АРМАВИРСКИЙ ГОСУДАРСТВЕННЫЙ
ПЕДАГОГИЧЕСКИЙ УНИВЕРСИТЕТ

Математический факультет

Кафедра Алгебры, геометрии и МПМ

ДЕРКАЧ Д.В.

ЭКОНОМЕТРИКА: задания и методические рекомендации по выполнению самостоятельных работ для студентов ОЗО

Учебно-методическое пособие

Армавир, 2010

УДК 378.147
ББК 74.58
Д 36

Печатается по решению
Ученого совета
математического факультета АГПУ
протокол №5 от 26.01.2010 г.

Рецензенты: кандидат педагогических наук, доцент кафедры информатики Зайцева О.Б.; старший преподаватель кафедры алгебры, геометрии и МПМ Сморкачева Г.М.

Деркач Д.В. ЭКОНОМЕТРИКА: задания и методические рекомендации по выполнению самостоятельных работ для студентов ОЗО (учебно-методическое пособие). – Армавир, 2010. – 41 с.

Учебно-методическое пособие предназначено для студентов ОЗО специальностей «Прикладная информатика в экономике», «Профессиональное обучение: Экономика и управление», «Экономика и управление на предприятии», «Математические методы в экономике», изучающих дисциплину «Эконометрика». По каждой теме представлена краткая справка, подробное решение типового задания и задания на самостоятельную работу по вариантам (30 вариантов).

Оглавление

1. Линейная парная регрессия

Краткая теоретическая справка

Регрессия [regression] – зависимость среднего значения какой-либо случайной величины от некоторой другой величины (парная регрессия) или нескольких величин (множественная регрессия).

Уравнение линейной парной регрессии имеет вид: $\hat{y} = a + bx$.

Для оценки параметров *a*, *b* *методом наименьших квадратов* (МНК) необходимо решить систему нормальных уравнений:

$$\begin{cases} na + b\sum x = \sum y, \\ a\sum x + b\sum x^2 = \sum yx. \end{cases} \tag{1.1}$$

Можно воспользоваться готовыми формулами решения системы:

$$b = \frac{\operatorname{cov}(x;y)}{\sigma_x^2} = \frac{\overline{y \cdot x} - \bar{y} \cdot \bar{x}}{\overline{x^2} - \bar{x}^2}, \ a = \bar{y} - b \cdot \bar{x}, \tag{1.2}$$

где $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ – среднее значение фактора X;

$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ – среднее значение результативной переменной Y;

$\overline{y \cdot x} = \frac{1}{n}\sum_{i=1}^{n} y_i \cdot x_i$ – среднее значение произведения переменных X и Y;

$\overline{x^2} = \frac{1}{n}\sum_{i=1}^{n} x_i^2$ – среднее значение квадрата переменной X;

$\operatorname{cov}(x;y) = \overline{x \cdot y} - \bar{x} \cdot \bar{y}$ – ковариация переменных X и Y;

$\sigma_x^2 = D_x = \overline{x^2} - \bar{x}^2$ – дисперсия переменной X.

Коэффициент регрессии *b* показывает, на сколько единиц в среднем по совокупности изменится результирующая переменная Y, если факторная переменная X увеличится на одну единицу.

Для оценки тесноты линейной связи между переменными используют *линейный коэффициент парной корреляции*:

$$r_{xy} = b\frac{\sigma_x}{\sigma_y} = \frac{\overline{x \cdot y} - \bar{x} \cdot \bar{y}}{\sigma_x \cdot \sigma_y}, \tag{1.3}$$

где $\sigma_x = \sqrt{D_x} = \sqrt{\overline{x^2} - \bar{x}^2}$ – срднеквадратическое отклонение (СКО) переменной X;

$\sigma_y = \sqrt{D_y} = \sqrt{\overline{y^2} - \bar{y}^2}$ – срднеквадратическое отклонение (СКО) переменной Y.

Можно считать, что:

1) если $r_{xy} > 0$, то имеется *прямая* линейная связь между переменными X и Y;

2) если $r_{xy} < 0$, то имеется *обратная* линейная связь между переменными X и Y;

3) если $r_{xy} \approx 0$ ($|r_{xy}| < 0{,}1$), то линейная связь между переменными X и Y *отсутствует*.

Качественная оценка тесноты связи величин X и Y может быть выявлена на основе шкалы Чеддока:

Тестона связи	**Значение коэффициента корреляции**
Слабая	0,1-0,3
Умеренная	0,3-0,5
Заметная	0,5-0,7
Высокая	0,7-0,9
Весьма высокая	0,9-0,99

Для оценки качества уравнения регрессии использую *коэффициент детерминации* R^2.

Коэффициент детерминации характеризует долю дисперсии, объясняемую регрессией, в общей дисперсии результативного признака (квадрат коэффициента корреляции):

$$R^2 = r_{xy}^2. \tag{1.4}$$

Коэффициент детерминации показывает, какую часть вариации (изменения) результативной переменной Y объясняет вариация (изменение) фактора X. Чем ближе R^2 к единице, тем лучше регрессионная модель.

Оценка статистической значимости уравнения регрессии в целом осуществляется с помощью *F-критерия Фишера*. Проверяется гипотеза H_0 о статистической незначимости уравнения регрессии. Для этого рассчитывается фактическое значение критерия по формуле:

$$F_{факт} = \frac{\sum(\hat{y} - \bar{y})/m}{\sum(y - \hat{y})^2/(n-m-1)}, \tag{1.5}$$

где n – число единиц совокупности;

m – число параметров при переменных x.

Если применяется линейное уравнение регрессии, то расчет $F_{\text{факт}}$ упрощается:

$$F_{факт} = \frac{R^2}{1-R^2} \cdot (n-2). \tag{1.6}$$

$F_{табл}$ – это максимально возможное значение критерия, которое могло сформироваться под влиянием случайных факторов при данных степенях свободы и уровне значимости α. Уровень значимости α – вероятность отвергнуть правильную гипотезу при условии, что она верна. Имеются таблицы критических (табличных) значений F-критерия: $F(\alpha;\ k_1;\ k_2)$, где $k_1 = m$, $k_2 = n - m - 1$. Для линейного уравнения парной регрессии с уровнем значимости α = 0,05 необходимо в таблице значений (приложение №4) найти значение $F(0{,}05;\ 1;\ n-2)$.

Если $F_{табл} < F_{факт}$, то гипотеза H_0 о случайной природе оцениваемых характеристик отклоняется и признается их статистическая значимость и надежность.

Для оценки статистической значимости коэффициентов регрессии и коэффициента корреляции рассчитывается t-критерий Стьюдента. Выдвигается гипотеза H_0 о случайной природе показателей, т.е. о незначимом их отличии от нуля. Наблюдаемые значения t-критерия рассчитываются по формулам:

$$t_b = \frac{b}{m_b},\ t_a = \frac{a}{m_a},\ t_r = \frac{|r_{xy}|}{m_r}, \tag{1.7}$$

где m_b, m_a, m_r – случайные ошибки параметров линейной регрессии и коэффициента корреляции.

Для линейной парной регрессии выполняется равенство $t_b = t_r = \sqrt{F}$, поэтому проверки гипотез о значимости коэффициента регрессии при факторе и коэффициента корреляции равносильны проверке гипотезы о статистической значимости уравнения регрессии в целом.

Вообще, случайные ошибки рассчитываются по формулам:

$$m_r = \sqrt{\frac{1 - r_{xy}^2}{n-2}},\ m_b = \frac{S_{ост}}{\sigma_x \sqrt{n}},\ m_a = S_{ост}\frac{\sqrt{\sum x^2}}{n\sigma_x}. \tag{1.8}$$

где $S_{ост}^2$ – остаточная дисперсия на одну степень свободы:

$$S_{ост} = \sqrt{\frac{\sum (y - \hat{y})^2}{n-2}}. \tag{1.9}$$

Табличное (критическое) значение t-статистики находят по таблицам распределения t-Стьюдента при уровне значимости α = 0,05 и числе степеней свободы $k = n - 2$ (приложение №5). Если $t_{табл} < t_{факт}$, то H_0 отклоняется, т.е. коэффициенты регрессии не случайно отличаются от нуля и сформировались под влиянием систематически действующего фактора.

Решение типового примера

Имеются статистические данные (приложение 1, ст. 0, ст. 31) о численности населения (оценка на конец 2008 года; сотни тысяч человек) (фактор X) и об обороте розничной торговли в этих же регионах (в фактически действовавших ценах, млрд. руб.) (результативный признак Y) в некоторых регионах Российской федерации.

№	Регион	X	Y	№	Регион	X	Y
1	Республика Адыгея	4,4	25,6	9	Ставропольский край	27,1	203,6
2	Республика Дагестан	27,1	217,3	10	Астраханская область	10,1	77,7
3	Кабардино-Балкарская Республика	8,9	48	11	Волгоградская область	26	186,1
4	Республика Калмыкия	2,8	7,4	12	Ростовская область	42,4	423,4
5	Карачаево-Черкесская Республика	4,3	22,6	13	Ивановская область	10,7	57,9
6	Республика Северная Осетия - Алания	7	37,4	14	Калужская область	10	80,7
7	Чеченская Республика	12,4	25,2	15	Костромская область	6,9	38,1
8	Краснодарский край	51,4	500,7				

Требуется:

1. Построить график зависимости между переменными, по которому необходимо подобрать модель уравнения регрессии.

2. Рассчитать параметры уравнения регрессии $y = a + bx$ методом наименьших квадратов.

3. Оценить тесноту связи между переменными с помощью показателей корреляции и детерминации.

4. Охарактеризовать статистическую надежность результатов регрессионного анализа с использованием F-критерия Фишера при уровне значимости $\alpha = 0,05$.

5. Оценить значимость коэффициентов регрессии и корреляции по t-критерию Стьюдента при уровне значимости $\alpha = 0,05$.

6. Определить прогнозное значение результативного признака, если возможное значение факторного признака составит 1,3 от его среднего уровня по совокупности.

Решение:

1. Построим график зависимости оборота розничной торговли от численности населения в среде MS Excel с помощью точечной диаграммы (рис. 1.1). По оси абсцисс откладываем значения фактора X (численность населения), по оси ординат – значения результативного признака Y (оборот розничной торговли). Визуальный анализ диаграммы показывает, что можно использовать линейную регрессионную модель.

