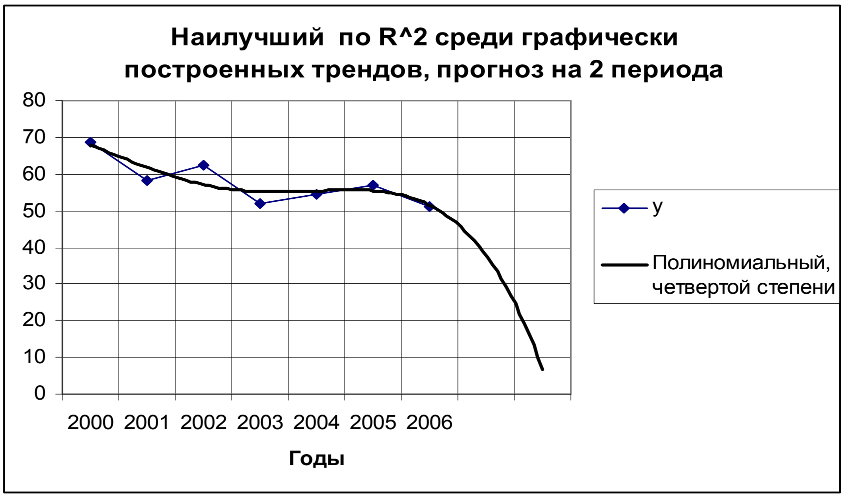
№	Тренд	Уравнение регрессии		Коэффициент детерминации (Величина достоверности аппроксимации R^2)
7	Полиномиальный, второй степени	$y = 0{,}4679x^2 - 6{,}025x + 72{,}514$		0,7006
8	Полиномиальный, третьей степени	$y = -0{,}2361x^3 + 3{,}3012x^2 - 15{,}706x + 81{,}014$		0,7520
9	Полиномиальный, четвертой степени	$y = -0{,}0708x^4 + 0{,}8972x^3 - 2{,}8208x^2 - 2{,}996x + 73$		0,7617
Тренд	**ЛГРФПРИБЛ(y,x,1,1)**		**Имеет максимум R^2**	**0,9024**
Среди трендов, найденных только графически	**Полиномиальный, четвертой степени**		**Имеет максимум R^2**	**0,7617**

Ниже приведены формулы для получения результатов сравнения.

	C	D	E	F	G
5	**№**	**Тренд**	**Уравнение регрессии**		**Коэффициент детерминации (Величина достоверности аппроксимации R^2)**
6	1	Линейный	$y = -2{,}2821x + 66{,}9$		0,6222
7	2	ЛИНЕЙН(y,x,1,1)	$y_{лин\ тр}=91{,}9158 - 5{,}51987*x$		0,8876
8	3	Логарифмический (показательный)	$y = -7{,}7602Ln(x) + 67{,}222$		0,7229
9	4	ЛГРФПРИБЛ(y,x,1,1)	$y_{логар\ тр}=102{,}73403 * 0{,}91042x$		0,9024
10	5	Степенной	$y = 67{,}327x^{-0{,}1297}$		0,7098
11	6	Экспоненциальный	$y = 67{,}067e^{-0{,}0385x}$		0,6231
12	7	Полиномиальный, второй степени	$y = 0{,}4679x^2 - 6{,}025x + 72{,}514$		0,7006
13	8	Полиномиальный, третьей степени	$y = -0{,}2361x^3 + 3{,}3012x^2 - 15{,}706x + 81{,}014$		0,7520
14	9	Полиномиальный, четвертой степени	$y = -0{,}0708x^4 + 0{,}8972x^3 - 2{,}8208x^2 - 2{,}996x + 73$		0,7617
15	**Тренд**	**ЛГРФПРИБЛ(y,x,1,1)**		**Имеет максимум R^2**	**0,9024**
16		=ИНДЕКС(B4:B9;ПОИСКПОЗ(D11;D4:D9;0))			=МАКС(G6:G14)
17	**Среди трендов, найденных только графически**	**Полиномиальный, четвертой степени**		**Имеет максимум R^2**	**0,7617**

ПРОГНОЗ НЕСКОЛЬКИХ ПЕРИОДОВ ВПЕРЕД

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



Лабораторная работа №4
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.

Именно поэтому в экономических исследованиях важное место заняла проблема описания структуры связей между переменными системой так называемых одновременных уравнений или структурных уравнений. Структурная форма модели описывает реальное экономическое явление или процесс.

