Министерство путей сообщения Российской Федерации

ДАЛЬНЕВОСТОЧНЫЙ ГОСУДАРСТВЕННЫЙ
УНИВЕРСИТЕТ ПУТЕЙ СООБЩЕНИЯ

Кафедра: «Бухгалтерский учет и аудит»

Л.Н. Корзова

ЭКОНОМЕТРИКА

Методическое пособие на выполнение контрольной работы
для студентов экономических специальностей ИИФО

Хабаровск
Издательство ДВГУПС
2002

УДК
ББК
К

Рецензент:
Кандидат экономических наук, доцент кафедры
«Бухгалтерский учет и аудит»
Дальневосточного государственного университета путей сообщения
К.И. Алексеева

Корзова Л.Н.
К… Эконометрика. Методическое пособие на выполнение контрольной работы для студентов экономических специальностей ИИФО. – Хабаровск: Изд-во ДВГУПС, 2002. - с.

Методическое пособие соответствуют государственному стандарту дисциплины «Эконометрика» для всех экономических специальностей ИИФО.

Рассмотрены вопросы построения модели линейных и нелинейных регрессий, качественного исследования их, определения прогнозных значений и доверительных интервалов результативных признаков. Рассмотрена проблема идентификации систем эконометрических уравнений и методы нахождения оценок структурных коэффициентов модели. Исследованы вопросы анализа временных рядов и способы устранения тенденций ряда. Главное внимание уделено качественному анализу моделей и использование их для прогнозирования.

Предназначено для студентов 2,3,4 курсов, экономических специальностей ИИФО.

УДК…
ББК…

© Издательство Дальневосточного государственного
университета путей сообщения (ДВГУПС), 2002

ВВЕДЕНИЕ

Эконометрика одна из базовых дисциплин экономического образования во всем мире.

Общепринятое определение «эконометрика» пока отсутствует, но есть высказывания признанных авторитетов в экономике и эконометрике.

«Основная задача эконометрики – наполнить эмпирическим содержанием априорных экономических рассуждений» (Л. Клейн – лауреат Нобелевской премии 1980 года по экономике (эконометрике)).

«Эконометрика не является синонимом приложения математики к экономике. Как показывает опыт, каждая из отправных точек – статистика, экономическая теория и математика – необходимое, но не достаточное условие для понимания количественных соотношений в современной экономической жизни. Это – единство всех трех составляющих. И это единство образует эконометрику» (Р. Фриш – лауреат Нобелевской премии 1969 года по экономике (эконометрике)).

Данная наука возникла на стыке 3–х дисциплин экономической теории, методов математического анализа и математической статистики. Основными дисциплинами, использующими математические методы применительно к экономике являются:

- многомерный статистический анализ;
- финансовая математика;
- математические модели в экономике;
- и математические методы в экономике.

Эконометрика содержит два больших раздела: моделирование данных, неупорядоченных во времени и теорию временных рядов.

Методические указания предназначены для студентов всех курсов по направлению подготовки 38.03.01 «Экономика».

В ходе изучения курса «Эконометрика» студент должен выполнить одну контрольную работу, содержащую четыре задачи, номера которых выбираются из четырех таблиц по последней цифре шифра зачетной книжки студента.

ТЕМАТИЧЕСКИЙ ПЛАН КУРСА И МЕТОДИЧЕСКИЕ УКАЗАНИЯ К ВЫПОЛНЕНИЮ КОНТРОЛЬНОЙ РАБОТЫ

Глава 1. Парная регрессия и корреляция

Парная регрессия представляет регрессию между двумя переменными y и x, то есть модель вида:

$$y = \hat{f}(x),$$

где y- результативный признак,

x – объясняющая переменная.

Практически в каждом отдельном случае:

$$y_j = \hat{y}_{x_j} + \varepsilon_j, \; j = \overline{1,n},$$

где y_j –фактическое значение результативного признака;

$\hat{y}_{x_j}$ - теоретическое значение результативного признака, найденное по построенному уравнению регрессии;

ε_i - случайная величина, определяющая отклонения реального значения (y) от теоретического ($\hat{y}$), включающая влияние не учтенных в модели факторов, случайных ошибок и ошибок измерения.

1.1. Линейная регрессия и корреляция. Оценка параметров

Основные гипотезы:

1. $y=a+bx+\varepsilon$ - спецификация модели
2. $M\varepsilon = 0$ (M- математическое ожидание)

$M(\varepsilon^2) = \sigma^2$, не зависит от j, ($\sigma^2$ - дисперсия).

3. $M(\varepsilon_i \cdot \varepsilon_{\hat{u}}) = 0$, если $i \neq s$, т.е. ошибки не коррелированы для разных наблюдений.
4. ε - есть случайная величина, нормально распределенная с параметрами $(0, \sigma^2)$, $\varepsilon \sim N(0, \sigma^2)$

Линейная регрессия сводится к нахождению ее параметров a и b, что можно осуществить различными методами. Одним из широко употребляемых методов является метод наименьших квадратов (МНК).

Суть метода состоит в выполнении условия:

$$S = \sum_j (y_j - \hat{y}_j)^2 \to \min$$

Для определения a и b строится система нормальных уравнений:

$$\begin{cases} \dfrac{\partial S}{\partial a} = 0 \\ \dfrac{\partial S}{\partial b} = 0 \end{cases} \quad \text{или} \quad \begin{cases} a + b\bar{x} = \bar{y} \\ a\bar{x} + b\overline{x^2} = \overline{yx} \end{cases}$$

где $\bar{x} = \frac{\Sigma x_j}{n}$; $\bar{y} = \frac{\Sigma y_j}{n}$; $\overline{x^2} = \frac{\Sigma x_j^2}{n}$; $\overline{xy} = \frac{\Sigma x_j y_j}{n}$ - средние значения соответствующих величин.

Из системы нормальных уравнений:

$$b = \frac{\overline{yx} - \bar{x} \cdot \bar{y}}{\overline{x^2} - (\bar{x})^2} = \frac{\overline{yx} - \bar{x} \cdot \bar{y}}{\sigma^2_x}, \; a = \bar{y} - b\bar{x}$$

b – параметр модели. Его величина показывает среднее изменение результативного признака с изменением объясняющий переменной на единицу.

Уравнение регрессии необходимо дополнить показателем тесноты связи, в качестве его выступает коэффициент корреляции r_{xy}.

$r_{xy} = b\frac{\sigma_x}{\sigma_y} = \frac{\overline{yx} - \bar{x} \cdot \bar{y}}{\sigma_x \cdot \sigma_y}$, обладающий свойством: $-1 \le r_{xy} \le 1$, причем,

если $b>0$, то $0 \le r_{xy} \le 1$,

если $b<0$, то $-1 \le r_{xy} \le 0$.

Близость r_{xy} к нулю означает отсутствие линейной связи между признаками, но возможна и другая спецификация модели.

Величина r^2_{xy} называется коэффициентом детерминации, характеризующим долю дисперсии признака y, объясненную регрессией, в общий дисперсии признака y.

Для проверки гипотезы H_0 : b=0 (отсутствие линейной связи между признаками) используют F-статистику (F-распределение).

F-статистика для этой гипотезы имеет вид:

$$F_{расч.} = \frac{r^2}{1 - r^2}(n-2),$$

где n – объем совокупности или

$$F_{расч.} = \frac{D_{факт.}}{D_{ост.}},$$

где $D_{факт.}$ и $D_{ост.}$ - дисперсии на одну степень свободы.

$F_{расч.}$ сравнивают с $F_{табл.}$, значение $F_{табл.}$ находят в таблице (см. приложение 1) по параметрам: α, ν_1, ν_2, где α - уровень значимости, обычно выбирают 0,05 или 0,01,

ν_1 - число степеней свободы дисперсии

$$D_{факт.} = \frac{\Sigma(\hat{y}_x - \bar{y})^2}{1}, \; (\nu_1 = 1);$$

ν_2 - число степеней свободы дисперсии

$$D_{ост.} = \frac{\Sigma(y - \hat{y})^2}{n-2}, \; (\nu_2 = n-2);$$

Гипотеза H_0 отвергается и принимается альтернативная, если $F_{расч.} > F_{табл.}$

Для оценки не только уравнения в целом, но и его параметров используют t- статистику (t- критерий) Стьюдента.

Для этого рассчитывают стандартные ошибки m_a, m_b, m_r, $m_{\hat{y}x}$

$$m_a = \sqrt{\frac{\Sigma(y-\hat{y}_x)^2}{n-2}\cdot\frac{\Sigma x^2}{n\Sigma(x-\bar{x})^2}};$$

$$m_b = \sqrt{\frac{\Sigma(y-\hat{y}_x)^2/(n-2)}{\Sigma(x-\bar{x})^2}};$$

$$m_r = \sqrt{\frac{1-r^2}{n-2}}$$

С помощью стандартных ошибок находят расчетные значение t-критерия.

$$t_a = \frac{a}{m_a},\ t_b = \frac{b}{m_b},\ t_r = \frac{r}{\sqrt{1-r^2}}\sqrt{n-2}$$

Далее сравнивают их с табличным значением t-статистики (см. приложение 2).

$t(\alpha,\nu) = t(\alpha=0{,}01 \text{ или } 0{,}05\ ;\ \nu=n-2)$, если $t_{расч.} > t_{табл.}$ Гипотеза H_0 о несущественности коэффициентов отвергается, принимается альтернативная.

Если гипотезы о существенности уравнения регрессии в целом и его параметров в частности принимаются, уравнение можно использовать как для точечного прогноза, так и для интервальной оценки прогнозного значения. Для этого рассчитывается стандартная ошибка $m_{\hat{y}x}$,

где $m_{\hat{y}_x} = \sqrt{\left(\frac{\Sigma(y-\hat{y}_x)^2}{n-2}\right)\cdot\left(\frac{1}{n}+\frac{(x_k-\bar{x})^2}{\Sigma(x-\bar{x})^2}\right)}$,

доверительный интервал прогноза: $(\hat{y}_{прогн} - t_{табл}\cdot m_{\hat{y}\,хпрог.}, \hat{y}_{прогн.} + t_{табл.}\cdot m_{\hat{y}\,хпрог.})$,

где $t_{табл.} = t(\alpha=0{,}01(0{,}05)),\ \nu = n-2)$

Замечание.

Условие независимости дисперсии ошибки от номера наблюдения называется гомоскедастичностью. Случай, когда это условие не выполняется, называется гетероскедаксичностью. В последнем случае используют обобщенный метод наименьших квадратов (ОМНК).

$$\sigma_{x_i} = \sigma_i k_i,\quad y_i = \alpha_0 + \alpha_i x_i + \sqrt{k_i}\varepsilon_i,$$

k_i – коэффициенты, характер которых зависит от выдвинутых гипотез.

1.2. Нелинейная регрессия и корреляция для нелинейной регрессии

Нелинейные связи между признаками выражаются с помощью соответствующих функций, например:

$$y = a + b\sqrt{x} + \varepsilon,$$

$$y = a + b_1 x + b_2 x^2 + \varepsilon$$

$$y = a + \frac{b}{x} + \varepsilon$$

и другие регрессии, линейные по параметрам;

$$y = ab^x \cdot \varepsilon, \quad y = ae^{bx} \cdot \varepsilon$$

и другие регрессии нелинейные по оцениваемым параметрам. Для полиномиальных регрессий можно непосредственно использовать МНК, для некоторых других нелинейных регрессий возможна их предварительная линеаризация, например, $y = a + b\sqrt{x} + \varepsilon$, обозначив $\sqrt{x} = z,$ получим $y = a + bz + \varepsilon$, $y = a \cdot e^{bx} \cdot \varepsilon, \quad \ln y = \ln a + bx + \ln \varepsilon,$ обозначив ln y=z, ln a=a*, ln $\varepsilon = \varepsilon^x$, получим z=a*+bx+ ε *

Уравнение нелинейной регрессии оценивается с помощью индекса корреляции R:

$$R = \sqrt{1 - \frac{\sigma^2_{ост.}}{\sigma^2_y}} = \sqrt{1 - \frac{\Sigma(y - \hat{y})^2}{\Sigma(y - \bar{y})^2}}, \; 0 \leq R \leq 1$$

Для оценки качества моделей определяется средняя ошибка аппроксимации: $\bar{A} = \frac{1}{n}\Sigma\left|\frac{y - \hat{y}}{y}\right| \times 100\%$, допустимые значения которой 8-10 %.