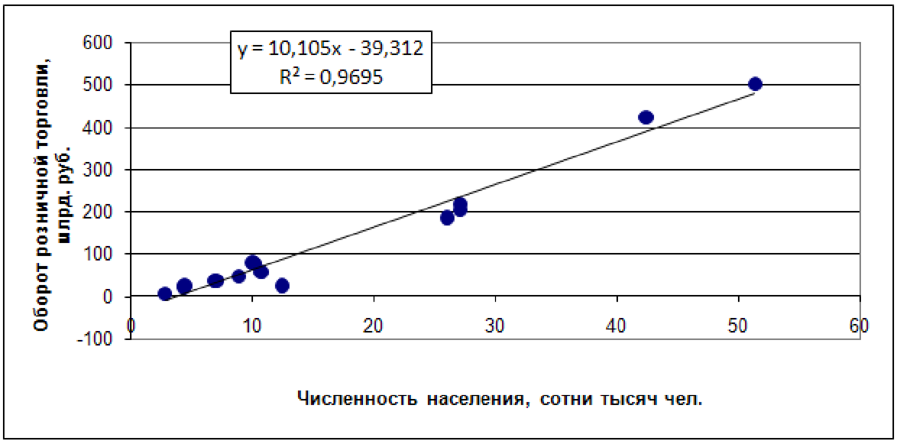


Рис. 1.1. Зависимость оборота розничной торговли от численности населения по регионам РФ.

2. Составим расчетную таблицу, заполним столбцы 1-6.

Таблица 1.1

№	x	y	x^2	y^2	xy	$\hat{y}$	$y-\hat{y}$	$(y-\hat{y})^2$
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1.	4,4	25,6	19,36	655,36	112,64	5,15	20,45	418,22
2.	27,1	217,3	734,41	47219,29	5888,83	234,53	-17,23	296,89
3.	8,9	48,0	79,21	2304,00	427,20	50,62	-2,62	6,87
4.	2,8	7,4	7,84	54,76	20,72	-11,02	18,42	339,23
5.	4,3	22,6	18,49	510,76	97,18	4,14	18,46	340,81
6.	7,0	37,4	49,00	1398,76	261,80	31,42	5,98	35,73
7.	12,4	25,2	153,76	635,04	312,48	85,99	-60,79	3695,26
8.	51,4	500,7	2641,96	250700,49	25735,98	480,08	20,62	425,21
9.	27,1	203,6	734,41	41452,96	5517,56	234,53	-30,93	956,70
10.	10,1	77,7	102,01	6037,29	784,77	62,75	14,95	223,58
11.	26,0	186,1	676,00	34633,21	4838,60	223,42	-37,32	1392,42
12.	42,4	423,4	1797,76	179267,56	17952,16	389,14	34,26	1174,07
13.	10,7	57,9	114,49	3352,41	619,53	68,81	-10,91	119,04
14.	10,0	80,7	100,00	6512,49	807,00	61,74	18,96	359,60
15.	6,9	38,1	47,61	1451,61	262,89	30,41	7,69	59,11
Сумма	251,5	1951,7	7276,31	576185,99	63639,34			9842,74
Среднее	**16,77**	**130,11**	**485,09**	**38412,40**	**4242,62**			
Дисперсия	203,97	21482,92						
СКО	**14,28**	**146,57**						

Найдем дисперсию и среднее квадратическое отклонение переменных:

Дисперсия: $D_x = \overline{x^2} - \bar{x}^2 = 485{,}09 - 16{,}77^2 = 203{,}97$;

$D_y = \overline{y^2} - \bar{y}^2 = 38412{,}4 - 130{,}11^2 = 21482{,}92$.

Среднее квадратическое отклонение (СКО):

$\sigma_x = \sqrt{D_x} = \sqrt{203{,}97} = 14{,}28$;

$\sigma_y = \sqrt{D_y} = \sqrt{21482{,}92} = 146{,}57$.

Найдем параметры уравнения регрессии:

$$b = \frac{\operatorname{cov}(x;y)}{\sigma_x^2} = \frac{\overline{x \cdot y} - \bar{x} \cdot \bar{y}}{D_x} = \frac{4242{,}62 - 16{,}77 \cdot 130{,}11}{203{,}97} = 10{,}1;$$

$a = \bar{y} - b \cdot \bar{x} = 130{,}11 - 10{,}1 \cdot 16{,}77 = -39{,}31$.

Таким образом, уравнение регрессии имеет вид: $\hat{y} = 10{,}1x - 39{,}31$.

Вывод: с увеличением численности населения на 1 единицу (т.е. на 100 000 человек) оборот розничной торговли увеличивается в среднем на 10 100 млн. рублей.

3. Оценим тесноту связи между переменными с помощью показателей корреляции и детерминации.

Линейный коэффициент корреляции:

$$r_{xy} = b\frac{\sigma_x}{\sigma_y} = 10{,}1 \cdot \frac{14{,}28}{146{,}57} = 0{,}98.$$

Вывод: линейная связь между оборотом розничной торговли и численностью населения по регионам весьма высокая, т.к. $r_{xy} = 0{,}98 \geq 0{,}9$.

Коэффициент детерминации:

$$R^2 = r_{xy}^2 = 0{,}98^2 = 0{,}97.$$

Вывод: Коэффициент детерминации показывает, что 97% различий в обороте розничной торговли по регионам (y) объясняется вариацией численности населения в этих регионах (x), а 3% другими, неучтенными факторами.

4. Охарактеризуем статистическую надежность результатов регрессионного анализа с использованием F-критерия Фишера при уровне значимости α = 0,05.

Рассчитаем фактическое значение критерия:

$$F_{факт} = \frac{R^2}{1-R^2} \cdot (n-2) = \frac{0{,}97}{1-0{,}97} \cdot (15-2) = 412{,}61.$$

Найдем табличное значение критерия при уровне значимости α = 0,05 и степенях свободы: $k_1 = m = 1$, $k_2 = n - m - 1 = 15 - 1 - 1 = 13$:

$F_{табл}(\alpha = 0{,}05;\ k_1 = 1;\ k_2 = 13) = 4{,}67$ (см. приложение №4).

Т.к. $F_{табл} = 4{,}67 < 412{,}61 = F_{факт}$, то уравнение регрессии является статистически значимым, надежным.

5. Оценим значимость коэффициентов регрессии и корреляции по t-критерию Стьюдента при уровне значимости α = 0,05.

Поскольку сделан вывод о статистической значимости уравнения регрессии в целом по критерию Фишера, то статистическими значимо отличны от нуля и коэффициент корреляции r_{xy} и коэффициент регрессии b при факторе.

Продолжим заполнять расчетную таблицу 1.1, заполним столбцы 7-9.

Для заполнения столбца 7, необходимо рассчитать теоретические значение переменной y соответствующие имеющимся значениям фактора x с помощью полученного уравнения регрессии $y = 10{,}1x - 39{,}31$.

Так, для $x = 4{,}4$: $\hat{y} = 10{,}1 \cdot 4{,}4 - 39{,}31 = 5{,}15$.

Для $x = 27,1$: $\hat{y} = 10,1 \cdot 27,1 - 39,31 = 234,53$ и т.д.

Столбец 8 заполняется как разность соответствующих значений столбца 2 и столбца 7, т.е. рассчитывается по формуле $y - \hat{y}$.

Так, для $i = 1$: $y - \hat{y} = 25,6 - 5,15 = 20,45$;

для $i = 2$: $y - \hat{y} = 217,3 - 234,53 = -17,23$ и т.д.

Далее заполним 9 столбец расчетной таблицы по формуле $(y - \hat{y})^2$ и сумму этого столбца.

Теперь можно рассчитать фактическое (наблюдаемое) значение t-критерия для свободного члена a в уравнении регрессии. Найдем остаточную дисперсию на одну степень свободы (в числителе дроби под корнем сумма значений 9 столбца расчетной таблицы 1.1):

$$S_{ост} = \sqrt{\frac{\sum (y - \hat{y})^2}{n-2}} = \sqrt{\frac{9842,74}{15-2}} = 27,52.$$

Случайная ошибка свободного члена уравнения регрессии:

$$m_a = S_{ост} \frac{\sqrt{\sum x^2}}{n\sigma_x} = 27,52 \cdot \frac{\sqrt{7276,31}}{15 \cdot 14,28} = 10,96.$$

Значение t-статистики для параметра a:

$$t_a = \frac{|a|}{m_a} = \frac{39,31}{10,96} = 3,59.$$

Табличное значение t-критерия для уровня значимости $\alpha = 0,05$ и числа степеней свободы $k = n - 2 = 13$ равно 2,1604 (см. приложение 5). Поскольку $t_{табл} = 2,1604 < 3,59 = t_{факт}$, то значение свободного члена уравнения регрессии статически значимо отличается от нуля.

7. Определим прогнозное значение результативного признака, если возможное значение факторного признака составит 1,3 от его среднего уровня по совокупности.

Прогнозное значение результативного признака определяется путем подстановки в уравнение регрессии возможного значения факторного признака $x_p = \bar{x} \cdot 1,3 = 16,77 \cdot 1,3 = 21,8$.

Прогнозное значение оборота розничной торговли составит:

$y = 10,1 \cdot 21,8 - 39,31 = 180,94$.

Значит, при численности населения в регионе в 2,18 млн. человек возможный оборот розничной торговли в регионе составит 180,94 млрд. руб.[1]

[1] Для прогнозного значения результативного признака корректнее рассчитать доверительные интервалы. См. Елисеева И.И. Практикум по эконометрике. – М.: Финансы и статистика, 2005. – С. 9.

Задание для самостоятельной работы

Имеются статистические данные (приложение №1) по некоторым регионам Российской Федерации. Для факторной переменной X (численность населения, столбец 0) и результативной переменной Y (столбец N, где N – номер варианта) требуется:

1. Рассчитать параметры уравнения регрессии $\hat{y} = a + bx$ методом наименьших квадратов.
2. Оценить тесноту связи между переменными с помощью показателей корреляции и детерминации.
3. Охарактеризовать статистическую надежность результатов регрессионного анализа с использованием F-критерия Фишера при уровне значимости $\alpha = 0{,}05$.
4. Оценить значимость коэффициентов регрессии и корреляции по t-критерию Стьюдента при уровне значимости $\alpha = 0{,}05$.
5. Определить прогнозное значение результативного признака, если возможное значение факторного признака составит 1,3 от его среднего уровня по совокупности.