Простейшая структурная форма модели имеет вид

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

где ***y*** – эндогенные переменные;

x – экзогенные переменные.

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

$$\begin{cases} y_1 = \delta_{11} \cdot x_1 + \delta_{12} \cdot x_2 + \ldots + \delta_{1m} \cdot x_m + \varepsilon_1, \\ y_2 = \delta_{21} \cdot x_1 + \delta_{22} \cdot x_2 + \ldots + \delta_{2m} \cdot x_m + \varepsilon_2, \\ \ldots, \\ y_k = \delta_{k1} \cdot x_1 + \delta_{k2} \cdot x_2 + \ldots + \delta_{km} \cdot x_m + \varepsilon_k. \end{cases}$$

Приведённая форма уравнений – система независимых уравнений, в которой все текущие эндогенные переменные модели выражены через предопределённые переменные модели.

Параметры δ_{ij} называются приведёнными коэффициентами. 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.

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 сверхидентифицировано.
2. Если $M - m = k - 1$ и ранг матрицы ***A*** равен $K - 1$, то уравнение точно идентифицировано.
3. Если $M - m \geq k - 1$ и ранг матрицы ***A*** меньше $K - 1$, то уравнение неидентифицировано.
4. Если $M - m < k - 1$, то уравнение неидентифицировано.

Для определения параметров такой системы применяется косвенный метод наименьших квадратов. Рассмотрим КМНК для простейшей идентифицируемой эконометрической модели с двумя эндогенными и двумя экзогенными переменными:

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

Структурная форма модели преобразуется в приведённую форму, в которой коэффициенты при ***x*** определяются методом наименьших квадратов:

$$\begin{cases} y_1 = \delta_{11} \cdot x_1 + \delta_{12} \cdot x_2, \\ y_2 = \delta_{21} \cdot x_1 + \delta_{22} \cdot x_2. \end{cases}$$

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

$$\begin{cases} \sum y_1 x_1 = \delta_{11} \cdot \sum x_1^2 + \delta_{12} \cdot \sum x_1 x_2, \\ \sum y_1 x_2 = \delta_{11} \cdot \sum x_1 x_2 + \delta_{12} \cdot \sum x_2^2. \end{cases}$$

Пример. Имеются данные за 5 лет.

Номер года	Годовое потребление продукта, у	Оптовая цена за кг, у2	Доход на душу населения, х1	Расходы по обработке продукта, х2
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

По имеющимся данным построить систему эконометрических уравнений вида

$$\begin{cases} y_1 = f(y_2, x_1), \\ y_2 = f(y_1, x_2). \end{cases}$$

- ✓ Провести идентификацию модели.
- ✓ Рассчитать соответствующие структурные коэффициенты.

Система одновременных уравнений с двумя эндогенными и двумя экзогенными переменными имеет вид

$$\begin{cases} y_1 = b_{12} \cdot y_2 + a_{11} \cdot x_1 + e_1, \\ y_2 = b_{21} \cdot y_1 + a_{22} \cdot x_2 + e_2. \end{cases}$$

ИДЕНТИФИКАЦИЯ МОДЕЛИ

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 идентифицировано.

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 $K - 1 = 1$, следовательно первое уравнение точно идентифицировано.

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 идентифицировано.

Д: во втором уравнении отсутствует x_1 . $M = 2$, $m = 1$, $K = 2$, $k = 2$. Проверяем: $M - m = 1$, $k - 1 = 1$, т.е. выполняется правило 2. Ранг матрицы $A = a_{11}$, определитель матрицы не равен нулю, ранг матрицы равен $K - 1 = 1$, следовательно второе уравнение точно идентифицировано.