При анализе построенной и оцененной модели используется средний показатель эластичности: $\bar{Э} = f'(x)\frac{\bar{x}}{\bar{y}}$, показывающий на сколько процентов изменится в среднем результат, если фактор изменится на 1%.

ПРИМЕР 1

Таблица 1.1

Год	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Выручка от реализации (млн. у.е.)	14	15	17	20	24	30	48	49	59	67

В таблице 1.1 приведены данные о годовой выручке (y) от реализации продукции за десятилетний период.

Требуется:

1. Для определения вида зависимости рассчитать параметры a и b функций:

$$а)\ y=a+bx+\varepsilon,$$

$$б)\ y=ae^{bx}\varepsilon.$$

2. Оценить тесноту связи и значимости параметров управления, используя F-распределение (Фишера) и t-критерий (Стьюдента), выбрав уровень значимости $\alpha=0{,}01$.

3. Рассчитать среднюю ошибку аппроксимации.

4. Выполнить прогноз выручки от продаж при прогнозном значении признака x, составляющим 102% от среднего уровня при условии, что ошибка аппроксимации не превысит 8-10%.

5. Оценить точность прогноза, рассчитав ошибку прогноза и его доверительный интервал.

6. Сделать выводы по результатам расчетов.

Присвоим каждому году x соответствующий код 1, 2, …, 10.

Составим таблицу расчетов 1.2.

Таблица 1.2

	x	x^2	y	xy	y^2	$y-\bar{y}$	$x-\bar{x}$	$(y-\bar{y})^2$	$(x-\bar{x})^2$	$\hat{y}$	$y-\hat{y}$	$(y-\hat{y})^2$	A (%)
	1	1	14	14	196	-20,3	-4,5	412,09	20,25	6,073	7,927	62,8373	56,62
	2	4	15	30	225	-19,3	-3,5	372,49	12,25	12,346	2,054	7,0437	17,69
	3	9	17	51	289	-17,3	-2,5	299,29	6,25	18,619	-1,619	2,6212	9,52
	4	16	20	80	400	-14,3	-1,5	204,49	2,25	24,892	-4,892	23,9317	24,46
	5	25	24	120	576	-10,3	-0,5	106,09	0,25	31,165	-7,165	51,3372	29,85
	6	36	30	180	900	-4,3	0,5	18,49	0,25	37,438	-7,438	55,3223	24,79
	7	49	48	336	2304	13,7	1,5	187,69	2,25	43,710	4,290	18,4041	8,93
	8	64	49	392	2401	14,7	2,5	216,09	6,25	49,984	-0,984	0,9682	2,00
	9	81	59	531	3481	24,7	3,5	610,09	12,25	56,257	2,743	7,5240	4,65
	10	100	67	670	4489	32,7	4,5	1069,29	20,25	62,530	4,470	119,9809	6,67
Σ	55	385	343	2404	15261	0	0	34,96,1	82,50		0,014	249,9706	185,18
Среднее значение	$\bar{x}=5,5$	$\overline{x^2}=38,5$	$\bar{y}=34,3$	$\overline{xy}=240,4$	$\overline{y^2}=1526,1$							24,9971	18,518
σ	2,872		18,698										
σ^2	8,250		349,610										

Все расчеты велись по формулам:

$$\bar{x} = \frac{\sum x_i}{n}; \qquad \sigma_x^2 = \overline{x^2} - (\bar{x})^2; \sigma_x = \sqrt{\sigma_x^2}$$

$$\bar{y} = \frac{\sum y_i}{n}; \qquad \sigma_y^2 = \overline{y^2} - (\bar{y})^2; \sigma_y = \sqrt{\sigma_y^2}$$

$$b = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sigma_x^2} = \frac{240{,}4 - 5{,}5 \cdot 34{,}3}{8{,}25} = 6{,}273$$

$$a = \bar{y} - b\bar{x} = 34{,}3 - 6{,}273 \cdot 5{,}5 = -0{,}2$$

1а) Линейное уравнение регрессии имеет вид: $\hat{y} = -0.2 + 6{,}273x$

Рассчитаем линейный коэффициент парной корреляции:

$$r_{xy} = b\frac{\sigma_x}{\sigma_y} = 6{,}273\frac{2{,}872}{18{,}698} = 0{,}963$$

т.к. значение коэффициента корреляции близко к единице, связь между признаками y и x тесная.

2. Рассчитаем значение F-критерия Фишера.

$$F_{факт.} = F_{расч.} = \frac{\Sigma(\hat{y} - \bar{y})^2 / m}{\Sigma(y - \hat{y})^2 / (n - m - 1)} = \frac{r_{xy}^2}{1 - r_{xy}^2} \cdot \frac{(n - m - 1)}{m}$$, где

m – число параметров уравнения регрессии (число коэффициентов при объясняющей переменной x);

n – объем совокупности.

$$F_{факт.} = \frac{(0{,}963)^2}{1 - (0{,}963)^2} \cdot 8 = \frac{0{,}927}{0{,}073} \cdot 8 = 101{,}629$$

По таблице приложения находим

$$F_{табл.} = F_{крит.}(\alpha = 0{,}01; \nu_1 = 1; \nu_2 = 8) = 11{,}26$$

т.к. $F_{факт.} > F_{табл.}$, то гипотеза H_0 о статистической незначимости параметра b в уравнении регрессии отклоняется.

Т.к. $r_{xy}^2 = 0{,}927$, это означает, что 92,7% результата объясняется вариациями объясняющей переменной.

Средняя ошибка аппроксимаций $\overline{A} = 18{,}5\%$ вышла за допустимые пределы (8-10%), что говорит о неудачном выборе модели регрессии.

Выберем в качестве модели уравнение регрессии уравнение 1б) ($y = ae^{bx}\varepsilon$), предварительно линеаризировав модель. Для этого прологарифмируем обе части уравнения.

$$\ln y = \ln a + bx + \ln\varepsilon,$$

Введем обозначения

$$Z = \ln y,\ a^* = \ln a,\ \varepsilon^* = \ln\varepsilon,$$

Получим линейное уравнение парной регрессии

$$Z = a^* + bx + \varepsilon^*.$$

Все промежуточные расчеты поместим в таблицу 1.3.

Таблица 1.3

	x	Z=lny	xZ	Z^2	$Z-\overline{Z}$	$(Z-\overline{Z})^2$	$\hat{Z}$	$\hat{y}=e^{\hat{z}}$	$Z-\hat{Z}$	$(Z-\hat{Z})^2$	$y-\hat{y}$	$(y-\hat{y})^2$	A_i
	1	2,6381	2,6381	6,9596	-0,7410	0,5491	2,5106	12,2432	0,1275	0,0162	1,7568	3,0863	4,83
	2	2,7071	5,4142	7,3284	-0,6720	0,4515	2,7036	14,6509	0,0035	0,0000	0,3491	0,1219	0,13
	3	2,8322	8,4966	8,0214	-0,5469	0,2991	2,8966	18,0832	-0,0644	0,0041	-1,0832	1,1733	2,27
	4	2,9947	11,9788	8,9682	-0,3844	0,1478	3,0896	21,8700	-0,0949	0,0090	-1,8700	3,4969	3,17
	5	3,1771	15,8855	10,0939	-0,2020	0,0408	3,2826	26,4465	-0,1055	0,0111	-2,4465	5,9854	3,32
	6	3,4002	20,4012	11,5614	0,0211	0,0004	3,4756	31,9748	-0,0754	0,0057	-1,9748	3,8998	2,22
	7	3,8702	27,0914	14,9784	0,4911	0,2412	3,6686	391574	0,2016	0,0406	8,8426	78,1916	5,21
	8	3,8908	31,1264	15,1382	0,5117	0,2618	3,8616	47,2291	0,0292	0,0008	1,7709	3,1361	0,75
	9	4,0766	36,6894	16,6187	0,6975	0,4865	4,0546	57,6852	0,0220	0,0005	1,3148	1,7257	0,54
	10	4,2037	42,0370	17,6711	0,8246	0,6799	4,2476	69,7572	-0,0439	0,0019	-2,7572	7,6022	1,04
Σ	55	33,7907	201,7586	117,3393	-0,003	3,1583			0,0003	0,0899		108,4222	23,48
Среднее значение	$\overline{x}=5,5$	$\overline{Z}=3,3791$	$\overline{xz}=20,1759$	$\overline{Z^2}=11,7339$									$\overline{A}=2,348$
σ	2,872						0,5618						
σ^2	8,250						0,3156						

1. Рассчитаем параметры уравнения, предварительно рассчитав

$$\sigma_Z^2 = \overline{Z}^2 - (\overline{Z})^2 = 11{,}7339 - 11{,}4183 = 0{,}3156$$

$$\sigma_Z = 0{,}5618$$

$$b = \frac{\overline{xz} - \overline{x} \cdot \overline{z}}{\sigma_x} = \frac{1{,}5907}{8{,}250} = 0{,}193,\ a^* = \overline{Z} - b\overline{x} = 2{,}3176$$

$$\hat{Z} = 2{,}3176 + 0{,}193x$$

2. Коэффициент корреляции

$$r_{xz} = b\frac{\sigma_x}{\sigma_z} = 0{,}986,$$

значение r_{xz} близко к единице, что говорит о тесной связи признаков.

Коэффициент детерминации

$$r^2=0{,}972,$$

следовательно 97,2% результата объясняется вариацией объясняющей переменной x.

$$F_{факт.} = F_{расчет.} = \frac{r^2}{1-r^2}(n-2) = \frac{0{,}972}{0{,}028} \cdot 8 = 277{,}71$$

$$F_{табл.}(\alpha = 0{,}01; v_1 = 1; v_2 = 8) = 11{,}26$$

$$F_{факт.} \rangle F_{табл.},$$

следовательно, гипотеза H_0 о статистической незначимости уравнения регрессии отклоняется. Уравнение регрессии статистически значимо в целом.

Оценим значимость каждого параметра уравнения регрессии.

Используем для этого t-распределения (Стьюдента).

Выдвигаем гипотезу H_0 о статически незначимом отличии параметров от нуля: $a=b=r_{xz}=0$.

$$t_{табл.}\ (\alpha=0{,}01;\ v=8)=3{,}355$$

Определим ошибки m_a, m_b, $m_{r_{xz}}$.

$$m_a = \sqrt{\frac{\Sigma(y-\hat{y})^2}{n-2} \cdot \frac{\Sigma x^2}{n\Sigma(x-\overline{x})^2}} = \sqrt{\frac{0{,}0899 \cdot 385}{8 \cdot 10 \cdot 82{,}5}} = 0{,}0707$$

$$m_b = \sqrt{\frac{\Sigma(y-\hat{y})^2/(n-2)}{\Sigma(x-\overline{x})^2}} = \sqrt{\frac{0{,}0899}{8 \cdot 82{,}5}} = 0{,}0116$$

$$m_r = \sqrt{\frac{1-r^2}{n-2}} = \sqrt{\frac{0{,}028}{8}} = 0{,}0591$$

$$t_a = \frac{a}{m_a} = 32{,}78 \rangle t_{табл.}$$

$$t_b = \frac{b}{m_b} = 16{,}64 \rangle t_{табл.}$$

$$t_r = \frac{r}{m_r} = 16{,}68 \rangle t_{табл.}$$

следовательно, a, b, r не случайно отличаются от 0, а сформировались под влиянием систематически действующей объясняющей переменной.

3. $\overline{A} = 2{,}348\%$ (т.е. в среднем расчетные значения отклоняются от фактических на 2,3%), следовательно, качество модели хорошее.

4. Полученные оценки модели и ее параметров позволяют использовать ее для прогноза.