2. Линейная множественная регрессия

Краткая теоретическая справка по теме

Для оценки параметров уравнения линейной множественной регрессии

$$y = a + b_1 x_1 + b_2 x_2 + ... + b_p x_p \qquad (2.1)$$

применяют метод наименьших квадратов – строится система нормальных уравнений, решение которой позволяет получить оценки параметров регрессии:

$$\begin{cases} \sum y = na + b_1 \sum x_1 + b_2 \sum x_2 + ... + b_p \sum x_p, \\ \sum y x_1 = a \sum x_1 + b_1 \sum x_1^2 + b_2 \sum x_2 x_1 + ... + b_p \sum x_p x_1, \\ ... \\ \sum y x_p = a \sum x_p + b_1 \sum x_1 x_p + b_2 \sum x_2 x_p + ... + b_p \sum x_p^2. \end{cases} \qquad (2.2)$$

Другой вид уравнения множественной регрессии – уравнение регрессии в стандартизированном масштабе:

$$t_y = \beta_1 t_{x_1} + \beta_2 t_{x_2} + ... + \beta_p t_{x_p}, \qquad (2.3)$$

где $t_y = \dfrac{y - \bar{y}}{\sigma_y}$, $t_{x_i} = \dfrac{x_i - \overline{x_i}}{\sigma_{x_i}}$ – стандартизированные переменные;

β_i – стандартизированные коэффициенты регрессии.

К уравнению множественной регрессии в стандартизированном масштабе применим МНК, что приводит к решению системы уравнений:

$$\begin{cases} r_{yx_1} = \beta_1 + \beta_2 r_{x_2 x_1} + \beta_3 r_{x_3 x_1} + ... + \beta_p r_{x_p x_1}, \\ r_{yx_2} = \beta_1 r_{x_1 x_2} + \beta_2 + \beta_3 r_{x_3 x_2} + ... + \beta_p r_{x_p x_2}, \\ ... \\ r_{yx_p} = \beta_1 r_{x_1 x_p} + \beta_2 r_{x_2 x_p} + \beta_3 r_{x_3 x_p} + ... + \beta_p. \end{cases} \qquad (2.4)$$

Для двухфакторной модели линейной регрессии $t_y = \beta_1 t_{x_1} + \beta_2 t_{x_2}$ расчет β-коэффициентов можно выполнить по формулам (следуют из решения системы (2.4)):

$$\beta_1 = \frac{r_{yx_1} - r_{yx_2} r_{x_1 x_2}}{1 - r_{x_1 x_2}^2}, \quad \beta_2 = \frac{r_{yx_2} - r_{yx_1} r_{x_1 x_2}}{1 - r_{x_1 x_2}^2} \qquad (2.5)$$

Связь коэффициентов множественной регрессии b_i со стандартизированными коэффициентами β_i описывается соотношением:

$$b_i = \beta_i \frac{\sigma_y}{\sigma_{x_i}}, \quad \beta_i = b_i \frac{\sigma_{x_i}}{\sigma_y}. \qquad (2.6)$$

При этом: $a = \bar{y} - b_1 \overline{x_1} - b_2 \overline{x_2}$.

Тесноту совместного влияния факторов на результат оценивает *коэффициент множественной корреляции*, который можно определить по формуле:

$$R_{yx_1x_2\ldots x_p}=\sqrt{\sum\beta_i r_{yx_i}}\,, \qquad (2.7)$$

где β_i – стандартизированные коэффициенты регрессии,

r_{yx_i} – парные коэффициенты корреляции между переменными y и x_i.

Качество построенной модели в целом оценивает коэффициент (индекс) детерминации. *Коэффициент множественной детерминации* рассчитывается как квадрат индекса множественной корреляции:

$$R^2_{yx_1x_2\ldots x_p}\,. \qquad (2.8)$$

Частные коэффициенты корреляции характеризуют тесноту связи между результатом и соответствующим фактором при устранении влияния (при закреплении их влияния на постоянном уровне) других факторов, включенных в уравнение регрессии. Для двухфакторной модели их можно определить по формулам:

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

$$r_{x_1x_2\cdot y}=\frac{r_{x_1x_2}-r_{yx_1}\cdot r_{yx_2}}{\sqrt{(1-r^2_{yx_1})(1-r^2_{yx_2})}}\,. \qquad (2.9)$$

При построении уравнения множественной регрессии может возникнуть проблема мультиколлениарности факторов (тесная линейная зависимость более двух факторов). Считается, что две переменные явно *коллинеарны*, если $r_{x_ix_j}>0{,}7$.

Статистическая значимость уравнения множественной регрессии в целом оценивается с помощью общего F-критерия Фишера:

$$F=\frac{R^2_{yx_1x_2\ldots x_p}}{1-R^2_{yx_1x_2\ldots x_p}}\cdot\frac{n-m-1}{m}\,, \qquad (2.10)$$

где m – число факторов в линейном уравнении регрессии;
n – число наблюдений.

Вывод о статистической значимости уравнения множественной регрессии в целом и коэффициента множественной детерминации можно сделать, если наблюдаемое значение критерия больше табличного, найденного для заданного уровня значимости (например, $\alpha=0{,}05$) и степенях свободы $k_1=m$, $k_2=n-m-1$.

Частный F-критерий оценивает статистическую значимость присутствия каждого из факторов в уравнении множественной регрессии. Для двухфакторной модели F_{x_1} оценивает целесообразность включения в уравнение фактора x_1 после того, как в него был включен фактор x_2; F_{x_2}

оценивает целесообразность включения в уравнение фактора x_2 после того, как в него был включен фактор x_1:

$$F_{x_1} = \frac{R^2_{yx_1x_2} - r^2_{yx_2}}{1 - R^2_{yx_1x_2}} \cdot \frac{n-m-1}{1}, \quad F_{x_2} = \frac{R^2_{yx_1x_2} - r^2_{yx_1}}{1 - R^2_{yx_1x_2}} \cdot \frac{n-m-1}{1}, \qquad (2.11)$$

где m – число факторов в линейном уравнении регрессии;
n – число наблюдений.

Фактическое значение частного F-критерия сравнивается с табличным при 5%-ном или 1%-ном уровне значимости и числе степеней свободы: $k_1 = 1$, $k_2 = n - m - 1$. Если фактическое значение превышает табличное, то дополнительное включение соответствующего фактора в модель статистически оправдано, в противном случае фактор в модель включать нецелесообразно.

Решение типового примера

Задание

Имеются данные[2] о стоимости автомобилей ВАЗ 2110 (результативная переменная *y*, тыс. руб.) в Краснодарском крае, о годе выпуска (возраст автомобиля – фактор x_1, лет) и о пробеге (фактор x_2, тыс. км):

№	Возраст, лет	Пробег, тыс. км	Цена, тыс. руб.
1	5	50	167
2	5	70	175
3	8	110	146
4	8	120	143
5	10	175	120
6	4	62	220
7	5	87,5	150
8	5	84	172
9	7	77	170
10	4	83	190
11	4	65	210
12	8	120	143
13	6	88	167
14	7	89	150
15	4	83	195

Требуется:

1) Найти уравнение линейной множественной регрессии в стандартизированной ($t_y = \beta_1 t_{x_1} + \beta_2 t_{x_2}$) и естественной форме ($y = a + b_1 x_1 + b_2 x_2$).

[2] Источник: поисковая система Яндекс, информация о продаже подержанных автомобилей, 2010 г.

2) Найти коэффициенты множественной и частной корреляции, множественной детерминации; дать их характеристику.

3) Рассчитать общий и частные F-критерии Фишера; оценить статистическую надежность уравнения регрессии и коэффициента множественной детерминации; оценить целесообразность включения в уравнение множественной регрессии фактора x_1 после фактора x_2 и целесообразность включения фактора x_2 после фактора x_1.

4) При необходимости найти уравнение парной регрессии (исключив статистически незначимый фактор).

Решение

1. Рассчитаем параметры уравнения линейной множественной регрессии в стандартизированной форме $t_y = \beta_1 t_{x_1} + \beta_2 t_{x_2}$ и естественной форме $y = a + b_1 x_1 + b_2 x_2$ методом наименьших квадратов.

Составим расчетную таблицу.

Таблица 2.1

№	y	x_1	x_2	yx_1	yx_2	x_1x_2	y^2	x_1^2	x_2^2
1	167	5	50	835	8350	250	27889	25	2500
2	175	5	70	875	12250	350	30625	25	4900
3	146	8	110	1168	16060	880	21316	64	12100
4	143	8	120	1144	17160	960	20449	64	14400
5	120	10	175	1200	21000	1750	14400	100	30625
6	220	4	62	880	13640	248	48400	16	3844
7	150	5	87,5	750	13125	437,5	22500	25	7656,25
8	172	5	84	860	14448	420	29584	25	7056
9	170	7	77	1190	13090	539	28900	49	5929
10	190	4	83	760	15770	332	36100	16	6889
11	210	4	65	840	13650	260	44100	16	4225
12	143	8	120	1144	17160	960	20449	64	14400
13	167	6	88	1002	14696	528	27889	36	7744
14	150	7	89	1050	13350	623	22500	49	7921
15	195	4	83	780	16185	332	38025	16	6889
Сумма	2518	90	1363,5	14478	219934	8869,5	433126	590	137078,3
Среднее	**167,87**	**6,00**	**90,90**	**965,20**	**14662,27**	**591,30**	**28875,07**	**39,33**	**9138,55**

Найдем средние квадратические отклонения переменных:

$$\sigma_y = \sqrt{\overline{y^2} - \bar{y}^2} = \sqrt{28875{,}07 - 167{,}87^2} = 26{,}38;$$

$$\sigma_{x_1} = \sqrt{\overline{x_1^2} - \overline{x_1}^2} = \sqrt{39{,}33 - 6{,}0^2} = 1{,}83;$$

$$\sigma_{x_2} = \sqrt{\overline{x_2^2} - \overline{x_2}^2} = \sqrt{9138{,}55 - 90{,}9^2} = 29{,}59.$$

Найдем коэффициенты парной корреляции:

$$r_{yx_1} = \frac{\operatorname{cov}(y, x_1)}{\sigma_y \cdot \sigma_{x_1}} = \frac{\overline{y \cdot x_1} - \bar{y} \cdot \overline{x_1}}{\sigma_y \cdot \sigma_{x_1}} = \frac{965{,}2 - 167{,}87 \cdot 6{,}0}{26{,}38 \cdot 1{,}83} = -0{,}87;$$

$$r_{yx_2} = \frac{\operatorname{cov}(y, x_2)}{\sigma_y \cdot \sigma_{x_2}} = \frac{\overline{y \cdot x_2} - \bar{y} \cdot \overline{x_2}}{\sigma_y \cdot \sigma_{x_2}} = \frac{14662{,}27 - 167{,}87 \cdot 90{,}9}{26{,}38 \cdot 29{,}59} = -0{,}76;$$

$$r_{x_1x_2} = \frac{\operatorname{cov}(x_1, x_2)}{\sigma_{x_1} \cdot \sigma_{x_2}} = \frac{\overline{x_1 \cdot x_2} - \overline{x_1} \cdot \overline{x_2}}{\sigma_{x_1} \cdot \sigma_{x_2}} = \frac{591{,}3 - 6{,}0 \cdot 90{,}9}{1{,}83 \cdot 29{,}59} = 0{,}85.$$

Стандартизированные β-коэффициенты определим по формулам (2.5):

$$\beta_1 = \frac{r_{yx_1} - r_{yx_2} r_{x_1x_2}}{1 - r^2_{x_1x_2}} = \frac{-0{,}87 - (-0{,}76) \cdot 0{,}85}{1 - 0{,}85^2} = -0{,}8;$$

$$\beta_2 = \frac{r_{yx_2} - r_{yx_1} r_{x_1x_2}}{1 - r^2_{x_1x_2}} = \frac{-0{,}76 - (-0{,}87) \cdot 0{,}85}{1 - 0{,}85^2} = -0{,}09.$$

Таким образом, уравнение регрессии в стандартизированной форме имеет вид: $t_y = -0{,}8t_{x_1} - 0{,}09t_{x_2}$.