ОЦЕНКА ПАРАМЕТРОВ СИСТЕМЫ

При решении системы

$$\begin{cases} \sum y_1 x_1 = \delta_{11} \cdot \sum x_1^2 + \delta_{12} \cdot \sum x_1 x_2, \\ \sum y_1 x_2 = \delta_{11} \cdot \sum x_1 x_2 + \delta_{12} \cdot \sum x_2^2 \end{cases}$$

выразим x и y через отклонения от средних уровней, и тогда матрица исходных данных будет иметь вид:

Среднее: 63 4,4 1500 57

y1	y2	x1	x2
-3	0,6	-200	3
-1	-0,4	-200	-1
2	-0,2	0	-1
-1	0,6	100	6
3	-0,6	300	-7

Соответственно, значения сумм:

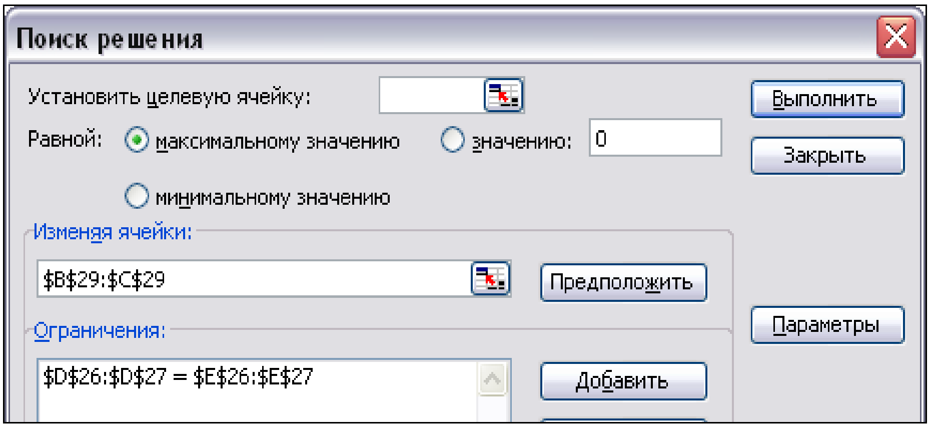
y1*x1	y1*x2	x1^2	x2^2	x1*x2
600	-9	40000	9	-600
200	1	40000	1	200
0	-2	0	1	0
-100	-6	10000	36	600
900	-21	90000	49	-2100
1600	-37	180000	96	-1900

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.**

	B	C	D	E
19				
20				
21				
22				
23				
24				
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		



Получаем:

$$\delta_{11} = 0{,}00609,$$

$$\delta_{12} = -0{,}26481.$$

Тогда $y_1 = 0{,}00609 \cdot x_1 - 0{,}26481 \cdot x_2$

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

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

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 решаем систему уравнений.

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

$$y_2 = 0{,}00029 \cdot x_1 + 0{,}11207 \cdot x_2 .$$

Приведённая форма модели примет вид

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

Из чего определяем коэффициенты структурной модели:

$$\begin{cases} y_1 = 0{,}00609 \cdot x_1 - 0{,}26481 \cdot x_2, \\ x_2 = \dfrac{y_2 - 0{,}00029 x_1}{0{,}11207}; \end{cases}$$

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

Аналогично

$$\begin{cases} y_2 = 0{,}00029 \cdot x_1 + 0{,}11207 \cdot x_2, \\ x_1 = \dfrac{y_1 + 0{,}26481 \cdot x_2}{0{,}00609}. \end{cases}$$

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

$$\begin{cases} y_1 = -2{,}36290 \cdot y_2 + 0{,}00678 \cdot x_1 + \varepsilon_1, \\ y_2 = 0{,}04762 \cdot y_1 + 0{,}12468 \cdot x_2 + \varepsilon_2. \end{cases}$$

СПИСОК ЛИТЕРАТУРЫ

1. Практикум по эконометрике / под ред. Н.И. Елисеевой. – М.: Финансы и статистика, 2006.

2. Эконометрика / под ред. Н.И. Елисеевой. – М.: Финансы и статистика, 2005.

3. Кремер, Н.Ш. Эконометрика / Н.Ш. Кремер, Б.А. Путко; под ред. Н.Ш. Кремера. – М.: ЮНИТИ, 2007.

4. Баклушина, О.А. Краткий курс по эконометрике / О.А. Баклушина. – М.: Окей-книга, 2007.

5. Айвазян, С.А. Прикладная статистика. Основы эконометрики: учебник для вузов: В 2 т. Т. 2. Основы эконометрики / С.А. Айвазян. – М.: ЮНИТИ-ДАНА, 2001. – 432 с.

6. Доугерти, К. Введение в эконометрику / К. Доугерти. – М.: ИНФРА-М, 1997. – 402 с.

7. Кремер, Н.Ш. Эконометрика / Н.Ш. Кремер, Б.А. Путко. – М.: ЮНИТИ, 2002. 311 с.