Рассчитаем $x_{прогн.} = \bar{x} \cdot 1{,}02 = 5{,}61$

$$\hat{z}_{прогн.} = 2{,}3176 + 0{,}193 \cdot 5{,}61 = 3{,}4003$$

5. Средняя ошибка прогноза

$$m_{\hat{z}_{прогн.}} = \sigma_{ост.} \cdot \sqrt{1 + \frac{1}{n} + \frac{(x_{прогн.} - \bar{x})^2}{\Sigma(x - \bar{x})^2}}$$, где

$$\sigma_{ост.} = \sqrt{\frac{\Sigma(z - \hat{z})^2}{n - m - 1}} = \sqrt{\frac{0{,}0899}{8}} = 0{,}106,$$

$$m_{\hat{z}_{прогн.}} = 0{,}106\sqrt{1 + 0{,}1 + \frac{0{,}11}{82{,}5}} = 0{,}1113.$$

Строим доверительный интервал (т.е. интервал, включающий в себе оцениваемое значение с вероятностью, близкой к единице).

$$\left(\gamma_{\hat{z}_{\min}}, \gamma_{\hat{z}_{\max}}\right)$$

$$\left(\hat{z}_{прогн.} - \Delta_{\hat{z}_{прогн.}}, \hat{z}_{прогн.} + \Delta_{\hat{z}_{прогн.}}\right)$$, где

$$\Delta_{\hat{z}_{прогн.}} = t_{табл.} \cdot m_{\hat{z}_{прогн.}} = 3{,}355 \cdot 0{,}1113 = 0{,}373$$

$$3{,}4003 - 0{,}373 \langle \hat{z}^*_{прогн.} \langle 3{,}4003 + 0{,}373$$

$$3{,}0273 \langle \hat{z}^*_{прогн.} \langle 3{,}7733$$

Выполненный процесс надежен (вероятность p=1-α=0,99) и достаточно точен, т.к. $\gamma_{\hat{z}_{\max}} / \gamma_{\hat{z}_{\min}} \approx 1{,}2$.

6. Вернемся к первоначальной переменной $y = ae^{bx} \cdot \varepsilon$,

$$\hat{y} = 10{,}176e^{0{,}193x}.$$

Воспользуемся формулой расчета индекса корреляции для нелинейной регрессии

$$\rho_{xy} = \sqrt{1 - \frac{\Sigma(y - \hat{y})^2}{\Sigma(y - \bar{y})^2}} = \sqrt{1 - \frac{108{,}4222}{3496{,}1}} = 0{,}984.$$

ρ_{xy} отличается от r_{xy} в третьем знаке из-за ошибок вычислений и округлений.

КОНТРОЛЬНОЕ ЗАДАНИЕ 1

Задачи 1 – 5

Имеются данные за 12 месяцев года по району города о рынке вторичного жилья (y – стоимость квартиры (тыс. у.е.), x– размер общей площади ($м^2$)). Данные приведены в таблице 1.4.

Таблица 1.4

Мес.	Задача 1		Задача 2		Задача 3		Задача 4		Задача 5	
	y	x	y	x	y	x	y	x	y	x
1	13,0	37,0	13,2	37,2	22,5	46,0	22,5	29,0	23,0	22,8
2	16,4	60,0	15,9	58,2	25,5	54,0	25,8	36,2	26,8	27,5
3	17,0	60,9	16,2	60,8	19,2	50,2	20,8	28,9	28,0	34,5
4	15,2	52,1	15,4	52,0	13,6	43,8	15,2	32,4	18,4	26,4
5	14,2	40,1	14,2	44,6	25,4	78,6	25,8	49,7	30,4	19,8
6	10,5	30,4	11,0	31,2	17,8	60,2	19,4	38,1	20,8	17,9
7	20,0	43,0	21,1	26,4	18,0	50,2	18,2	30,0	22,4	25,2
8	12,0	32,1	13,2	20,7	21,0	54,7	21,0	32,6	21,8	20,1
9	15,6	35,1	15,4	22,4	16,5	42,8	16,4	27,5	18,5	20,7
10	12,5	32,0	12,8	35,4	23,0	60,4	23,5	39,0	23,5	21,4
11	13,2	33,0	14,5	28,4	14,6	47,2	18,8	27,5	16,7	19,8
12	14,6	32,5	15,1	20,7	14,2	40,6	17,5	31,2	20,4	24,5

Задачи 6 - 10

Имеются данные о потребительских расходах на душу населения (y, руб.), средней заработной плате и социальных выплатах (x, руб.) по 16 районам региона. Данные приведены в таблице 1.5.

Таблица 1.5

районы	Январь		Февраль		Март		Апрель		Май	
	Задача 6		Задача 7		Задача 8		Задача 9		Задача 10	
	y	x	y	x	y	x	y	x	y	x
A	416	1288	420	1305	440	1310	445	1310	450	1380
B	501	1435	512	1440	525	1490	537	1490	540	1530
C	403	1210	430	1230	450	1250	463	1255	470	1310
D	208	1190	230	1275	240	1280	251	1287	275	1355
E	462	1640	505	1700	545	1710	553	1720	564	1810
F	386	1420	402	1480	447	1497	453	1500	462	1520

G	399	1250	430	1305	469	1312	478	1320	495	1410
H	342	870	400	895	435	903	448	918	450	1000
I	354	740	410	775	442	787	453	794	460	810
J	558	910	585	1000	605	1012	627	1012	633	1051
K	302	1020	37	1035	352	1049	364	1058	377	1078
L	360	1050	384	1150	405	1207	419	1213	422	1231
M	310	1205	345	1215	376	1221	392	1225	405	1230
N	415	990	445	1010	462	1035	478	1042	483	1050
O	452	1042	485	1059	505	1064	521	1071	537	1095
P	450	1037	491	1051	500	1072	517	1093	529	1101

Задание:

1. Рассчитайте параметры уравнений регрессий y=a+bx+ε и y=a+b$\sqrt{x}$+ε.
2. Оцените тесноту связи с показателем корреляции и детерминации.
3. Рассчитайте средний коэффициент эластичности и дайте сравнительную оценку силы связи фактора с результатом.
4. Рассчитайте среднюю ошибку аппроксимации и оцените качество модели.
5. С помощью F-статистики Фишера (при α=0,5) оцените надежность управления регрессий.
6. Рассчитайте прогнозное значение $\hat{y}_{пр.}$,если прогнозное значение фактора увеличится на 5% от его среднего значения. Определите доверительный интервал прогноза для α=0,01.
7. Расчеты должны быть подробны как показаны в примере 1 и сопровождены пояснительной запиской.

Глава 2. Множественная регрессия

Обобщением модели парной регрессии является модель множественной регрессии. Наиболее часто в эконометрике используется линейная модель, т.е. уравнение вида:

$$y = a + \sum_{j=1}^{k} b_j x_j + \varepsilon, \qquad (2.1)$$

Используются модели и нелинейных регрессий, например:

$$y = a \cdot e^{\Sigma b_j x_j} \cdot \varepsilon, \quad y = a x_1^{b_1} \cdot ..... \cdot x_k^{b_k} \cdot \varepsilon \text{ и др.}$$

Построение модели связано с выбором вида уравнения и отбором факторов. Факторы, включаемые в модель должны удовлетворять требованиям:

- должны быть количественно измеримы;

- не должны находиться в точной функциональной связи;
- между факторами не должно быть высокой корреляционной связи;
- факторы не должны быть коллинеарны и мультиколлениарны, то есть факторы не должны дублировать друг друга и более двух факторов не должны быть линейно зависимы.

Замечание. Для оценки мультиколлинеарности факторов можно использовать матрицу М:

$$M = \begin{pmatrix} r_{x_1x_1} r_{x_2x_1} ....r_{x_kx_1} \\ r_{x_1x_2} r_{x_2x_2} ....r_{x_kx_2} \\ r_{x_1x_k} r_{x_2x_{k1}} ....r_{x_kx_k} \end{pmatrix}$$

$r_{x_ix_j}$ - коэффициенты парной корреляции.

1. Если факторы не коррелированы, то определитель матрицы, det M=1

2. Если между факторами – функциональная связь det M=0

Чем ближе к нулю det M, тем сильнее мультиколлинеарность и ненадежнее результаты множественной регрессии.

Один из путей устранения мультиколлинеарности – исключение из модели одного или нескольких факторов.

2.1. Нахождение параметров линейного уравнения множественной регрессии

Параметры уравнения (2.1) оцениваются МНК, результаты применения которого приводят к следующей системе уравнений

$$\begin{cases} \overline{y} = a + b_1\overline{x_1} + ... + b_k\overline{x_k} \\ \overline{yx_1} = a\overline{x_1} + b_1\overline{x_1^2} + ... + b_k\overline{x_kx_1} \\ ..... \\ \overline{yx_k} = a\overline{x_k} + b_1\overline{x_1x_k} + ... + b_k\overline{x_k^2} \end{cases}$$

Решив которую, любым способом, например, Крамера, получают коэффициенты a, b_j, j=$\overline{1,k}$.

Коэффициенты уравнения множественной регрессии можно определить с помощью стандартизированных коэффициентов регрессии на основании уравнения регрессии в стандартизированном масштабе.

$$\hat{t}_y = \beta_1 t_{x_1} + \beta_2 t_{x_2} + ... + \beta_k t_{x_k} \qquad (2.2)$$

где t_y, t_j (j=$\overline{1,k}$) – стандартизированные переменные, причем

$$t_y = \frac{y - \overline{y}}{\sigma_y},\ t_{x_j} = \frac{x_j - \overline{x_j}}{\sigma_{x_j}},$$

Система нормальных уравнений для определения коэффициентов $\beta_j, j=\overline{1,k}$ уравнения (2.2) имеет вид:

$$\begin{cases} r_{yx_1} = \beta_1 + \beta_2 r_{x_2x_1} + \beta_3 r_{x_3x_1} + ... + \beta_k r_{x_kx_1} \\ r_{yx_2} = \beta_1 r_{x_1x_2} + \beta_2 + \beta_3 r_{x_3x_2} + ... + \beta_k r_{x_kx_2} \\ ..... \\ r_{yx_k} = \beta_1 r_{x_1x_k} + \beta_2 r_{x_2x_{k1}} + \beta_3 r_{x_3x_{k1}} + ... + \beta_k \end{cases} \quad (2.3)$$

где $r_{x_ix_j} = \beta_i \frac{\sigma_y}{\sigma_{x_i}}$ - парные коэффициенты корреляции,

$a = \overline{y} - \Sigma b_j \overline{x_j}$, j=$\overline{1,k}$.

2.2. Частные уравнения регрессии

На основе уравнения (2.1) могут быть найдены частные уравнения регрессии.

$$\begin{cases} y_{x_1 \cdot x_2, x_3, ..., x_k} = a + b_1 x_1 + b_2 \overline{x_2} + ... + b_k \overline{x_k} + \varepsilon \\ y_{x_2 \cdot x_1, x_3, ..., x_k} = a + b_1 \overline{x_1} + b_2 x_2 + ... + b_k \overline{x_k} + \varepsilon \\ y_{x_k \cdot x_1, x_2, ..., x_{k-1}} = a + b_1 \overline{x_1} + b_2 \overline{x_2} + ... + b_k \overline{x_k} + \varepsilon \end{cases} \quad (2.4)$$

Частные уравнения регрессии характеризуют изолированное влияние фактора, при условии, что остальные закреплены на неизменном уровне.

Система (2.4) имеет вид:

$$\begin{cases} \hat{y}_{x_1 \cdot x_2 x_3 ... x_k} = A_1 + b_1 x_1 \\ \hat{y}_{x_k \cdot x_1 x_2 ... x_{k-1}} = A_k + b_k x_k \end{cases} \quad (2.5)$$

если в (2.4) подставить $\overline{x_1}, ..., \overline{x_k}$

2.3. Множественная корреляция

Значимость уравнения множественной регрессии оценивается с помощью индекса множественной корреляции $R_{yx_1 ....x_k}$

$R_{yx_1 ....x_k} = \sqrt{1 - \frac{\sigma^2_{ост}}{\sigma^2 y}}$, обладающего свойствами: $0 \le R_{yx_1 ....x_k} \le 1$,

$R_{yx_1 ....x_k} \ge$ max $r_{y_{x_j}}$, j=$\overline{1,k}$.