Вывод: Сравнение модулей значений стандартизированных коэффициентов регрессии ($|\beta_1| = 0{,}8 > |\beta_2| = 0{,}09$) говорит о том, что на цену автомобиля возраст (фактор x_1) оказывает значительно большее влияние, нежели пробег (фактор x_2).

Рассчитаем естественные коэффициенты регрессии:

$$b_1 = \beta_1 \frac{\sigma_y}{\sigma_{x_1}} = -0{,}8 \cdot \frac{26{,}38}{1{,}83} = -11{,}56;$$

$$b_2 = \beta_2 \frac{\sigma_y}{\sigma_{x_2}} = -0{,}09 \cdot \frac{26{,}38}{29{,}59} = -0{,}08;$$

$$a = \bar{y} - b_1\overline{x_1} - b_2\overline{x_2} = 167{,}87 - (-11{,}56) \cdot 6{,}0 - (-0{,}08) \cdot 90{,}9 = 244{,}09.$$

Получаем уравнение линейной множественной (двухфакторной) регрессии в естественной форме: $y = 244{,}09 - 11{,}56x_1 - 0{,}08x_2$.

Вывод: с увеличением возраста машины на 1 год ее цена уменьшается в среднем на 11,56 тыс. рублей, а с увеличением пробега на 1 тыс. км цена уменьшается в среднем на 0,08 тыс. рублей (80 рублей).

2. Найдем коэффициенты множественной и частной корреляции, а также множественной детерминации.

Коэффициент множественной корреляции находится по формуле:

$$R_{yx_1x_2} = \sqrt{\beta_1 r_{yx_1} + \beta_2 r_{yx_2}} = \sqrt{-0{,}8 \cdot (-0{,}87) - 0{,}09 \cdot (-0{,}76)} = \sqrt{0{,}76} = 0{,}87.$$

$R^2_{yx_1x_2} = \left(\sqrt{0{,}76}\right)^2 = 0{,}76$ – коэффициент множественной детерминации.

Вывод: величина коэффициента множественной корреляции показывает, что связь между y, x_1, x_2 – высокая[3], причем 76,3% вариации цены на автомобиль объясняется вариацией возраста машины и пробега.

Коэффициенты частной корреляции определяются через парные коэффициенты корреляции по формулам:

---

[3] При качественной интерпретации коэффициента корреляции используется шкала Чеддока

$$r_{yx_1\cdot x_2}=\frac{r_{yx_1}-r_{yx_2}\cdot r_{x_1x_2}}{\sqrt{(1-r_{yx_2}^2)(1-r_{x_1x_2}^2)}}=\frac{-0{,}87-(-0{,}76)\cdot 0{,}85}{\sqrt{(1-(-0{,}76)^2)(1-0{,}85^2)}}=-0{,}65;$$

$$r_{yx_2\cdot x_1}=\frac{r_{yx_2}-r_{yx_1}\cdot r_{x_1x_2}}{\sqrt{(1-r_{yx_1}^2)(1-r_{x_1x_2}^2)}}=\frac{-0{,}76-(-0{,}87)\cdot 0{,}85}{\sqrt{(1-(-0{,}87)^2)(1-0{,}85^2)}}=-0{,}09;$$

$$r_{x_1x_2\cdot y}=\frac{r_{x_1x_2}-r_{yx_1}\cdot r_{yx_2}}{\sqrt{(1-r_{yx_1}^2)(1-r_{yx_2}^2)}}=\frac{0{,}85-(-0{,}76)\cdot(-0{,}87)}{\sqrt{(1-(-0{,}76)^2)(1-(-0{,}87)^2)}}=0{,}32.$$

Вывод: коэффициенты частной корреляции характеризуют тесноту связи между двумя переменными, исключив влияние третьей переменной. Значит, связь между ценой на ВАЗ 2110 и годом выпуска при исключении влияния величины пробега обратная и заметная; между ценой автомобиля и пробегом без учета возраста машины – обратная, но слабая; связь между факторами x_1 и x_2 – умеренная.

Сравним соответствующие коэффициенты парной и частной корреляции: $r_{yx_1}=-0{,}87$, $r_{yx_2}=-0{,}76$, $r_{x_1x_2}=0{,}85$;

$r_{yx_1\cdot x_2}=-0{,}65$, $r_{yx_2\cdot x_1}=-0{,}09$, $r_{x_1x_2\cdot y}=0{,}32$.

Вывод:

1) при закреплении фактора x_2 на постоянном уровне влияние на y фактора x_1 оказалось несколько менее сильным (– 0,65 против – 0,87), но все равно остается заметным;

2) при закреплении фактора x_1 на постоянном уровне влияние на y фактора x_2 стало весьма слабым (– 0,09 против – 0,76);

3) межфакторная связь ($r_{x_1x_2}=0{,}85$) говорит о высокой коллинеарности факторов, причем исключив влияние результативной переменной y эта связь становится умеренной.

3. Оценим значимость уравнения регрессии и коэффициента множественной детерминации с помощью F-критерия Фишера. Наблюдаемое значение критерия находится по формуле:

$$F_{набл}=\frac{R_{yx_1x_2}^2}{1-R_{yx_1x_2}^2}\cdot\frac{n-m-1}{m}=\frac{0{,}76}{1-0{,}76}\cdot\frac{15-2-1}{2}=19{,}27.$$

Табличное значение критерия при уровне значимости $\alpha=0{,}05$ и $k_1=m=2$, $k_2=n-m-1=15-2-1=12$:

$F_{табл}=F(0{,}05;\ 2;\ 12)=3{,}88$.

Вывод: т.к. $F_{табл}<F_{набл}$, то с вероятностью $1-\alpha=0{,}95$ делаем заключение о статистической значимости уравнения регрессии и коэффициента множественной детерминации, которые сформировались под неслучайным воздействием факторов x_1 и x_2.

Оценим целесообразность включения в уравнение множественной регрессии фактора x_1 после фактора x_2 и целесообразность включения фактора x_2 после фактора x_1 с помощью частных F-критериев F_{x_1} и F_{x_2}.

$$F_{x_1\ набл} = \frac{R^2_{yx_1x_2} - r^2_{yx_2}}{1 - R^2_{yx_1x_2}} \cdot \frac{n-m-1}{1} = \frac{0{,}76 - (-0{,}76)^2}{1 - 0{,}76} \cdot \frac{15-2-1}{1} = 9{,}00;$$

$$F_{x_2\ набл} = \frac{R^2_{yx_1x_2} - r^2_{yx_1}}{1 - R^2_{yx_1x_2}} \cdot \frac{n-m-1}{1} = \frac{0{,}76 - (-0{,}87)^2}{1 - 0{,}76} \cdot \frac{15-2-1}{1} = 0{,}10.$$

Найдем табличные значения критерия на уровне значимости $\alpha = 0{,}05$ и $k_1 = 1$, $k_2 = n - m - 1 = 15 - 2 - 1 = 12$: $F_{табл} = F(0{,}05;\ 1;\ 12) = 4{,}75$.

Вывод: 1) Поскольку $F_{x_1\ набл} > F_{табл}$, то включение в модель фактора x_1 (возраста автомобиля) после фактора x_2 статистически оправдано и коэффициент b_1 при факторе x_1 статистически значим.

2) Поскольку $F_{x_2\ набл} < F_{табл}$, то нецелесообразно включать в модель фактор x_2 (пробег) после фактора x_1. Это означает, что парная регрессия зависимости цены ВАЗ 2110 от возраста машины является достаточно статистически значимой, надежной и что нет необходимости улучшать ее, включая дополнительный фактор x_2.

Найдем уравнение парной регрессии $y = a + bx_1$, где y – цена автомобиля (тыс. руб), x_1 – возраст машины (лет):

$$b = \frac{\mathrm{cov}(x_1; y)}{\sigma^2_{x_1}} = \frac{\overline{x_1 \cdot y} - \overline{x_1} \cdot \overline{y}}{\sigma^2_{x_1}} = \frac{965{,}2 - 6 \cdot 167{,}87}{1{,}83^2} = -12{,}6;$$

$$a = \overline{y} - b \cdot \overline{x_1} = 167{,}87 - (-12{,}6) \cdot 6 = 243{,}47.$$

Получаем: $y = 243{,}47 - 12{,}6x_1$.

Задание для самостоятельной работы

Имеются данные[4] о стоимости автомобилей (результативная переменная y, тыс. руб.) в Краснодарском крае, о годе выпуска (возраст автомобиля – фактор x_1, лет) и о пробеге (фактор x_2, тыс. км).

Требуется:

1) Найти уравнение линейной множественной регрессии в стандартизированной $(t_y = \beta_1 t_{x_1} + \beta_2 t_{x_2})$ и естественной форме $(y = a + b_1x_1 + b_2x_2)$.

2) Найти коэффициенты множественной и частной корреляции, множественной детерминации; дать их характеристику.