СОДЕРЖАНИЕ

Лабораторная работа №1. ПАРНАЯ РЕГРЕССИЯ 3
ПАРНАЯ ЛИНЕЙНАЯ РЕГРЕССИЯ 3
ПАРНАЯ СТЕПЕННАЯ РЕГРЕССИЯ 8
ПАРНАЯ ЭКСПОНЕНЦИАЛЬНАЯ РЕГРЕССИЯ 11
ПАРНАЯ ПОКАЗАТЕЛЬНАЯ РЕГРЕССИЯ 14
ПАРНАЯ РЕГРЕССИЯ В ВИДЕ РАВНОСТОРОННЕЙ ГИПЕРБОЛЫ 18
ПАРНАЯ РЕГРЕССИЯ В ВИДЕ ОБРАТНОЙ ФУНКЦИИ 21
ВЫБОР НАИЛУЧШЕЙ МОДЕЛИ 24
Лабораторная работа №2. МНОЖЕСТВЕННАЯ РЕГРЕССИЯ И КОРРЕЛЯЦИЯ 25
ОЦЕНКА ПОКАЗАТЕЛЕЙ ВАРЬИРОВАНИЯ ПРИЗНАКОВ 26
АНАЛИЗ ЛИНЕЙНЫХ КОЭФФИЦИЕНТОВ ПАРНОЙ И ЧАСТНОЙ КОРРЕЛЯЦИИ 28
РАСЧЁТ КОЭФФИЦИЕНТОВ ПАРНОЙ И ЧАСТНОЙ КОРРЕЛЯЦИИ..30
ВЫЧИСЛЕНИЕ ПАРАМЕТРОВ ЛИНЕЙНОГО УРАВНЕНИЯ МНОЖЕСТВЕННОЙ РЕГРЕССИИ МЕТОДОМ СТАНДАРТИЗАЦИИ ПЕРЕМЕННЫХ 31
ОЦЕНКА КОЭФФИЦИЕНТОВ УРАВНЕНИЯ МНОЖЕСТВЕННОЙ ЛИНЕЙНОЙ РЕГРЕССИИ МЕТОДОМ НАИМЕНЬШИХ КВАДРАТОВ 34
РАСЧЁТ ЧАСТНЫХ КОЭФФИЦИЕНТОВ ЭЛАСТИЧНОСТИ. 37
РАСЧЁТ ОБЩЕГО И ЧАСТНОГО F-КРИТЕРИЯ ФИШЕРА. 37
Лабораторная работа №3. ВРЕМЕННЫЕ РЯДЫ В ЭКОНОМЕТРИЧЕСКИХ ИССЛЕДОВАНИЯХ 39
РАСЧЕТ ЛИНЕЙНОГО ТРЕНДА 39
РАСЧЕТ ЛОГАРИФМИЧЕСКОГО ТРЕНДА 41
ПОДБОР ТРЕНДОВ, ПОСТРОЕННЫХ ГРАФИЧЕСКИ 43
ВЫБОР НАИЛУЧШЕГО ТРЕНДА 44
ПРОГНОЗ НЕСКОЛЬКИХ ПЕРИОДОВ ВПЕРЕД 47
Лабораторная работа №4. СИСТЕМА ЭКОНОМЕТРИЧЕСКИХ УРАВНЕНИЙ 48
ПРАВИЛА ИДЕНТИФИКАЦИИ МОДЕЛИ. 49
ИДЕНТИФИКАЦИЯ МОДЕЛИ. 51
ОЦЕНКА ПАРАМЕТРОВ СИСТЕМЫ 51
СПИСОК ЛИТЕРАТУРЫ 55
ПРИЛОЖЕНИЯ 57

ПРИЛОЖЕНИЯ

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

Распределение Стьюдента(t-распределение)

k- число степеней свободы, α – уровень значимости

k \ α	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
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

Учебное издание

ЭКОНОМЕТРИКА

Методические указания
к лабораторному практикуму

Составители: ***Озерная Светлана Алексеевна***
Макаренко Татьяна Васильевна

Редактор Л. Я. Ч е г о д а е в а
Компьютерная верстка Т. Е. П о л о в н е в а

Подписано в печать 27.11.08. Формат 60х84 1/16.
Бумага офсетная. Печать офсетная.
Усл. печ. л. 3,75.
Тираж 150 экз. Заказ . Арт. С-82/2008.

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

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