Вывод о целесообразности включения фактора в уравнение множественной регрессии можно сделать после сравнения индексов коэффициентов множественной и парной регрессии.

Для оценки линейного уравнения регрессии можно воспользоваться формулой:

$$R_{yx_1 \ldots x_k} = \sqrt{\Sigma \beta_{x_j} r_{yx_j}}$$

Замечания.

Если число коэффициентов (в) при переменных х приближается к n, то коэффициент (индекс) корреляции близок к единице даже при слабой связи признаков.

В этом случае используется скорректированный индекс (коэффициент) корреляции: $\widetilde{R} = \sqrt{1 - \frac{\Sigma(y - \hat{y})^2 / (n - m - 1)}{\Sigma(y - \bar{y})^2 / (n - 1)}}$

2.4. Частная корреляция

Для оценки частных уравнений регрессии могут быть использованы индексы (коэффициенты) частной корреляции:

$$r_{yx_i \cdot x_1 \ldots x_{i-1} x_{i+1} \ldots x_k} = \sqrt{1 - \frac{R^2_{yx_1 x_2 \ldots x_2 \ldots x_k}}{R^2_{yx_1 \ldots x_{i-1} x_{i+1} \ldots x_k}}},$$

где $R_{yx_1 \ldots x_{i-1} x_{i+1} \ldots x_k}$ - множественный коэффициент детерминации всего набора к факторов с признаком у.

Возможны коэффициенты корреляции первого, второго, третьего, …, (к-1) порядков, например, влияние фактора x_1 можно оценивать при разных условиях независимости действия других факторов:

- $r_{yx_1 \cdot x_2}$ - при постоянном действии фактора x_2;

- $r_{yx_1 \cdot x_2 x_3 x_4 x_5}$ - при постоянном действии факторов x_2, x_3, x_4, x_5;….. .

- $r_{yx_1 \cdot x_2 x_3 x_4 x_5 \ldots\ldots x_r}$ - при постоянном действии всех факторов, включенных в регрессию.

При исследовании зависимостей предпочтение отдается показателям частной корреляции более высоких порядков, они являются дополнением к уравнению множественной регрессии. В эконометрике частные

коэффициенты корреляции не имеют самостоятельного значения. Их используют на этапе формирования модели, например, для отсева некоторых факторов.

2.5. Оценка надежности результатов множественной регрессии и корреляции

Аналогично парной регрессии оценка осуществляется с помощью F-статистики.

$$F_{\text{расч.}} = \frac{D_{факт.}}{D_{ост.}} = \frac{R^2 / m}{(1-R^2)/(n-m-1)}$$

Для оценки значимости влияние фактора x_i рассчитывается значение частной F- статистики.

$$F_{x_{i\,расч.}} = \frac{R^2_{yx_1 \cdots x_k} - R^2_{yx_1 \cdots x_{i-1}x_{i+1} \cdots x_k}}{1 - R^2_{yx_1 \ldots x_k}} \times \frac{n-m-1}{1}$$

Зная величину $Fx_{i\text{расч.}}$ можно определить значение t- статистики для коэффициентов b_i.

$$t_{b_i} = \sqrt{F_{x_i}}$$

t_{b_i} можно рассчитывать и иначе $t_{b_i} = \frac{b_i}{m_{b_i}}$ в частности для уравнения (2.1)

$$m_{b_i} = \frac{\sigma_y}{\sigma_{x_i}} \times \frac{\sqrt{1 - R^2_{yx_1 \ldots\ldots x_k}}}{\sqrt{1 - R^2_{yx_1 \ldots x_{i-1}x_{i+1} \ldots x_k}}} \times \frac{1}{\sqrt{n-m-1}}$$

2.6. Фиктивные переменные

Возможны случаи, когда в модель регрессии необходимо включить факторы, имеющие атрибутивные признаки, например, образование, тип изделия, профессию и т.д.

Чтобы использовать эти переменные им присваивают цифровые пометки. Такие искусственно сконструированные переменные в эконометрике называются фиктивными или структурными переменными.

Фиктивные переменные могут вводиться как в линейные, так и нелинейные модели при условии, что последние можно линеаризировать.

ПРИМЕР 2

Изучается влияние стоимости основных и оборотных средств на величину валового дохода торговых предприятий.

Для этого по 12 торговым предприятиям получены данные, приведенные в таблице 2.3.

Таблица 2.3.

Валовой доход (y), млн. руб.	203	63	45	13	121	88	110	56	80	237	160	75
Среднегодовая стоимость основных фондов, (x_1)	118	28	17	50	56	102	116	124	114	154	115	98
Среднегодовая стоимость оборотных средств, (x_2)	105	56	54	63	28	50	54	42	36	106	88	46

Требуется:

1. Требуется построить уравнение множественной линейной регрессии в стандартизованной и естественной форме. Рассчитать частные коэффициенты эластичности.
2. Рассчитать линейные коэффициенты частной корреляции и коэффициент множественной корреляции.
3. Рассчитать общий и частные F- статистики Фишера.
4. По результатам расчетов сделать соответствующие выводы.

Решение:

Результаты расчетов приведены в таблице 2.4.

Таблица 2.4

	y	x_1	x_2	yx_1	yx_2	x_1x_2	x_1^2	x_2^2	y^2
1	2	3	4	5	6	7	8	9	10
	203	118	105	23954	21315	12390	13924	11025	41209
	63	28	56	1764	3528	1568	784	3136	3969
	45	17	54	765	2430	918	289	2916	2025
	113	50	63	5650	7119	3150	2500	3969	12769
	121	56	28	6776	3388	1568	3136	784	14641
	88	102	50	8976	4400	5100	10404	2500	7744
	110	116	54	12760	5940	6264	12100	2916	12100
	56	124	42	6944	2352	5208	15376	1764	3136

Продолжение табл.2.4

1	2	3	4	5	6	7	8	9	10
	80	114	36	9120	2880	4104	12996	1296	6400
	237	154	106	36498	23716	16324	23716	11236	56169
	160	115	88	18400	14080	10120	13225	7740	25600
	75	98	46	7350	3450	4508	9604	2116	5625
$\sum$	1351	1092	728	138957	94598	71222	118054	51348	191387
Средн.	112,583	91	60,667	11579,75	7883,167				15948,916
σ	57,219	39,457	24,549						
σ^2	3273,985	1556,833	602,6821						

Рассматриваем уравнение вида:

$$Y= a+в_1x_1+в_2x_2+\varepsilon \qquad (1)$$

Параметры уравнения можно найти из решения системы уравнений:

$$\begin{cases} \overline{y} = a + b_1\overline{x_1} + b_2\overline{x_2} \\ \overline{yx_1} = a\overline{x_1} + b_1\overline{x_1^2} + b_2\overline{x_2x_1} \\ \overline{yx_2} = a\overline{x_2} + b_1\overline{x_1x_2} + b_2\overline{x_2^2} \end{cases}$$

Или перейдя к уравнению в стандартизированном масштабе:

$$t_y=\beta_1t_{x1}+\beta_2t_{x2}+\varepsilon',$$

где $t_y = \frac{y-\overline{y}}{\sigma_y}$, $t_{xj} = \frac{x_j-\overline{x_j}}{\sigma x_j}$ - стандартизированные переменные,

β_j - стандартизированные коэффициенты,

$$b_j= \beta_j \frac{\sigma_y}{\sigma_{x_j}}, \quad j=\overline{1,k}.$$

Коэффициенты β_j, в частности для к=2, определяются из системы уравнений:

$$\begin{cases} r_{yx_1} = \beta_1 + \beta_2 r_{x_2x_1} \\ r_{yx_2} = \beta_1 r_{x_1x_2} + \beta_2 \end{cases}$$

$$\beta_1 = \frac{r_{yx_1} - r_{yx_2}r_{x_1x_2}}{1-r^2_{x_1x_2}}, \qquad \beta_2 = \frac{r_{yx_2} - r_{yx_1}r_{x_1x_2}}{1-r^2_{x_1x_2}};$$

$$r_{x_1x_2} = \frac{\overline{x_2x_1} - \overline{x_2}\cdot\overline{x_1}}{\sigma_{x_1}\sigma_{x_2}}, \qquad r_{yxj} = \frac{\overline{yx_j} - \overline{y}\cdot\overline{x_j}}{\sigma_{x_j}\sigma_y};$$

$$r_{yx_1} = \frac{11579,750 - 112,583\cdot 91}{57,219\cdot 39,457} = 0,591, \qquad r^2_{yx_1} = 0,349;$$

$$r_{yx_2} = \frac{7883{,}167 - 112{,}583 \cdot 60{,}667}{57{,}219 \cdot 24{,}549} = 0{,}749 \quad r^2_{yx_2} = 0{,}561;$$

,

$$r_{x_1x_2} = \frac{5935{,}167 - 5520{,}697}{968{,}629} = 0{,}428, \qquad r^2_{x_1x_2} = 0{,}183;$$

$$\beta_1 = \frac{0{,}591 - 0{,}749 \cdot 0{,}428}{0{,}817} = 0{,}331, \qquad \beta_2 = \frac{0{,}749 - 0{,}591 \cdot 0{,}428}{0{,}817} = 0{,}607;$$

$$b_1 = 0{,}331\frac{57{,}219}{39{,}457} = 0{,}480, \qquad b_2 = 0{,}607\frac{57{,}219}{24{,}549} = 1{,}415;$$

$$a = \overline{y} - b_1\overline{x_1} - b_2\overline{x_2} = -16{,}941$$

Стандартизованная форма уравнения регрессии имеет вид:

$$\hat{t}_y = 0{,}331t_{x_1} + 0{,}607t_{x_2}$$

Естественная форма уравнения регрессии имеет вид:

$$\hat{y} = -16{,}941 + 0{,}480x_1 + 1{,}415x_2$$

Для выяснения относительной силы влияния факторов на результативный признак рассчитываются средние коэффициенты эластичности:

$$\overline{Э_j} = b_j\frac{\overline{x_j}}{\overline{y_j}}$$

$$\overline{Э_1} = 0{,}48\frac{91}{112{,}583} = 0{,}388 \qquad \overline{Э_2} = 1{,}415\frac{60{,}667}{112{,}583} = 0{,}762$$

Следовательно, при увеличении среднегодовой стоимости основных фондов (x_1) на 1 % валовой доход (у) увеличивается на 0,388 % от своего среднего уровня.

При повышении среднегодовой стоимости оборотных средств (x_2) на 1 % валовой доход повышается на 0,762 % от своего среднего уровня.

2. Линейные коэффициенты частной корреляции для уравнения (1) определяются следующим образом:

$$r_{y x_1 x_2} = \frac{r_{yx_1} - r_{yx_2} \cdot r_{x_1x_2}}{\sqrt{(1 - r^2_{yx_2})(1 - r^2_{x_1x_2})}} = 0{,}452$$

$$r_{y x_2 x_1} = \frac{r_{yx_2} - r_{yx_1} \cdot r_{x_1x_2}}{\sqrt{(1 - r^2_{yx_1})(1 - r^2_{x_1x_2})}} = 0{,}680$$

Отличие коэффициентов частной корреляции объясняется не слабой межфакторной связью ($r_{x_1x_2}$ =0,428)

Линейный коэффициент множественной корреляции рассчитывается из формулы:

$$R_{yx_1x_2} = \sqrt{r_{yx_1}\beta_1 + r_{yx_2}\beta_2} = 0{,}807$$

Коэффициент детерминации R^2= 0,651

3. $F_{факт}=\frac{R^2_{yx_1x_2}}{1-R^2_{yx_1x_2}}\times\frac{n-m-1}{m}$,

n – число единиц совокупности,

m – число коэффициентов при переменных x.

$$F_{факт.}=\frac{0{,}651}{0{,}349}\times\frac{9}{2}=8{,}394$$

$F_{табл.}$ (α=0,05 ν_1=2; ν_2=9)=4,26

$F_{факт.} > F_{табл.}$, следовательно, уравнение значимо в целом.

Выясним статистическую значимость каждого фактора во множественном уравнении регрессии.