3) Рассчитать общий и частные F-критерии Фишера; оценить статистическую надежность уравнения регрессии и коэффициента

[4] Источник: поисковая система Яндекс, информация о продаже подержанных автомобилей, 2010 г.

множественной детерминации; оценить целесообразность включения в уравнение множественной регрессии фактора x_1 после фактора x_2 и целесообразность включения фактора x_2 после фактора x_1.

4) При необходимости найти уравнение парной регрессии (исключив статистически незначимый фактор).

3. Анализ временных рядов

Краткая теоретическая справка

Временным рядом (*рядом динамики*, *динамическим рядом*) называется упорядоченная во времени последовательность наблюдений:

$$y_1, y_2, ..., y_t, ..., y_n .$$

Всякий ряд динамики включает два обязательных элемента: *уровни ряда* – числовые значения того или иного статистического показателя и *время*, выраженное моментами или периодами, к которым относятся уровни.

Под *прогнозом* понимается научно обоснованное описание возможных состояний объектов в будущем, а также альтернативных путей и сроков достижения этого состояния.

Время (*период*) *упреждения прогноза* – отрезок времени от момента, для которого имеются последние статистические данные об изучаемом объекте, до момента, к которому относится прогноз.

По времени упреждения экономические прогнозы делятся на:

- оперативные (с периодом упреждения до одного месяца);
- краткосрочные (период упреждения – от одного, нескольких месяцев до года);
- среднесрочные (период упреждения более 1 года, но не превышает 5 лет);
- долгосрочные (с периодом упреждения более 5 лет).

Каждый уровень временного ряда формируется из *трендовой* (*T*), *циклической* (*S*) и *случайной* (*E*) компонент.

Аддитивная модель временного ряда: уровни ряда представлены в виде суммы указанных компонент $Y = T + S + E$.

Мультипликативная модель временного ряда: уровни ряда представлены в виде произведения указанных компонент $Y = T \cdot S \cdot E$.

Автокорреляция уровней ряда – это корреляционная зависимость между последовательными уровнями временного ряда:

$$r_i = \frac{\operatorname{cov}(y_{t-i};\, y_t)}{\sigma_{y_{t-i}}\sigma_{y_t}} = \frac{\overline{y_{t-i} \cdot y_t} - \overline{y_{t-i}} \cdot \overline{y_t}}{\sqrt{\overline{y_{t-i}^2} - \overline{y_{t-i}}^2} \cdot \sqrt{\overline{y_t^2} - \overline{y_t}^2}},$$

где $\overline{y_t} = \frac{1}{n-i}\sum_{t=i+1}^{n} y_t$, $\overline{y_{t-i}} = \frac{1}{n-i}\sum_{t=1}^{n-i} y_t$.

Последовательность коэффициентов автокорреляции называется *автокорреляционной функцией* временного ряда, а график ее значений от величины лага порядка коэффициента автокорреляции) – *коррелограммой*.

Автокорреляция в остатках – корреляционная зависимость между значениями остатков ε_t за текущий и предыдущие моменты времени.

Построение аддитивной модели выполняется в следующей последовательности:

1) выравнивание исходного ряда скользящей средней;

2) оценка сезонной компоненты с учетом того, что для аддитивной модели сумма сезонных компонент за год (весь период) равна нулю;

3) удаление сезонных компонент из исходных уровней ряда $Y_t - S_i$;

4) оценка параметров тренда по полученным уровням ряда (без сезонной компоненты);

5) оценка качества полученной модели.

Решение типового примера (аддитивная модель)

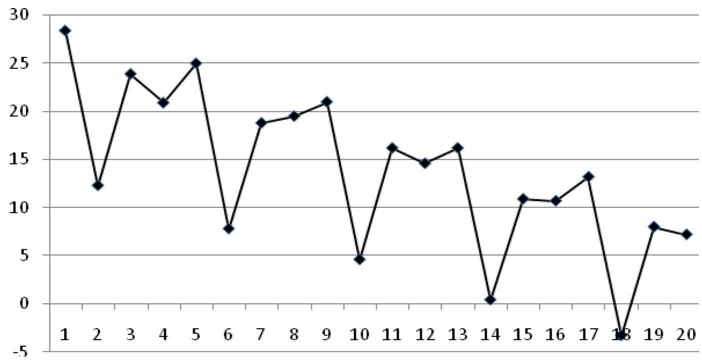
Задан временной ряд ежеквартальных значений переменной X:

t	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
X	28,4	12,3	23,9	20,9	25,0	7,8	18,8	19,5	21,0	4,6
t	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
X	16,2	14,6	16,2	0,4	10,9	10,7	13,2	-3,3	8,0	7,2

1. Построить график временного ряда.
2. Рассчитать автокорреляционную функцию.
3. Оценить параметры линейного тренда и сезонной компоненты.
4. Сделать прогноз на 21, 22, 23, 24 кварталы по полученной модели.

Решение

1. Построим график временного ряда:



Визуальный анализ позволяет сделать вывод о наличии сезонных колебаний, с периодом 4, наличие убывающего тренда и о возможности использования аддитивной модели, т.к. амплитуда колебаний со временем не меняется.

Замечание: для построения графика временного ряда в среде MS Excel необходимо выделить диапазон данных и добавить диаграмму *График*.

2. Рассчитаем автокорреляционную функцию.

Коэффициент автокорреляции первого порядка r_1:

$$r_1 = \frac{\operatorname{cov}(y_{t-1};\, y_t)}{\sigma_{y_{t-1}}\sigma_{y_t}} = \frac{\overline{y_{t-1}\cdot y_t} - \overline{y_{t-1}}\cdot\overline{y_t}}{\sqrt{\overline{y_{t-1}^2} - \overline{y_{t-1}}^2}\cdot\sqrt{\overline{y_t^2} - \overline{y_t}^2}}, \text{ где } \overline{y_t} = \frac{1}{n-1}\sum_{t=2}^{n} y_t,\ \overline{y_{t-1}} = \frac{1}{n-1}\sum_{t=1}^{n-1} y_t .$$

Заполним расчетную таблицу:

Таблица 3.1

t	y_t	y_{t-1}	$y_t \cdot y_{t-1}$	y_t^2	y_{t-1}^2
1.	28,4	-	-	-	-
2.	12,3	28,4	349,32	151,29	806,56
3.	23,9	12,3	293,97	571,21	151,29
4.	20,9	23,9	499,51	436,81	571,21
5.	25	20,9	522,5	625	436,81
6.	7,8	25	195	60,84	625
7.	18,8	7,8	146,64	353,44	60,84
8.	19,5	18,8	366,6	380,25	353,44
9.	21	19,5	409,5	441	380,25
10.	4,6	21	96,6	21,16	441
11.	16,2	4,6	74,52	262,44	21,16
12.	14,6	16,2	236,52	213,16	262,44
13.	16,2	14,6	236,52	262,44	213,16
14.	0,4	16,2	6,48	0,16	262,44
15.	10,9	0,4	4,36	118,81	0,16
16.	10,7	10,9	116,63	114,49	118,81
17.	13,2	10,7	141,24	174,24	114,49
18.	-3,3	13,2	-43,56	10,89	174,24
19.	8	-3,3	-26,4	64	10,89
20.	7,2	8	57,6	51,84	64
Среднее	**13,05**	**14,16**	**193,87**	**227,02**	**266,75**

Примечание: в столбце y_{t-1} находятся исходные уровни временного ряда, смещенные на одну строку вниз.

$$r_1 = \frac{193{,}87 - 13{,}05 \cdot 14{,}6}{\sqrt{227{,}02 - 13{,}05^2} \cdot \sqrt{266{,}75 - 14{,}16^2}} = 0{,}148$$

Аналогично, можно рассчитать остальные коэффициенты автокорреляции, получаем автокорреляционную функцию:

$$r_1 = 0{,}148;\ r_2 = 0{,}377;\ r_3 = 0{,}122;\ r_4 = 0{,}991;$$

$$r_5 = -0{,}021;\ r_6 = 0{,}251;\ r_7 = -0{,}029;\ r_8 = 0{,}996.$$

Значения коэффициентов автокорреляции подтверждают гипотезу о наличии сезонных колебаний с периодом 4, т.к. наибольшее значение имеют r_4 и r_8.

Замечание: Для последовательного расчета значений автокорреляционной функции в MS Excel целесообразно использовать функцию КОРРЕЛ(массив1;массив2). Так, если уровни исходного временного ряда располагаются в ячейках A1:A20, то для расчета коэффициентов автокорреляции можно вводить функции:

r_1: =КОРРЕЛ(A1:A19;A2:A20);

r_2: =КОРРЕЛ(A1:A18;A3:A20);

r_3: =КОРРЕЛ(A1:A17;A4:A20);

r_4: =КОРРЕЛ(A1:A16;A5:A20).

3. Проведем оценку тренда и сезонной компоненты временного ряда.

Результаты расчета коэффициентов автокорреляции показали, что данный временной ряд содержит сезонные колебания периодичностью 4.

Для выделения сезонной компоненты проведем выравнивание исходных уровней ряда методом скользящей средней. Заполним расчетную таблицу:

Таблица 3.2

t	y_t	Скользящая средняя за 4 квартала	Центрированная скользящая средняя	Оценка сезонной компоненты
1	**2**	**3**	**4**	**5**
1	28,4	-	-	-
2	12,3	21,38	-	-
3	23,9	20,53	20,95	2,95
4	20,9	19,40	19,96	0,94
5	25	18,13	18,76	6,24
6	7,8	17,78	17,95	-10,15
7	18,8	16,78	17,28	1,53
8	19,5	15,98	16,38	3,13
9	21	15,33	15,65	5,35
10	4,6	14,10	14,71	-10,11
11	16,2	12,90	13,50	2,70
12	14,6	11,85	12,38	2,23
13	16,2	10,53	11,19	5,01
14	0,4	9,55	10,04	-9,64
15	10,9	8,80	9,18	1,73
16	10,7	7,88	8,34	2,36
17	13,2	7,15	7,51	5,69
18	-3,3	6,28	6,71	-10,01
19	8	-	-	-
20	7,2	-	-	-

В столбце 3 рассчитаны средние значения четырех последовательных уровней ряда столбца 2:

– во второй строке: $21{,}38 = \dfrac{28{,}4 + 12{,}3 + 23{,}9 + 20{,}9}{4}$;

– в третьей строке: $20{,}53 = \dfrac{12{,}3 + 23{,}9 + 20{,}9 + 25}{4}$;

– в четвертой строке: $19{,}40 = \dfrac{23{,}9 + 20{,}9 + 25 + 7{,}8}{4}$ и т.д.