Для этого рассчитаем частные F- статистики.

$$Fx_{1расч.}=\frac{R^2_{yx_1x_2}-r^2_{yx_2}}{1-R^2_{yx_1x_2}}\times\frac{n-m-1}{m}=2{,}321$$

$F_{табл.}$ (α=0,05 ν_1=1; ν_2=9)=5,12

$Fx_{1расч} < F_{табл.}$

Из неравенства следует вывод о нецелесообразности включения в модель фактора x_1 после фактора x_2

$$Fx_{2расч.}=\frac{R^2_{yx_1x_2}-r^2_{yx_1}}{1-R^2_{yx_1x_2}}\times\frac{n-m-1}{m}=7{,}788$$

$Fx_{2расч} > F_{табл.}$

Из неравенства следует вывод о целесообразности включения в модель фактора x_2 после фактора x_1.

4.Результаты позволяют сделать вывод:

1) о незначимости фактора x_1 и нецелесообразности включения его в уравнение регрессии после фактора x_2.

2) о значимости фактора x_2 и целесообразности включения его в уравнение регрессии.

Значимой оказалась модель: $\hat{y} = -18{,}249 + 1{,}415x_2$

КОНТРОЛЬНОЕ ЗАДАНИЕ 2

Задачи 11-15.

Имеются данные о деятельности крупнейших компаний в течение 12 месяцев 199X года.

Известны – чистый доход (y), оборот капитала (x_1), использованный капитал (x_2) в млрд. у.е.

Таблица 2.1.

Задача 11			Задача 12			Задача 13			Задача 14			Задача 15		
y	x_1	x_2	y	x_1	x_2	y	x_1	x_2	y	x_1	x_2	y	x_1	x_2
5,5	53,1	27,1	6,6	6,9	83,6	3,6	16,2	13,3	1,5	5,9	5,9	3,0	18,0	6,5
2,4	18,8	11,2	3,0	18,0	6,5	1,5	5,9	5,9	5,5	53,1	27,1	3,3	16,7	15,4
3,0	35,3	16,4	6,5	107,9	50,4	5,5	53,1	27,1	2,4	18,8	11,2	3,6	16,2	13,3
4,2	71,9	32,5	3,3	16,7	15,4	2,4	18,8	11,2	3,0	35,3	16,4	5,5	53,1	27,1
2,7	93,6	25,4	0,1	76,6	29,6	3,0	35,3	16,4	4,2	71,9	32,5	3,0	35,3	16,4
1,6	10,0	6,4	3,6	16,2	13,3	4,2	71,9	32,5	2,7	93,6	25,4	2,7	93,6	25,4
2,4	31,5	12,5	2,4	18,8	11,2	2,7	93,6	25,4	1,6	10,0	6,4	2,4	31,5	12,5
3,3	36,7	14,3	3,0	35,3	16,4	1,6	10,0	6,4	2,4	31,5	12,5	1,8	13,8	6,5
1,8	13,8	6,5	1,8	13,8	6,5	2,4	31,5	12,5	3,3	36,7	14,3	1,6	30,4	15,8
2,4	64,8	22,7	2,4	64,8	22,7	3,3	36,7	14,3	1,8	13,8	6,5	0,9	31,3	18,9
1,6	30,4	15,8	1,6	30,4	15,8	1,8	13,8	6,5	2,4	64,8	22,7	6,5	107,9	50,4
1,4	12,1	9,3	1,4	12,1	9,3	2,4	64,8	22,7	1,6	30,4	15,8	3,6	16,2	13,3

Задачи 16-20.

Имеются данные 12 месяцев по району города о рынке вторичного жилья, (y – стоимость квартиры, тыс. у.е., x_1 – размер жилой площади, $м^2$, x_2 – размер кухни, $м^2$).

Таблица 2.2.

Задача 16			Задача 17			Задача 18			Задача 19			Задача 20		
y	x_1	x_2	y	x_1	x_2	y	x_1	x_2	y	x_1	x_2	y	x_1	x_2
13,0	37,0	6,2	13,2	46,0	5,8	23,0	22,8	5,0	22,5	37,2	7,6	22,7	28,8	5,4
16,4	60,9	10,0	15,9	54,1	8,5	26,8	27,7	5,2	25,5	58,0	9,4	25,8	36,2	7,2
17,0	60,0	8,5	16,2	50,6	8,0	28,0	34,5	6,0	19,2	60,2	9,5	20,8	28,9	5,6
15,2	52,1	7,4	15,4	43,8	5,2	18,4	26,4	5,1	13,6	52,0	8,1	15,2	32,4	6,4
14,2	40,1	7,0	14,2	78,6	12,0	30,4	19,8	4,8	25,4	44,6	7,4	25,4	49,7	7,5
10,5	30,4	6,2	11,0	60,2	7,2	20,8	17,9	4,5	17,8	31,2	6,3	19,4	38,1	6,7
20,0	43,0	7,5	21,1	50,2	7,0	22,4	25,2	5,4	18,0	26,4	5,9	18,2	30,2	6,2
12,0	32,1	6,4	13,4	54,7	7,3	21,8	20,4	4,9	21,1	20,7	5,5	21,0	32,6	6,4
15,6	35,1	7,0	15,6	42,8	5,5	18,5	20,7	5,0	16,5	22,4	5,7	16,4	27,5	5,5
12,5	32,0	6,2	12,8	60,4	7,3	23,5	21,4	5,2	23,0	35,4	6,8	23,5	39,0	6,9

13,2	33,0	6,0	14,5	47,2	5,8	16,7	19,6	4,5	16,2	28,4	6,5	18,8	27,5	5,4
14,6	32,5	5,8	15,1	40,6	5,2	20,4	24,5	4,9	17,2	22,7	6,0	17,5	31,2	6,3

Задание:

1. Рассчитайте параметры линейного уравнения множественной регрессии.

2. Дайте оценку силы связи факторов с результатом с помощью средних коэффициентов эластичности.

3. Оцените статистическую значимость параметров и уравнения регрессии в целом с помощью соответственно критериев Стьюдента и Фишера (α=0,01).

4. Рассчитайте среднюю ошибку аппроксимации. Сделайте вывод.

5. Составьте матрицы парных и частных коэффициентов корреляции и укажите информативные факторы.

6. Оцените полученные результаты, выводы оформите в аналитической записке.

ГЛАВА 3. СИСТЕМЫ ЭКОНОМЕТРИЧЕСКИХ УРАВНЕНИЙ

Система взаимосвязанных регрессионных уравнений и тождеств, в которых одни и те же переменные в различных регрессионных уравнениях могут одновременно выступать и в роли результирующих показателей, и в роли объясняющих переменных называют системой одновременных (эконометрических уравнений).

Среди переменных систем различают как эндогенные (внутрисистемные), так и экзогенные (внешние по отношению к рассматриваемой системе).

Рассматриваются различные виды систем одновременных уравнений. Система вида:

$$\begin{cases} y_1 = b_{12}y_2 + b_{13}y_3 + ... + b_{1m}y_m + a_{11}x_1 + a_{12}x_2 + ... + a_{1k}x_k + \varepsilon_1 \\ y_2 = b_{21}y_1 + b_{23}y_3 + ... + b_{2m}y_m + a_{21}x_1 + a_{22}x_2 + ... + a_{2k}x_k + \varepsilon_2 \\ ..... \\ y_m = b_{m1}y_1 + b_{m2}y_2 + ... + b_{mm-1}y_{m-1} + a_{m1}x_1 + ... + a_{mk}x_k + \varepsilon_m \end{cases} \quad (3.1)$$

называется системой совместных одновременных уравнений или структурной формой модели.

Система вида:

$$\begin{cases} y_1 = a_{11}x_1 + ... + a_{1k}x_k + \varepsilon_1 \\ ..... \\ y_m = a_{m1}x_1 + ... + a_{mk}x_k + \varepsilon_m \end{cases} \quad (3.2)$$

называется системой независимых уравнений.

МНК, использованной для определения параметров (коэффициентов) системы (3.1) дает смещенные и несостоятельные оценки. Поэтому для определения коэффициентов системы (3.1) – структурных коэффициентов, систему (3.1) приводят к виду:

$$\begin{cases} \hat{y}_1 = \delta_{11}x_1 + \delta_{12}x_2 + ... + \delta_{1k}x_k \\ ..... \\ \hat{y}_m = \delta_{m1}x_1 + \delta_{m2}x_2 + ... + \delta_{mk}x_k \end{cases} \quad (3.3)$$

Система (3.3) называется приведенной формой модели.

Параметры системы (3.3) могут быть оценены МНК.

Оценки, полученные по МНК могут использованы для оценивания структурных параметров (косвенный метод наименьших квадратов-КМНК).

При этом возможны ситуации:

- однозначное выражение структурных коэффициентов через коэффициенты приведенной модели;
- структурный коэффициент допускает несколько разных оценок КМНК;
- структурный коэффициент не может быть выражен через коэффициенты приведенной модели.

В последнем случае структурное уравнение является не идентифицированным.

Необходимое условие идентифицируемости модели:

$D+1=H$ – уравнение идентифицировано;

$D+1\langle H$ - уравнение неидентифицировано;

$D+1\rangle H$ - уравнение сверхидентифицировано,

где H – число эндогенных переменных в уравнении, D – число предопределенных (экзогенных и лаговых) переменных, которые содержаться в системе, но не входят в данное уравнение.

Достаточным условием идентифицируемости является следующее:

- уравнение идентифицируемо, если по отсутствующим в нем переменным можно из коэффициентов при них в других уравнениях системы получить матрицу, ранг которой не меньше числа эндогенных переменных в системе минус единица.
- система, в которой для каждого уравнения выполнены необходимые и достаточные условия – точно идентифицируема.

Система, в которой каждое уравнение идентифицировано, идентифицируема. Для оценки ее параметров используют косвенный метод наименьших квадратов (КМНК).

Система в которой хотя бы одно уравнение сверхидентифицировано, а остальные идентифицируемы, сверхидентифицирована.

Для оценки параметров сверхидентифицированной модели используется двухшаговый метод наименьших квадратов (ДМНК).

Система, в которой хотя бы одно уравнение не идентифицировано, не идентифицирована. В этом случае необходимо или менять спецификацию системы или использовать трехшаговый метод наименьших квадратов (ТМНК).

Замечание.

Если в модели есть тождества, то статистические исследования исключают их**.**

ПРИМЕР 3

1. Оценить структурную модель на идентификацию:

$$Y_1=b_{12}y_2+a_{11}x_1+a_{12}x_2$$
$$Y_2=b_{21}y_1+b_{23}y_3+a_{22}x_2 \qquad (1)$$
$$Y_3=b_{32}y_2+a_{31}x_1+a_{33}x_3$$

2. На основе приведенной модели

$$Y_1=2x_1+4x_2+10x_3$$
$$Y_2=3x_1-6x_2+2x_3 \qquad (2)$$
$$Y_3=-5x_1+8x_2+5x_3$$

найти структурные коэффициенты модели.

1. Модель имеет 3 эндогенные ($y_1y_2y_3$) и 3 экзогенные переменные ($x_1x_2x_3$).

Проверим необходимое условие идентификации:

D+1=H уравнение идентифицировано

D+1<H уравнение не идентифицировано

D+1>H уравнение сверх идентифицировано

где H – число эндогенных переменных

D – число предопределенных (экзогенных и лаговых), отсутствующих в уравнении, но присутствующих в системе.

1 уравнение: D=1(x_3), H=2 (y_1,y_2), D+1=H уравнение идентифицировано

2 уравнение: D=2 (x_1, x_3), H=3 (y_1,y_2,y_3), D+1=H уравнение идентифицировано

3 уравнение: D=1(x_2), H=2 (y_2,y_3), D+1=H уравнение идентифицировано

Следовательно, необходимое условие идентифицируемости выполнено.

Проверим достаточное условие:

В 1 уравнении нет переменных x_3,y_3

Строим матрицу:

	x_3	y_3
2 ур.	0	b_{23}
3 ур.	a_{33}	-1

det M = det $\begin{pmatrix} 0 & b_{23} \\ a_{33} & -1 \end{pmatrix} = \begin{vmatrix} 0 & b_{23} \\ a_{33} & -1 \end{vmatrix} \neq 0$, rang M = 2

Во 2 уравнении нет переменных x_1, x_3

Строим матрицу:

	x_1	x_3
1 ур.	a_{11}	0
3 ур.	a_{31}	a_{33}

det M = det $\begin{pmatrix} a_{11} & 0 \\ a_{31} & a_{33} \end{pmatrix} = \begin{vmatrix} a_{11} & 0 \\ a_{31} & a_{33} \end{vmatrix} \neq 0$, rang M = 2

В 3 уравнении нет переменных y_1, x_2

Строим матрицу:

	y_1	x_2
1 ур.	-1	a_{12}
2 ур.	b_{21}	a_{22}

det M = det $\begin{pmatrix} -1 & a_{12} \\ b_{21} & a_{22} \end{pmatrix} = \begin{vmatrix} -1 & a_{12} \\ b_{21} & a_{22} \end{vmatrix} \neq 0$, rang M = 2

Следовательно, достаточное условие идентифицируемости выполнено. Система точно идентифицируема.

2. Находим структурные коэффициенты модели.

Для этого:

1) Находим x_3 из 2 уравнения приведенной системы (2) и подставляем ее (x_3) в 1 уравнение системы (2).

$Y_1=2,6x_1-6,8x_2+0,2y_2$

2) Из 3 уравнения системы (2) находим x_3 и подставляем во 2 уравнение системы.

$Y_2=12x_1-8x_2+2y_3$

из системы

$Y_1=2,6x_1-6,8x_2+0,2y_2$

$Y_2=12x_1-8x_2+2y_3$

исключаем x_1 и определяем y_2: $Y_2=2,4y_1+12,16x_2+1,04y_3$

x_2 выразим из последнего уравнения и подставим в 3 уравнение системы (2).

$Y_3=-8x_1+1,5y_2-10x_3$

Выводы:

Структурная модель получена:

$Y_1=2,6x_1-6,8x_2+0,2y_2$

$Y_2=12,16x_2+2,4y_1++1,04y_3$

$Y_3= -8x_1-10x_3+1,5y_2$

Замечание. К приведенной модели можно применить МНК. Для точно идентифицируемой структурной модели, возможно осуществить переход к приведенной форме. Эти действия составляют косвенный метод наименьших квадратов (КМНК).

КОНТРОЛЬНОЕ ЗАДАНИЕ 3

Задача 21

Гипотетическая модель экономики:
$C_t=a_1+b_{11}Y_t+b_{12}\ Y_t+\varepsilon_1$
$J_t=a_2+b_{21}Y_{t-1}+\varepsilon_2$
$T_t=a_3+b_{31}Y_t++\varepsilon_3$
$Y_t=C_t+Y_t+G_t$,
где C – совокупное потребление в период t;
Y – совокупный доход в период t;
J – инвестиции в период t;
T – налоги в период t;
G – государственные доходы в период t.

Задача 22

Модель спроса и предложения на деньги:
$R_t=a_1+b_{11}M_t+b_{12}\ Y_t+\varepsilon_1$
$Y_t=a_2+b_{21}R_t+\varepsilon_2$,
где R – процентные ставки в период t;
Y – ВВП в период t;
M- денежная масса в период t

Задача 23

Макроэкономическая модель:
$C_t=a_1+b_{12}Y_t+b_{13}\ T_t+\varepsilon_1$
$I_t=a_2+b_{21}Y_t+b_{24}\ K_{t-1}+\varepsilon_1$
$Y_t= C_t+I_t$,
где C – потребление;
I - инвестиции
Y – доход
T - налоги
K- запас капитала

Задача 24

Модель денежного и товарного рынков:
$R_t=a_1+b_{12}Y_t+b_{14}\ M_t+\varepsilon_1$
$Y_t=a_2+b_{21}R_t+b_{23}\ I_t+b_{25}G_t+\varepsilon_2$
$I_t=a_3+b_{31}R_t+\varepsilon_3$,
где R – процентные ставки;
Y – реальный ВВП
M – денежная масса
I – внутренние инвестиции
G- реальные государственные расходы

Задача 25

Модель денежного рынка:
$R_t=a_1+b_{11}M_t+b_{12}\ Y_t+\varepsilon_1$
$Y_t=a_2+b_{21}R_t+b_{22}\ I_t+\varepsilon_2$
$I_t=a_3+b_{33}R_t+\varepsilon_t$,
где R – процентные ставки
Y - ВВП
M – денежная масса
I – внутренние инвестиции.

Задача 26

Модель имеет вид:
$Y_1=a_1+b_{12}Y_2+\varepsilon_1$
$Y_2=a_2+b_{21}Y_1+C_{21}\ X_1+\varepsilon_2$
$Y_3=Y_2+X_2$

Задача 27

Модель имеет вид:
$Y_1=a_1+b_{11}x_1+b_{12}\ x_2+C_{12}Y_2+\varepsilon_1$
$Y_2=a_2+b_{22}x_2+C_{21}Y_1+\varepsilon_2$
$Y_3=a_3+b_{31}x_1+b_{33}\ x_3+\varepsilon_3$

Задача 28

Модель имеет вид:
$Y_1=a_1+b_{11}x_1+b_{13}\ x_3+C_{12}Y_2+\varepsilon_1$

$Y_2=a_2+b_{22}x_2+C_{21}Y_1+\varepsilon_2$
$Y_3=a_3+b_{32}x_2+b_{33}\ x_3+\varepsilon_3$

Задача 29

Модель имеет вид:
$Y_1=b_{12}y_2+a_{11}\ x_1+a_{12}x_2+\varepsilon_1$
$Y_2=b_{21}y_1+b_{23}\ Y_3+a_{22}x_2+\varepsilon_2$
$Y_3=b_{31}y_1+a_{31}\ x_1+a_{33}x_3+\varepsilon_3$

Задача 30

Модель имеет вид:
$Y_1=b_{12}y_2+a_{11}\ x_1+a_{12}x_2+\varepsilon_1$
$Y_2=b_{21}y_1+a_{22}\ x_2+a_{23}x_3+\varepsilon_2$
$Y_3=b_{31}y_1+a_{33}\ x_3+\varepsilon_3$

Задание:

1. Используя необходимое и достаточное условие идентификации, определить, идентифицировано ли каждое уравнение модели.
2. Определите тип модели.
3. Определите метод оценки параметров модели.
4. Опишите последовательность действий при использовании указанного метода.
5. Результаты оформите в виде пояснительной записки.

ГЛАВА 4. ВРЕМЕННЫЕ РЯДЫ (ВР)

При построении эконометрической модели можно использовать данные:
- описывающие совокупность различных объектов в определенный момент (период) времени;
- описывающие один объект за ряд последовательных моментов (периодов) времени.

Соответственно определяются и модели:
- пространственные модели;
- модели временных рядов (или динамические ряды).

Для оценки поведения временных рядов значения ряда разделяют на составляющие:
- Тренд;
- Циклические колебания;

- Сезонные колебания;
- Случайные колебания.

Тренд рассматривают в качестве общей направленности изменений определенных значений, взятых на протяженном отрезке времени.

Циклические колебания. Эти составляющие показывают колебания относительно линии тренда для периодов выше одного года. Сезонные колебания показывают периодичность колебаний на протяжении года и более. Сезонные колебания можно выявить после анализа тренда и циклических колебаний. Случайные колебания представляют собой случайные величины. Их выявляют путем снятия тренда, циклических и сезонных колебаний для данного значения. И хотя такое значение нельзя предугадать заранее, его целесообразно учитывать при определении вероятной точности принятой модели прогнозирования.

4.1. Моделирование тенденции временного ряда (выделение тренда)

Наиболее распространенный способ моделирования тенденции ВР является построение функции, характеризующей зависимость уровней ряда от времени или тренда. Этот прием называют аналитическим выравниванием ВР.

Для построения тренда чаще других используют функции:

$$\hat{y}_t = a + bt;$$

$$\hat{y}_t = ab^t;$$

$$\hat{y}_t = a + \frac{b}{t};$$

$$\hat{y}_t = at^b;$$

$$\hat{y}_t = a + b_1 t + \ldots + b_k t^k .$$

Параметры трендов можно определить МНК, в котором в роли объясняющей переменной выступает время t=1,2,…,n, а в качестве зависимой переменной – фактические уровни временного ряда y_t.

Тип тенденции можно определить несколькими способами, например:

- качественным анализом изучаемого процесса;
- расчетом основных показателей динамики;
- расчетом и сравнением коэффициентов автокорреляции уровней ВР.

Согласно последнему подходу, если ВР имеет линейную тенденцию, то его соседние уровни y_t и y_{t-1} тесно коррелируют, коэффициент автокорреляции первого порядка по логарифмам натуральным уровней исходного ряда будет выше, чем соответствующий коэффициент, рассчитанный по уровням ряда.

4.2. Моделирование циклических и сезонных колебаний

Анализ структуры ВР, содержащего циклические и сезонные колебания, можно осуществить расчетом сезонной компоненты методом скользящей средней и построением аддитивной (Y=T+S+E) или мультипликативной (Y=T×S×E) составляющих.

Модели предполагают, что каждый уровень ВР может быть представлен с помощью трендовой (T), сезонной (S) и случайной (E) компонент.

Выбор модели производится на основе анализа структуры сезонных колебаний. Если амплитуда колебаний возрастает или уменьшается – строят мультипликативную модель, если амплитуда приблизительно постоянна, строят аддитивную модель. Построение моделей сводится к расчету значений T, S и E для каждого уровня ряда.

Моделирование циклических колебаний в основном осуществляется аналогично моделированию сезонных колебаний.

ПРИМЕР 4

Имеются данные об урожайности овощных культур в хозяйствах региона. Данные приведены в таблице 4.1.

Таблица 4.1

Год	1989	1990	1991	1992	1993	1994	1995	1996
Урожайность т/га	10,2	10,7	11,7	13,1	14,9	17,2	20,0	23,2

Для данного временного ряда требуется:
1. Обосновать тип уравнения тренда.
2. Рассчитать параметры тренда.
3. Дать прогноз на 1997 год.

Расчеты приведем в таблице 4.2.

Решение:

Определим коэффициент корреляции между рядами y_t и y_{t-1}.

$$r_1 = \frac{\sum_{t=2}^{n}(y_t - \overline{y_1})(y_{t-1} - \overline{y_2})}{\sqrt{\sum_{t=2}^{n}(y_t - \overline{y_1})^2 \times \sum_{t=2}^{n}(y_{t-1} - \overline{y_2})^2}}$$

где $\overline{y_1} = \frac{\sum_{t=2}^{n} y_t}{n-1}$; $\overline{y_2} = \frac{\sum_{t=2}^{n} y_{t-1}}{n-1}$

$$\overline{y_1} = \frac{10,7 + 11,7 + 13,1 + 14,9 + 17,2 + 20,0 + 23,2}{7} = \frac{110,8}{7} = 15,828$$

$$\overline{y_2} = \frac{10,2 + 10,7 + 11,7 + 13,1 + 14,9 + 17,2 + 20,0 +}{7} = \frac{97,8}{7} = 13,971$$

$$r_1 = \frac{99,035}{99,149} = 0,998$$

Результат говорит о тесной зависимости между урожайностью текущего и непосредственно предшествующего годов и наличии во временном ряде сильной линейной тенденции.