В столбце 4 рассчитаны средние значения двух последовательных средних из столбца 3:

– в третьей строке: $20{,}95 = \dfrac{21{,}38 + 20{,}53}{2}$;

– в четвертой строке: $19{,}96 = \dfrac{20{,}53 + 19{,}4}{2}$ и т.д.

В столбце 2 находятся исходные уровни ряда, а в столбце 4 соответствующие уровни, но без сезонного отклонения. Поэтому оценить сезонную компоненту можно разностью между соответствующими исходными и выровненными уровнями: в столбце 5 рассчитаны разности значений столбца 2 и столбца 4:

– в третьей строке: 23,9 – 20,95 = 2,95;

– в четвертой строке: 20,9 – 19,96 = 0,94 и т.д.

Таким образом, получили следующие оценки:

S_3: 2,95; 1,53; 2,7; 1,73.

S_4: 0,94; 3,13; 2,23; 2,36.

S_1: 6,24; 5,35; 5,01; 5,69.

S_2: -10,15; -10,11; -9,64; -10,01.

Будем считать, что сезонная компонента для каждого квартала является константой, причем за все четыре квартала суммарное воздействие сезонных компонент равно нулю: $S_1 + S_2 + S_3 + S_4 = 0$.

Найдем средние значения сезонных компонент и рассчитаем корректирующий коэффициент:

	Среднее значение	Корректирующий коэффициент	Окончательное значение сезонной компоненты
S_3	2,23	-0,00469	2,23
S_4	2,16		2,17
S_1	5,57		5,58
S_2	-9,98		-9,97

Корректирующий коэффициент рассчитан следующим образом:

$$k = \frac{S_1 + S_2 + S_3 + S_4}{4} = \frac{2{,}23 + 2{,}16 + 5{,}57 - 9{,}98}{4} = -0{,}00469 .$$

Корректирующий коэффициент вычли из среднего значения сезонной компоненты и получили окончательные значения:

$$S_1 = 5{,}58;\ S_2 = -9{,}97;\ S_3 = 2{,}23;\ S_4 = 2{,}17 .$$

Следующий шаг построения модели – оценка тренда. Для этого вычтем из исходных уровней временного ряда y_t сезонные компоненты и проведем аналитической выравнивание ряда – методом наименьших квадратов (МНК) найдем оценки параметров линейного тренда $y = a + bt$.

Оформим расчеты следующим образом:

Таблица 3.3

t	y_t	S	$y_t - S$	T = a+bt	$\hat{y}$ =T+S	E= $y_t - \hat{y}$	E^2	$(y_t - \bar{y}_t)^2$
1	2	3	4	5	6	7	8	9
1	28,4	5,58	22,82	22,93	28,51	-0,11	0,01	212,72
2	12,3	-9,97	22,27	21,97	12,00	0,30	0,09	2,30
3	23,9	2,23	21,67	21,01	23,24	0,66	0,43	101,71
4	20,9	2,17	18,73	20,05	22,22	-1,32	1,74	50,20
5	25	5,58	19,42	19,09	24,67	0,33	0,11	125,10
6	7,8	-9,97	17,77	18,13	8,16	-0,36	0,13	36,18
7	18,8	2,23	16,57	17,17	19,40	-0,60	0,36	24,85

1	2	3	4	5	6	7	8	9
8	19,5	2,17	17,33	16,21	18,38	1,12	1,25	32,32
9	21	5,58	15,42	15,25	20,83	0,17	0,03	51,62
10	4,6	-9,97	14,57	14,29	4,32	0,28	0,08	84,92
11	16,2	2,23	13,97	13,34	15,56	0,64	0,40	5,69
12	14,6	2,17	12,43	12,38	14,54	0,06	0,00	0,62
13	16,2	5,58	10,62	11,42	16,99	-0,79	0,63	5,69
14	0,4	-9,97	10,37	10,46	0,48	-0,08	0,01	179,96
15	10,9	2,23	8,67	9,50	11,73	-0,83	0,68	8,50
16	10,7	2,17	8,53	8,54	10,70	0,00	0,00	9,70
17	13,2	5,58	7,62	7,58	13,15	0,05	0,00	0,38
18	-3,3	-9,97	6,67	6,62	-3,35	0,05	0,00	292,92
19	8	2,23	5,77	5,66	7,89	0,11	0,01	33,81
20	7,2	2,17	5,03	4,70	6,87	0,33	0,11	43,76
Сумма							**6,09**	**1302,95**
Среднее	**13,82**							

Столбец 3 заполняется последовательными значениями S_1, S_2, S_3, S_4.

В столбце 4 находится разность исходных уровней ряда (столбец 2) и соответствующей сезонной компоненты S_1, S_2, S_3, S_4 (столбец 3). Полученные выровненные уровни ряда позволяют оценить параметры линейного тренда $y = a + bt$. Сделаем это двумя способами: 1) по формулам решения системны нормальных уравнений: $b = \dfrac{\overline{Y \cdot t} - \bar{Y} \cdot \bar{t}}{\overline{t^2} - \bar{t}^2}$, $a = \bar{Y} - b \cdot \bar{t}$[5]; 2) с помощью функции ЛИНЕЙН в MS Excel, которая рассчитывает параметры уравнения линейной регрессии.

1) Заполним расчетную таблицу для оценки параметров $y = a + bt$:

Таблица 3.4

	t	$Y = y_t - S$	$Y \cdot t$	t^2
	1	22,82	22,82	1
	2	22,27	44,55	4
	3	21,67	65,01	9
	4	18,73	74,93	16
	5	19,42	97,12	25
	6	17,77	106,64	36
	7	16,57	115,99	49
	8	17,33	138,66	64
	9	15,42	138,81	81
	10	14,57	145,73	100
	11	13,97	153,67	121
	12	12,43	149,19	144
	13	10,62	138,10	169
	14	10,37	145,23	196
	15	8,67	130,05	225

[5] Подробно алгоритм оценки параметров парной линейной регрессии приводится в соответствующей теме пособия.

	t	$Y = y_t - S$	$Y \cdot t$	t^2
	16	8,53	136,53	256
	17	7,62	129,60	289
	18	6,67	120,12	324
	19	5,77	109,64	361
	20	5,03	100,66	400
Среднее	**10,5**	**13,82**	**113,15**	**143,5**

Получаем параметры линейного тренда:

$$b = \frac{\overline{Y \cdot t} - \overline{Y} \cdot \bar{t}}{\overline{t^2} - \bar{t}^2} = \frac{113{,}15 - 13{,}82 \cdot 10{,}5}{143{,}5 - 10{,}5^2} = -0{,}96;$$

$$a = \overline{Y} - b \cdot \bar{t} = 13{,}82 + 0{,}96 \cdot 10{,}5 = 23{,}89.$$

Таким образом, уравнение тренда имеет вид: $y = 23{,}89 - 0{,}96t$ [6].

2) Также можно воспользоваться средствами MS Excel, в частности, функцией ЛИНЕЙН, которая проводит оценку методом наименьших квадратов (МНК) параметров уравнения линейной парной (множественной) регрессии.

Синтаксис: ЛИНЕЙН(известные_значения_*y*; известные_значения_*x*; конст; статистика).

Аргументы:

- *известные_значения_y*: множество значений результативного признака *Y*;
- *известные_значения_x*: множество значений факторных признаков X_i (необязательный аргумент); если этот аргумент опущен, то предполагается, что это массив {1; 2; 3; …} такого же размера, как и аргумент *известные_значения_y*;
- *конст*: логическое значение, которое указывает, требуется ли, чтобы свободный член a_0 был равен 0 (необязательный аргумент);
- *статистика*: логическое значение, которое указывает, требуется ли вернуть дополнительную статистику по регрессии (необязательный аргумент).

Результаты функции выводятся в диапазон ячеек 5х2 (5 строк, 2 столбца) в следующем виде:

Значение параметра *b*	Значение параметра *a*
Стандартное отклонение *b*	Стандартное отклонение *a*
Коэффициент детерминации R^2	Стандартное отклонение *y*
F-статистика	Число степеней свободы
Регрессионная сумма квадратов	Остаточная сумма квадратов

Если аргумент *статистика* = 0 (или пустой), то достаточно диапазона 1х2 (будут получены только оценки параметров регрессии). После ввода функции в ячейку необходимо выделить соответствующий диапазон (верхняя левая ячейка содержит введенную функцию), нажать F2, а затем комбинацию CTRL+SHIFT+ENTER.

[6] Наиболее простым способом получения уравнения тренда в табличном процессоре MS Excel является построение графика ряда и добавление линии тренда с отображением его уравнения

Если считать, что уровни ряда с исключенной сезонной компонентой $Y = y_t - S$ находятся в диапазоне ячеек D2:B21, то можно ввести функцию ЛИНЕЙН в следующем виде: =ЛИНЕЙН(D2:B21;;1;1). Получаем:

$b =$	-0,96	$a =$	23,89
$\sigma_b =$	0,02	$\sigma_a =$	0,27
$R^2 =$	0,99	$\sigma_y =$	0,58
$F_{набл} =$	1809,02	$df =$	18,00
RSS =	612,26	ESS =	6,09

Таким образом, получили уравнение тренда $y = 23{,}89 - 0{,}96t$. Большое значение коэффициента детерминации $R^2 = 0{,}99$ говорит о хорошем качестве полученного уравнения, другие статистики подтверждают этот вывод.

Вернемся к расчетной таблице 3.3.

Столбец 5 в ней рассчитывается по полученному уравнению тренда:

при t = 1: $y = 23{,}89 - 0{,}96 \cdot 1 = 22{,}93$;

при t = 2: $y = 23{,}89 - 0{,}96 \cdot 2 = 21{,}97$;

при t = 3: $y = 23{,}89 - 0{,}96 \cdot 3 = 21{,}01$ и т.д.

В столбце 6 рассчитаны теоретические (смоделированные, объясненные) значения уровней ряда, как сумма сезонной и трендовой компонент $\hat{y}$=T+S (сумма значений столбца 3 и столбца 5).

В столбце 7 найдем значения случайной компоненты E = y_t – T – S (разность значений столбца 2 и столбца 6). Таким образом, определение компонент аддитивной модели ряда можно считать завершенной.