Таблица 4.2

	Год	y_t	y_{t-1}	y_{t-2}	$(y_t - \bar{y}_1)^2$	$(y_{t-1} - \bar{y}_2)^2$	$y_t - \bar{y}_3$	$y_{t-2} - \bar{y}_4$	$(y_t - \bar{y}_3)^2$	$(y_{t-2} - \bar{y}_4)^2$	$(y_t - \bar{y}_1) \times \times (y_{t-1} - \bar{y}_2)$	$(y_t - \bar{y}_3) \times \times (y_{t-2} - \overline{y_4})$
	1	10,2	-	-	-	-	-	-	-	-	-	-
	2	10,7	10,2	-	26,296	14,220	-	-	-	-	19,338	-
	3	11,7	10,7	10,2	17,040	10,699	-4,983	-2,767	24,830	7,656	13,503	13,788
	4	13,1	11,7	10,7	7,442	5,157	-3,583	-2,267	12,838	5,139	6,195	8,123
	5	14,9	13,1	11,7	0,861	0,759	-1,783	-1,267	3,179	1,605	0,808	2,259
	6	17,2	14,9	13,1	1,882	0,863	0,517	0,133	0,267	0,018	1,274	0,069
	7	20,0	17,2	14,9	47,416	10,426	3,317	1,933	11,002	3,736	13,471	6,412
	8	23,2	20,0	17,2	54,346	36,349	6,517	4,233	42,471	17,917	44,446	27,586
Σ	36				125,273	78,773	0,002	-0,002	94,587	36,072	99,035	58,237
Средн.	4,5	$\bar{y}_1 = 15{,}828$ $\bar{y}_3 = 16{,}683$	$\bar{y}_2 = 13{,}971$	$\bar{y}_4 = 12{,}967$								

Определим коэффициент автокорреляции второго порядка.

$$r_2 = \frac{\sum_{t=3}^{n}(y_t - \overline{y_3})(y_{t-2} - \overline{y_4})}{\sqrt{\sum_{t=3}^{n}(y_t - \overline{y_3})^2 \times \sum_{t=3}^{n}(y_{t-2} - \overline{y_4})^2}}$$

где $\overline{y_3} = \frac{\sum_{t=3}^{n} y_t}{n-2}$; $\overline{y_4} = \frac{\sum_{t=3}^{n} y_{t-2}}{n-2}$

$$\overline{y_3} = \frac{11{,}7 + 13{,}1 + 14{,}9 + 17{,}2 + 20{,}0 + 23{,}2}{6} = 16{,}683$$

$$\overline{y_4} = \frac{10{,}2 + 10{,}7 + 11{,}7 + 13{,}1 + 14{,}9 + 17{,}2}{6} = 12{,}967$$

$$r_2 = \frac{58{,}237}{58{,}412} = 0{,}997$$

Результаты подтверждает наличие линейной тенденции. Выбираем линейное уравнение тренда:

$$\hat{y} = a + bt$$

3. Параметры тренда определим, используя МНК. Результаты расчетов приведены в таблице 4.3.

Таблица 4.3

	t	y_t	t^2	y_t^2	t y_t	$t - \bar{t}$	$(t - \bar{t})^2$	$\hat{y}_t$	$y - \hat{y}_t$	$(y - \hat{y}_t)^2$
	1	10,2	1	104,04	10,2	-3,5	12,25	8,800	1,400	1,960
	2	10,7	4	114,49	21,4	-2,5	6,25	10,607	0,093	0,009
	3	11,7	9	136,89	35,1	-1,5	2,25	12,414	-0,714	0,09
	4	13,1	16	171,61	52,4	-0,5	0,25	14,221	-1,121	1,257
	5	14,9	25	222,01	72,5	0,5	0,25	16,028	-1,128	1,272
	6	17,2	36	295,84	103,2	1,5	2,25	17,835	-0,635	0,403
	7	20,0	49	400,00	140	2,5	6,25	19,642	0,358	0,128
	8	23,2	64	538,24	185,6	3,5	12,25	21,449	1,751	3,066
Σ	36	121	204	1983,12	620,4	0,00	32,00			8,604
Средн.	4,5	15,125	25,5	247,890	77,55					

$$\sigma_t^2 = \overline{t^2} - (\bar{t})^2 = 5{,}25, \quad \sigma_t = 2{,}291$$

$$\sigma_{y_t}^2 = \overline{y_t^2} - (\overline{y_t})^2 = 19{,}124 \quad \sigma_{y_t} = 4{,}373$$

$$b = \frac{\overline{ty_t} - \bar{t} \cdot \overline{y_t}}{\sigma_t^2} = 1{,}807$$

$$a = \bar{y} - b\bar{t} = 15{,}125 - 1{,}807 \cdot 4{,}5 = 6{,}993$$

$$\hat{y}_t = 6{,}993 + 1{,}807t$$

$$r = b\frac{\sigma_t}{\sigma_{y_t}} = 0{,}947$$

Расчетное значение критерия Фишера равно Fрасч.= 52,25;

Fтабл.(α =0,01; ν1=1; ν2=6) = 13,74, следовательно, уравнение статистически значимо, прогноз имеет смысл.

Прогнозное значение tпрогн.=9, тогда $\hat{y}_t = 23{,}256$, средняя ошибка прогноза

$$m_{\hat{y}_{прогн.}} = \sqrt{D_{ост.} \cdot \left(1 + \frac{(t_{прогн.} - \bar{t})^2}{\Sigma(t - \bar{t})^2}\right)}$$

где остаточная дисперсия $D_{ост}$ рассчитывается по формуле: $D_{ост.}$=

$$\frac{\Sigma(y_t - \hat{y}_t)^2}{n-2};\ m_{\hat{y}_{t_{прогн.}}} = 0{,}854$$

Предельная ошибка прогноза $\Delta\hat{y}_{t_{прогн.}} = t_{табл.} \cdot m_{\hat{y}_{t_{прогн.}}}$,

$t_{табл.}$(α =0,01; ν=6) = 3,707, $\Delta\hat{y}_{t_{прогн.}} = 3{,}165$

Доверительный интервал прогноза (20,091;26,494)

Выполненный прогноз надежный и достаточно точный. Следовательно, применительно к рассмотренному временному ряду темпы роста урожайности за восемь рассмотренных лет изменялись от 6,993 (т/га) со средним за год приростом, равным 1,807 (т/га).

КОНТРОЛЬНОЕ ЗАДАНИЕ 4

Задачи 31-35.

Имеются данные за 15 дней по количеству пациентов клиники, прошедших через соответствующие отделения в течении дня

Таблица 4.4

День	Задача 31	Задача 32	Задача 33	Задача 34	Задача 35
	Терапевтическое отделение	Хирургическое отделение	Стоматологическое отделение	Гласное отделение	Отделение пластической хирургии
1	29	35	41	30	22
2	40	29	52	22	19
3	30	22	30	19	11
4	52	19	47	28	12

5	47	30	28	24	16
6	28	47	22	18	28
7	16	28	51	35	30
8	51	12	40	29	18
9	40	13	57	40	17
10	35	15	33	34	20
11	57	18	43	31	21
12	28	19	51	29	19
13	33	20	36	35	24
14	42	16	19	23	13
15	39	35	42	27	16

Задачи 36-40.

Имеются данные за 12 лет по странам о годовом объеме продаж автомобилей.

Таблица 4.5

Год Объем продаж 100 тыс.	Задача 36 Страна А	Задача 37 Страна В	Задача 38 Страна С	Задача 39 Страна Д	Задача 40 Страна Е
1986	3,8	4,1	5,2	2,8	4,2
1987	4,7	5,2	6,3	3,6	5,4
1988	3,9	4,3	4,5	2,7	4,0
1989	2,7	3,2	3,9	2,0	3,1
1990	2,9	3,0	3,8	1,8	2,9
1991	2,3	2,8	3,0	1,4	2,4
1992	3,0	4,2	4,8	2,1	3,7
1993	3,6	4,6	5,0	2,5	4,1
1994	2,9	3,7	4,6	2,1	1,4
1995	3,7	4,8	6,1	3,0	2,2
1996	4,5	5,6	6,7	3,7	2,9
1997	4,2	5,0	6,9	3,1	2,6

Требуется:

1. Определить коэффициенты автокорреляции уровней ряда первого и второго порядка.
2. Обоснуйте выбор уравнение тренда и определите его параметры.
3. Сделайте выводы.
4. Результаты оформите в виде пояснительной записки.

Приложение 1

Распределение Фишера-Снедекора (F-распределение)

Значения $F_{табл}$, удовлетворяющие условию P ($F > F_{табл}$). Первое значение соответствует вероятности 0,05, второе - вероятности 0,01 и третье - вероятности 0,001, где v_1 - число степеней свободы числителя, а v_2 - число степеней свободы знаменателя.

v_2 \ v_1	1	2	3	4	5	6	8	12	24	∞	t
1	161,4	199,5	215,7	224,6	230,2	234,0	238,9	243,9	249,0	253,3	12,71
	4052	4999	5403	5625	5764	5859	5981	6106	6234	6366	63,66
	406523	500016	536700	562527	576449	585953	598149	610598	623432	636535	636
2	18,51	19,00	19,16	19,25	19,30	19,33	19,37	19,41	19,45	19,50	4,30
	98,49	99,01	00,17	99,25	99,30	99,33	99,36	99,42	99,46	99,50	9,92
	998,46	999,00	999,20	999,20	999,20	999,20	999,40	999,60	999,40	999,40	31,00
3	10,13	9,55	9,28	9,12	9,01	8,94	8,84	8,74	8,64	8,53	3,18
	34,12	30,81	29,46	28,71	28,24	27,91	27,49	27,05	26,60	26,12	5,84
	67,47	148,51	141,10	137,10	134,60	132,90	130,60	128,30	125,90	123,50	12,94
4	7,71	6,94	6,59	6,39	6,26	6,16	6,04	5,91	5,77	5,63	2,78
	21,20	18,00	16,69	15,98	15,52	15,21	14,80	14,37	13,93	13,46	4,60
	74,13	61,24	56,18	53,43	51,71	50,52	49,00	47,41	45,77	44,05	8,61
5	6,61	5,79	5,41	5,19	5,05	4,95	4,82	4,68	4,53	4,36	2,57
	16,26	13,27	12,06	11,39	10,97	10,67	10,27	9,89	9,47	9,02	4,03
	47,04	36,61	33,20	31,09	20,75	28,83	27,64	26,42	25,14	23,78	6,86

Продолжение

v_2 \ v_1	1	2	3	4	5	6	8	12	24	∞	t
6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4,00	3,84	3,67	2,45
	13,74	10,92	9,78	9,15	8,75	8,47	8,10	7,72	7,31	6,88	3,71
	35,51	26,99	23,70	21,90	20,81	20,03	19,03	17,99	16,89	15,75	5,96
7	5,59	4,74	4,35	4,12	3,97	3,87	3.73	3,57	3,41	3,23	2,36
	12,25	9,55	8,45	7,85	7,46	7,19	6,84	6,47	6,07	5,65	3,50
	29,22	21,69	18,77	17,19	16,21	15,52	14,63	13,71	12,73	11,70	5,40
8	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,28	3,12	2,99	2,31
	11,26	8,65	7,59	7,10	6,63	6.37	6,03	5,67	5,28	4,86	3,36
	25,42	18,49	15,83	14,39	13,49	12,86	12,04	11,19	10,30	9,35	5,04
9	5,12	4,26	3,86	3,63	3,48	3,37	3,23	3,07	2,90	2,71	2,26
	10,56	8,02	6,99	6,42	6,06	5,80	5,47	5,11	4,73	4,31	3,25
	22,86	16,39	13,90	12,56	11,71	11,13	10,37	9,57	8,72	7,81	4,78
10	4,96	4,10	3,71	3,48	3,33	3,22	3,07	2,91	2,74	2,54	2,23
	10,04	7,56	6,55	5,99	5,64	5,39	5,06	4,71	4,33	3,91	3,17
	21,04	14,91	12,55	11,28	10,48	9,92	9,20	8,45	7,64	6,77	4,59
11	4,84	3,98	3,59	3,36	3,20	3,09	2,95	2,79	2,61	2,40	2,20
	9,65	7,20	6,22	5,67	5,32	5,07	4,74	4,40	4,02	3,60	3,11
	19,69	13,81	11,56	10,35	9,58	9,05	8,35	7,62	6,85	6,00	4,49
12	4,75	3,88	3,49	3,26	3,11	3,00	2,85	2,69	2,50	2,30	2,18
	9,33	6,93	5,95	5,41	5,06	4,82	4,50	4,16	3,78	3,36	3,06
	18,64	12,98	10,81	9,63	8,89	8,38	7,71	7,00	6,25	5,42	4,32

Продолжение

v_2 \ v_1	1	2	3	4	5	6	8	12	24	∞	t
13	4,67	3,80	3,41	3,18	3,02	2,92	2,77	2,60	2,42	2,21	2,16
	9,07	6,70	5,74	5,20	4,86	4,62	4,30	3,96	3,59	3,16	3,01
	17,81	12,31	10,21	9,07	8,35	7,86	7,21	6,52	5,78	4,97	4,12
14	4,60	3,74	3,34	3,11	2,96	2,85	2,70	2,53	2,35	2,13	2,14
	8,86	6,51	5,56	5,03	4,69	4,46	4,14	3,80	3,43	3,00	2,98
	17,14	11,78	9,73	8,62	7,92	7,44	6,80	6,13	5,41	4,60	4,14
15	4,45	3,68	3,29	3,06	2,90	2,79	2,64	2,48	2,29	2,07	2,13
	8,68	6,36	5,42	4,89	4,56	4,32	4,00	3,67	3,29	2,87	2,95
	16,59	11,34	9,34	8,25	7,57	7,09	6,47	5,81	5,10	4,31	4,07
16	4,41	3,63	3,24	3,01	2,85	2,74	2,59	2,42	2,24	2,01	2,12
	8,53	6,23	5,29	4,77	4,44	4,20	3,89	3.55	3,18	2,75	2,92
	16,12	10,97	9,01	7,94	7,27	6,80	6,20	5,55	4,85	4,06	4,02
17	4,45	3,59	3,20	2,96	2,81	2,70	2,55	2,38	2,19	1,96	2,11
	8,40	6,11	5,18	4,67	4,34	4,10	3,79	3,45	3,08	2,65	2,90
	15,72	10,66	8,73	7,68	7,02	6,56	5,96	5,32	4,63	3,85	3,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,51	2,34	2,15	1,92	2,10
	8,28	6,01	5,09	4,58	4,25	4,01	3,71	3,37	3,01	2,57	2,88
	15,38	10,39	8,49	7,46	8,49	6,35	5,76	5,13	4,45	3,67	3,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,48	2,31	2,11	1,88	2,09
	8,18	5,93	5,01	4,50	4,17	3,94	3,63	3,30	2,92	2,49	2,86
	15,08	10,16	8,28	7,26	6,61	6,18	5,59	4,97	4,29	3,52	3,88

Продолжение

v_2 \ v_1	1	2	3	4	5	6	8	12	24	∞	t
20	4,35	3,49	3,10	2,87	2,71	2,60	2,45	2,28	2,08	1,84	2,09
	8,10	5,85	4,94	4,43	4,10	3,87	3,56	3,23	2,86	2,42	2,84
	14,82	9,95	8,10	7,10	6,46	6,02	5,44	4,82	4,15	3,38	3,85
21	4,32	3,47	3,07	2,84	2,68	2,57	2,42	2,25	2,05	1,82	2,08
	8,02	5,78	4,87	4,37	4 04	3,81	3,51	3,17	2,80	2,36	2,83
	14,62	9,77	7,94	6,95	6,32	5,88	5,31	4,70	4,03	3,26	3,82
22	4,30	3,44	3,05	2.82	2,66	2,55	2,40	2,23	2,03	1,78	2,07
	7,94	5,72	4,82	4,31	3 99	3,75	3,45	3,12	2,75	2,30	2,82
	14,38	9,61	7,80	6,81	6,19	5,76	5,19	4,58	3,92	3,15	3.79
23	4,28	3,42	3,03	2,80	2,64	2,53	2,38	2,20	2,00	1,76	2,07
	7,88	5,66	4,76	4,26	3 94	3,71	3,41	3,07	2,70	2,26	2,81
	14,19	9,46	7,67	6,70	6,08	5,56	5,09	4,48	3,82	3,05	3,77
24	4,26	3,40	3,01	2,78	2,62	2,51	2,36	2,18	1.98	1,73	2,06
	7,82	5,61	4,72	4,22	3 90	3,67	3,36	3,03	2,66	2,21	2,80
	14,03	9,34	7,55	6,59	5,98	5,55	4,99	4,39	3,74	2,97	3,75
25	4,24	3,38	2,99	2,76	2,60	2,49	2,34	2,16	1,96	1,71	2,06
	7,77	5,57	4,68	4,18	3,86	3,63	3,32	2,99	2,62	2,17	2,79
	13,88	9,22	7,45	6,49	5,89	5,46	4,91	4,31	3,66	2,89	3,72
26	4,22	3,37	2,98	2,74	2,59	2,47	2,32	2,15	1,95	1,69	2,06
	7,72	5,53	4,64	4,14	3,82	3,59	3,29	2,96	2,58	2,13	2,78
	13,74	9,12	7,36	6,41	5.80	5,38	4,83	4,24	3,59	2,82	3,71

Продолжение

v_2 \ v_1	1	2	3	4	5	6	8	12	24	∞	t
27	4,21	3,35	2,96	2,73	2,57	2,46	2,30	2,13	1,93	1,67	2,05
	7,68	5,49	4,60	4,11	3,78	3,56	3,26	2,93	2,55	2,10	2,77
	13,61	9,02	7,27	6,33	5,73	5,31	4,76	4,17	3,52	2,76	3,69
28	4,19	3,34	2,95	2,71	2,56	2,44	2,29	2,12	2,93	1,65	2,05
	7,64	5,54	4,57	4,07	3,75	3,53	3,23	2,90	4,17	2,06	2,76
	13,50	8,93	7,18	6,25	5,66	5,24	4,69	4,11	2,12	2,70	3,67
29	4,18	3,33	2,93	2,70	2,54	2,43	2,28	2,10	1,90	1,6	2,05
	7,60	5,42	4,54	4,04	3,73	3,50	3.20	2,87	2,49	2,03	2,76
	13,39	8,85	7,12	6,19	5,59	5,18	4.65	4,05	3,41	2,64	3,66
30	4,17	3,32	2,92	2,69	2,53	2,42	2,27	2,09	1,89	1,62	2,04
	7,56	5,39	4,51	4,02	3,70	3,47	3.17	2,84	2,47	2,01	2,75
	13,29	8,77	7,05	6,12	5,53	5,12	4,58	4,00	3,36	2,59	3,64
60	4,00	3,15	2,76	2,52	2,37	2,25	2,10	1,92	1,70	1,39	2,00
	7,08	4,98	4,13	3,65	3,34	3,12	2,82	2,50	2,12	1,60	2,66
	11,97	7,76	6,17	5,31	4,76	4,37	3,87	3,31	2,76	1,90	3,36
∞	3,84	2,99	2,60	2,37	2,21	2,09	1,94	1,75	1,52	1,03	1,96
	6,64	4,60	3,78	3,32	3,02	2,80	2,51	2,18	1,79	1,04	2,58
	10,83	6,91	5,42	4,62	4,10	3,74	3,27	2,74	2,13	1,05	3,29

Приложение2

Распределение Стьюдента (t-распределение)

v	Вероятность $\alpha=St(t)=P(\lvert T\rvert > t_{табл})$												
	0,9	0,8	0,7	0,6	0,5	0,4	0,3	0,2	0,1	0,05	0,02	0,01	0,001
1	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,657	636,61
2	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925	31,598
3	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,841	12,941
4	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604	8,610
5	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,043	6,859
6	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707	5,959
7	0,130	0,263	0,402	0,549	0,711	0,896	1,119	1,415	1,895	2,365	2,998	3,499	5,405
8	0,130	0,262	0,399	0,546	0,706	0,889	1,108	1,397	1,860	2,306	2,896	3,355	5,041
9	0,129	0,261	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250	4,781
10	0,129	0,260	0,327	0,542	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169	4,583
11	0,129	0,260	0,396	0,540	0,697	0,876	1,088	1,363	1,796	2,201	2,718	3,106	4,437
12	0,128	0,259	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055	4,318
13	0,128	0,259	0,394	0,538	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012	4,221
14	0,128	0,258	0,393	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977	4,140
15	0,128	0,258	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947	4,073
16	0,128	0,258	0,392	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,583	2,921	4,015
17	0,128	0,257	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,567	2,898	3,965
18	0,127	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	1,101	2,552	2,878	3,922
19	0,127	0,257	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,539	2,861	3,833

Окончание

v	Вероятность $\alpha=St(t)=P(|T| > t_{табл})$												
	0,9	0,8	0,7	0,6	0,5	0,4	0,3	0,2	0,1	0,05	0,02	0,01	0,001
20	0,127	0,257	0,391	0,533	0,687	0,860	1,064	1,325	1,725	2,086	2,528	2,845	3,850
21	0,127	0,257	0,391	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,518	2,831	3,81
22	0,127	0,256	0,390	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,508	2,819	3,792
23	0,127	0,256	0,390	0,532	0,685	0,868	1,060	1,319	1,714	2,069	2,500	2,807	3,767
24	0,127	0,256	0,390	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,402	2,797	3,745
25	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,316	1,708	2,060	2,485	2,787	3,725
26	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,315	1,706	2,056	2,479	2,779	3,707
27	0,127	0,256	0,389	0,531	0,684	0,855	1,057	1,314	1,703	2,052	2,473	2,771	3,690
28	0,127	0,256	0,389	0,530	0,683	0,855	1,056	1,313	1,701	2,048	2,467	2,763	3,674
29	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,311	1,699	2,045	2,462	2,756	3,659
30	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,310	1,697	2,042	2,457	2,750	3,646
40	0,126	0,255	0,388	0,529	0,681	0,851	1,050	1,303	1,684	2,021	2,423	2,704	3,551
60	0,126	0,254	0,387	0,527	0,679	0,848	1,046	1,296	1,671	2,000	2,390	2,660	3,460
120	0,126	0,254	0,386	0,526	0,677	0,845	1,041	1,289	1,658	1,980	2,358	2,617	3,373
∞	0,126	0,253	0,385	0,524	0,674	0,842	1,036	1,282	1,645	1,960	2,326	2,576	3,291

СПИСОК ЛИТЕРАТУРЫ

1. Айвазян С.А., Енюков И.С., Мешалкин Л.Д. Прикладная статистика. Классификация и снижение размерности. - М.: Финансы и статистика, 1989.

2. Айвазян С.А., Мхитарян В.С. Прикладная статистика и эконометрика. - М.: Юнити, 1998.

3. Вентцель Е.С. Теория вероятности. М., Наука, 1963.

4. Доугерти К. Введение в эконометрику. – М.: Финансы и статистика, 1999.

5. Елисеева И.И. Юзбашев М.М. Общая теория статистики. – 4-е издание, переработанное и дополненное. – М.: Финансы и статистика, 2001.

6. Магнус Я.Р., Катышев П.К., Пересетский А.А. Эконометрика. Начальный курс. - М.: Дело, 2000.

7. Практикум по эконометрике: Учебное пособие / И.И. Елисеева, С.В. Курышева, Н.М. Гордиенко и др.; Под ред. ИИ. Елисеевой. – М.: Финансы и статистика, 2001.

8. Пугачев В.С. Теория вероятностей и математическая статистика. – М.: Наука, 1979.

9. Трошин Л.И., Мхитарян В.С. Коррялиационный и регрессионный анализ. – М.: МЭСИ, 1991.

10. Эконометрика: Учебник / Под ред. И.И. Елисеевой. – М.: Финансы и статистика, 2001.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ 3
ТЕМАТИЧЕСКИЙ ПЛАН КУРСА И МЕТОДИЧЕСКИЕ УКАЗАНИЯ К ВЫПОЛНЕНИЮ КОНТРОЛЬНОЙ РАБОТЫ 4
ГЛАВА 1. ПАРНАЯ РЕГРЕССИЯ И КОРРЕЛЯЦИЯ 4
КОНТРОЛЬНОЕ ЗАДАНИЕ 1 14
ГЛАВА 2. МНОЖЕСТВЕННАЯ РЕГРЕССИЯ 15
КОТРОЛЬНОЕ ЗАДАНИЕ 2 24
ГЛАВА 3. СИСТЕМЫ ЭКОНОМЕТРИЧЕСКИХ УРАВНЕНИЙ 25
КОТРОЛЬНОЕ ЗАДАНИЕ 3 29
ГЛАВА 4. ВРЕМЕННЫЕ РЯДЫ 31
КОНТРОЛЬНОЕ ЗАДАНИЕ 4 37
ПРИЛОЖЕНИЕ 1
ПРИЛОЖЕНИЕ 2
СПИСОК ЛИТЕРАТУРЫ