Проведем оценку качества полученной модели. Найдем отношение остаточной суммы квадратов отклонений исходных уровней ряда от соответствующих теоретических значений $\sum E^2$ к общей сумме квадратов отклонений исходных уровней ряда от среднего значения $\sum (y_t - \overline{y_t})^2$ (столбец 8 и столбец 9) расчетной таблицы 3.3. Т.к. $\frac{\sum E^2}{\sum (y_t - \overline{y_t})^2} = \frac{6{,}09}{1302{,}95} = 0{,}005$, т.е. 0,5%, то аддитивная модель объясняет 99,5% общей вариации уровней данного временного ряда.

4. Сделаем прогноз на 21, 22, 23, 24 кварталы по полученной модели.

при t = 21: $\hat{y} = S_1 + T_{21} = 5{,}58 + (23{,}89 - 0{,}96 \cdot 21) = 9{,}31$;

при t = 22: $\hat{y} = S_2 + T_{22} = -9{,}97 + (23{,}89 - 0{,}96 \cdot 22) = -7{,}2$;

при t = 23: $\hat{y} = S_3 + T_{23} = 2{,}23 + (23{,}89 - 0{,}96 \cdot 23) = 4{,}04$;

при t = 24: $\hat{y} = S_4 + T_{24} = 2{,}17 + (23{,}89 - 0{,}96 \cdot 24) = 3{,}02$;

Ответ:

1) Автокорреляционная функция:

$$r_1 = 0{,}148;\ r_2 = 0{,}377;\ r_3 = 0{,}122;\ r_4 = 0{,}991;$$

$$r_5 = -0{,}021;\ r_6 = 0{,}251;\ r_7 = -0{,}029;\ r_8 = 0{,}996.$$

2) Сезонная компонента: $S_1 = 5{,}58;\ S_2 = -9{,}97;\ S_3 = 2{,}23;\ S_4 = 2{,}17$.

3) Уравнение тренда: $y = 23{,}89 - 0{,}96t$.

4) Прогнозные значения: $\hat{y}_{21} = 9{,}31$; $\hat{y}_{22} = -7{,}2$; $\hat{y}_{23} = 4{,}04$; $\hat{y}_{24} = 3{,}02$.

Задание для самостоятельной работы

Имеются поквартальные значения некоторого экономического показателя (приложение №3). Требуется:

1. Построить график временного ряда[7].

2. Получить аддитивную модель ряда:

– оценить сезонную компоненту (S_1, S_2, S_3, S_4);

– оценить параметры линейного тренда $y = a + bt$, исключив из исходных уровней ряда сезонную компоненту.

3. Сделать прогноз на 17, 18, 19, 20 кварталы по полученной модели.

[7] График целесообразно построить как диаграмму в MS Excel

Список литературы

1. Айвазян С.А., Мхитарян В.С. Прикладная статистика и основы эконометрики. – М.: ЮНИТИ, 1998. - 1005 с.
2. Вуколов Э.А. Основы статистического анализа. Практикум по статистическим методам и исследованию операций с использованием пакетов STATISTICA и EXCEL. – М.: Форум, 2008. - 464 с.
3. Дорохина Е.Ю., Преснякова Л.Ф., Тихомиров Н.П. Сборник задач по эконометрике: Учебное пособие для студентов экономических вузов. - М.: Издательство «Экзамен», 2003. - 224 с.
4. Елисеева И.И. Практикум по эконометрике: Учебное пособие. – М.: Финансы и статистика, 2005 - 192 с.
5. Елисеева И.И. Эконометрика: учебник. – М.: Финансы и статистика, 2004 - 344 с.
6. Кремер Н.Ш., Путко Б.А. Эконометрика: учебник для вузов. – М.: ЮНИТИ, 2005 - 311 с.
7. Макарова Н.В., Трофимец В.Я. Статистика в Excel. – М.: Финансы и статистика, 2002. - 368 с.
8. Тихомиров Н.П., Дорохина Е.Ю. Эконометрика: учебник. – М.: Экзамен, 2003 - 512 с.

Приложение 1

Статистические данные к теме «Линейная парная регрессия»

Все данные по регионам представлены за 2008 год[8]

№	Регион	0	1	2	3	4	5	6	7	8	9	10
1	Республика Адыгея	4,4	1,6	16	10,1	117	28,3	11	124	1,7	94	8,1
2	Республика Дагестан	27,1	5,7	173	7,6	513	3603,9	45	625	10,8	3109	24,6
3	Кабардино-Балкарская Республика	8,9	2,2	74	9,1	178	44,9	15	58	3,8	454	11,7
4	Республика Калмыкия	2,8	1,7	24	9,1	67	21,9	6	106	1,4	368	13,5
5	Карачаево-Черкесская Республика	4,3	2,0	32	9,4	112	17,1	8	111	1,6	44	6,4
6	Республика Северная Осетия - Алания	7,0	2,2	36	9,2	203	23,8	18	214	4,9	230	11,3
7	Чеченская Республика	12,4	1,7	168	11,8	276	0,2	8	138	3,0	397	9,0
8	Краснодарский край	51,4	13,9	126	13,2	1408	381,7	109	1469	22,0	3253	128,2
9	Ставропольский край	27,1	9,8	108	11,1	688	165,9	57	806	12,1	3091	55,2
10	Астраханская область	10,1	4,6	40	12,3	236	92,2	20	261	6,7	873	18,9
11	Волгоградская область	26,0	10,5	108	12,0	713	192,0	55	777	13,2	823	53,3
12	Ростовская область	42,4	17,2	147	12,5	1220	235,6	89	1256	16,2	2320	98,1
13	Ивановская область	10,7	4,3	29	10,2	324	126,2	25	379	5,5	185	28,9
14	Калужская область	10,0	4,9	26	14,1	291	110,0	24	265	4,0	135	26,3
15	Костромская область	6,9	4,8	19	11,5	205	90,1	16	365	2,5	1671	18,1

Расшифровка столбцов:

0 – Численность населения (оценка на конец года; сотни тысяч человек);
1 – Численность работников органов местного самоуправления (на конец года; тысяч человек);
2 – Численность безработных (тысяч человек);
3 – Среднемесячная номинальная начисленная заработная плата работников организаций (тысяч рублей);
4 – Численность пенсионеров (тысяч человек);
5 – Розничная продажа водки и ликероводочных изделий (десятки тысяч декалитров);
6 – Жилищный фонд (общая площадь жилых помещений; миллионов квадратных метров);
7 – Число дошкольных образовательных учреждений;
8 – Численность врачей всех специальностей (на конец года; тысяч человек);
9 – Использование свежей воды (миллионов кубических метров);
10 – Число предприятий и организаций (на конец года; тысяч).

[8] Источник: http://www.gks.ru

Продолжение таблицы

№	Регион	11	12	13	14	15	16	17	18	19	20	21
1	Республика Адыгея	16,1	9,6	21,2	11,9	3,1	6,7	54,3	41,6	33,2	21,8	80,8
2	Республика Дагестан	205,3	40,3	27,5	11,6	21,8	37,5	28,2	88,7	333,5	31,3	870,9
3	Кабардино-Балкарская Республика	27,4	19,5	29,7	19,6	12,5	16,6	67,2	35,1	192,0	6,9	296,3
4	Республика Калмыкия	5,5	9,7	29,2	22,5	0,9	1,5	43,7	10,6	10,0	0,3	18,0
5	Карачаево-Черкесская Республика	3,3	12,4	11,3	6,5	13,5	2,0	14,2	21,2	167,2	2,4	51,6
6	Республика Северная Осетия - Алания	8,9	14,2	16,8	12,2	7,5	3,1	52,1	10,1	130,4	6,6	32,8
7	Чеченская Республика	1,5	8,5	18,4	11,9	3,0	5,0	18,2	0,2	18,8	3,7	21,4
8	Краснодарский край	339,6	185,3	368,9	224,1	52,9	57,0	1163,4	318,0	512,0	81,0	613,8
9	Ставропольский край	161,0	76,4	294,5	225,5	25,9	18,1	841,3	183,4	258,8	35,7	199,5
10	Астраханская область	78,7	15,2	7,2	2,4	8,7	18,0	4,0	38,6	151,7	10,4	479,4
11	Волгоградская область	127,0	70,7	316,4	214,9	32,4	29,8	518,3	124,2	361,3	18,2	713,9
12	Ростовская область	209,3	115,8	455,2	286,6	36,0	38,4	887,1	213,7	380,2	75,0	552,9
13	Ивановская область	67,2	8,9	23,9	7,3	10,9	3,4	10,1	47,8	108,0	13,4	87,7
14	Калужская область	59,7	18,4	34,1	9,8	23,6	5,2	19,1	67,5	323,1	20,6	105,7
15	Костромская область	31,0	12,2	25,9	7,5	10,9	3,7	8,3	30,1	160,2	8,1	105,2

Расшифровка столбцов:

11 – Производство хлеба и хлебобулочных изделий (тысяч тонн);
12 – Продукция сельского хозяйства (в хозяйствах всех категорий; в фактически действовавших ценах; млрд. рублей);
13 – Посевные площади всех сельскохозяйственных культур (в хозяйствах всех категорий; десятки тысяч гектаров);
14 – Посевные площади зерновых и зернобобовых культур (в хозяйствах всех категорий; десятки тысяч гектаров);
15 – Посевные площади картофеля (в хозяйствах всех категорий; тысяч гектаров);
16 – Посевные площади овощей (в хозяйствах всех категорий; тысяч гектаров);
17 – Валовой сбор зерна (в весе после доработки) (в хозяйствах всех категорий; десятки тысяч тонн);
18 – Розничная продажа коньяков (тысяч декалитров);
19 – Валовой сбор картофеля (в хозяйствах всех категорий; тысяч тонн);
20 – Ввоз бензина автомобильного (десятки тысяч тонн);
21 – Валовой сбор овощей (в хозяйствах всех категорий; тысяч тонн).

Продолжение таблицы

№	Регион	22	23	24	25	26	27	28	29	30	31
1	Республика Адыгея	6,3	15,4	4,9	10,4	5,3	6,5	0,8	0,5	70	25,6
2	Республика Дагестан	52,3	96,7	90,6	52,6	43,8	45,7	9,1	2,7	523	217,3
3	Кабардино-Балкарская Республика	2,4	71,0	22,9	31,2	10,8	6,5	2,5	1,3	253	48,0
4	Республика Калмыкия	0,9	0,5	33,2	16,0	2,6	2,6	0,6	0,1	91	7,4
5	Карачаево-Черкесская Республика	0,9	6,7	20,5	22,1	6,1	5,9	0,9	2,2	121	22,6
6	Республика Северная Осетия - Алания	2,0	11,4	12,9	18,6	12,5	11,4	2,0	2,0	166	37,4
7	Чеченская Республика	22,1	11,4	23,4	25,9	6,6	28,0	0,6	2,2	173	25,2
8	Краснодарский край	77,4	256,1	68,7	136,8	137,9	182,3	39,4	32,9	1053	500,7
9	Ставропольский край	14,2	41,1	38,7	61,1	57,2	31,1	10,7	10,6	401	203,6
10	Астраханская область	37,2	11,2	23,4	15,3	18,5	28,8	3,8	8,0	268	77,7
11	Волгоградская область	28,6	143,6	31,7	47,9	71,0	55,7	8,2	17,6	373	186,1
12	Ростовская область	38,6	102,2	58,9	99,6	88,1	92,9	20,1	27,4	1336	423,4
13	Ивановская область	1,6	15,0	8,8	18,0	16,8	12,1	1,7	2,1	257	57,9
14	Калужская область	1,5	17,2	12,8	22,6	21,0	23,5	6,3	2,1	989	80,7
15	Костромская область	1,0	10,6	8,3	15,7	10,1	9,7	1,5	2,2	17	38,1

Расшифровка столбцов:
22 – Ввоз пиломатериалов (тысяч кубических метров);
23 – Валовой сбор плодов и ягод (в хозяйствах всех категорий; тысяч тонн);
24 – Поголовье крупного рогатого скота (в хозяйствах всех категорий; на конец года; десятки тысяч голов);
25 – Производство молока (в хозяйствах всех категорий; десятки тысяч тонн);
26 – Объем платных услуг населению (миллиардов рублей);
27 – Объем работ, выполненных по виду экономической деятельности «Строительство» (в фактически действовавших ценах; млрд. рублей);
28 – Ввод в действие жилых домов (десятки тысяч квадратных метров общей площади);
29 – Отправление грузов железнодорожным транспортом общего пользования (миллионов тонн);
30 – Численность исследователей с учеными степенями (человек);
31 – Оборот розничной торговли (в фактически действовавших ценах; млрд. рублей).

Приложение 2

Статистические данные к теме «Линейная множественная регрессия»

№	Вариант 1 ВАЗ 2107			Вариант 2 ВАЗ 2107			Вариант 3 ВАЗ 2107		
	y	x_1	x_2	y	x_1	x_2	y	x_1	x_2
1	55	14	100	62	10	85	90	7	100
2	80	11	150	45	16	90	70	6	107
3	55	14	195	65	10	130	45	16	90
4	83	8	118	72	9	88	80	6	120
5	86	8	77	55	17	62	80	6	60
6	107	7	100	60	6	64	85	6	96
7	85	7	100	90	7	73	97	6	59
8	50	14	180	70	7	63	70	6	115
9	75	7	79	85	6	115	70	9	110
10	50	17	90	70	8	80	50	6	117
11	95	7	100	73	7	98	68	11	71
12	40	14	184	75	9	110	60	9	90
13	100	6	76	82	8	90	90	8	136
14	40	18	83	48	17	68	82	7	83
15	60	12	100	75	7	130	50	16	170

№	Вариант 4 ВАЗ 2107			Вариант 5 ВАЗ 2107			Вариант 6 ВАЗ 2109		
	y	x_1	x_2	y	x_1	x_2	y	x_1	x_2
1	60	14	70	32	13	75	150	5	110
2	60	11	150	85	7	75	135	8	115
3	87	8	60	75	10	90	130	9	118
4	90	14	70	90	14	70	85	5	150
5	75	10	130	60	11	75	145	7	94
6	100	6	59	100	7	70	130	6	110
7	67	9	180	100	6	76	90	11	100
8	80	8	115	100	6	60	110	9	120
9	45	13	75	48	7	138	150	5	110
10	50	15	138	32	13	75	165	6	93
11	100	7	70	73	6	75	107	10	105
12	90	7	65	85	7	150	117	8	79
13	65	10	91	80	10	80	143	5	87
14	55	14	100	45	14	85	100	10	130
15	80	6	70	100	6	76	135	8	80

№	Вариант 7 ВАЗ 2109			Вариант 8 ВАЗ 2109			Вариант 9 ВАЗ 2109		
	y	x_1	x_2	y	x_1	x_2	y	x_1	x_2
1	105	9	105	55	17	181	96	11	88
2	140	8	160	90	9	180	100	12	160
3	130	7	85	100	7	153	140	8	160
4	80	12	160	143	6	112	110	9	173
5	120	10	184	56	19	100	135	10	140
6	135	6	92	107	11	140	70	19	192
7	103	10	160	105	10	85	81	11	61
8	100	11	98	115	8	130	105	9	95
9	130	7	90	55	16	150	60	11	80
10	140	8	150	161	6	88	83	18	87
11	110	7	95	65	14	100	65	13	175
12	130	6	103	83	11	88	90	9	120
13	80	16	85	117	7	65	90	17	190
14	155	5	71	55	14	100	80	12	112
15	110	11	98	70	15	160	70	15	100

№	Вариант 10 ВАЗ 2109			Вариант 11 ВАЗ 2110			Вариант 12 ВАЗ 2110		
	y	x_1	x_2	y	x_1	x_2	y	x_1	x_2
1	130	7	120	175	5	70	255	3	70
2	135	8	105	146	8	110	170	5	120
3	95	17	103	120	10	175	220	4	100
4	102	9	120	220	4	62	107	12	100
5	43	15	180	172	5	84	135	10	110
6	110	14	100	170	7	77	125	7	85
7	139	6	85	190	4	83	107	11	102
8	95	13	70	80	10	65	138	10	170
9	70	16	137	143	8	120	175	5	150
10	70	18	165	167	6	88	130	9	110
11	107	10	75	150	7	89	100	11	130
12	97	11	100	195	4	83	218	5	104
13	140	9	120	131	8	108	170	6	100
14	43	15	180	175	4	110	135	9	125
15	110	8	60	150	9	90	120	11	168

№	Вариант 13 ВАЗ 2110			Вариант 14 ВАЗ 2110			Вариант 15 ВАЗ 2112		
	y	x_1	x_2	y	x_1	x_2	y	x_1	x_2
1	130	9	100	120	11	168	175	4	78
2	195	4	71	165	7	90	270	3	62
3	105	12	80	135	10	160	160	8	140
4	118	11	179	120	9	55	235	3	64
5	120	9	103	95	12	150	200	5	84
6	146	8	100	160	8	100	200	4	110
7	165	6	130	120	12	95	150	8	82
8	110	11	180	140	9	150	180	5	100
9	140	8	95	135	10	120	200	4	110
10	195	4	71	170	6	167	203	5	94
11	165	7	95	110	13	160	145	9	145
12	123	10	120	100	11	130	160	7	100
13	180	4	91	140	9	160	140	8	119
14	130	10	160	135	10	98	165	7	100
15	98	13	150	190	5	155	220	3	99

№	Вариант 16 ВАЗ 2112			Вариант 17 ВАЗ 2112			Вариант 18 ВАЗ 2112		
	y	x_1	x_2	y	x_1	x_2	y	x_1	x_2
1	175	5	90	175	6	80	150	8	110
2	160	8	80	140	8	115	115	6	80
3	130	10	136	145	8	110	140	10	105
4	215	4	70	225	4	65	137	8	106
5	170	7	107	165	5	98	210	2	67
6	155	8	111	200	5	80	160	7	136
7	250	3	120	170	6	77	180	9	100
8	187	5	80	190	4	70	210	5	80
9	220	3	51	130	10	140	195	5	120
10	140	8	130	240	3	60	130	9	107
11	196	5	64	175	5	95	180	7	110
12	168	6	80	197	5	75	150	9	105
13	170	5	62	240	3	70	235	3	62
14	140	8	195	185	6	80	150	7	105
15	200	5	84	190	5	110	227	2	53

№	Вариант 19 ВАЗ 2114			Вариант 20 ВАЗ 2114			Вариант 21 ВАЗ 2115		
	y	x_1	x_2	y	x_1	x_2	y	x_1	x_2
1	175	3	70	175	5	140	155	6	140
2	222	2	56	190	3	74	150	7	116
3	105	5	70	170	5	63	180	4	100
4	190	3	68	203	2	60	179	3	65
5	150	6	72	160	6	130	167	3	38
6	130	5	103	190	3	74	180	6	77
7	155	6	130	176	5	140	172	5	75
8	175	5	78	180	3	83	160	6	90
9	160	4	110	169	6	87	200	4	35
10	170	5	60	170	4	70	150	7	118
11	125	6	79	155	5	110	165	4	40
12	155	4	100	205	3	65	165	6	90
13	200	3	70	150	5	65	125	7	89
14	177	4	65	157	6	79	130	9	95
15	165	5	77	165	4	66	115	10	132

№	Вариант 22 ВАЗ 2115			Вариант 23 ВАЗ Лада Калина			Вариант 24 ВАЗ 2106		
	y	x_1	x_2	y	x_1	x_2	y	x_1	x_2
1	139	6	83	185	4	61	65	12	160
2	130	9	100	205	3	47	47	17	125
3	170	5	75	170	5	130	60	7	160
4	145	4	89	210	3	45	50	14	73
5	189	3	65	233	2	120	65	13	197
6	120	10	140	215	3	50	57	12	120
7	100	5	80	170	4	62	85	6	105
8	155	6	89	200	3	38	60	13	60
9	150	5	50	235	1	18	35	16	120
10	180	4	60	200	2	42	65	12	160
11	160	5	120	210	3	20	60	13	60
12	220	2	40	178	4	104	50	16	180
13	150	6	120	215	3	20	40	17	90
14	135	7	195	195	4	55	60	17	125
15	165	5	91	200	3	40	75	7	130

№	Вариант 25 ВАЗ 2111			Вариант 26 ВАЗ 2111			Вариант 27 ВАЗ 2108		
	y	x_1	x_2	y	x_1	x_2	y	x_1	x_2
1	260	3	40	136	8	97	96	10	142
2	190	3	43	220	4	42	95	10	150
3	178	5	89	133	9	168	45	17	82
4	119	11	117	230	4	45	85	11	170
5	230	4	45	210	4	42	60	18	100
6	125	9	130	180	4	96	70	11	100
7	175	7	45	100	10	85	75	17	73
8	185	4	65	220	3	80	60	13	110
9	242	2	35	210	2	22	65	17	100
10	175	7	45	135	11	90	85	10	107
11	185	4	65	230	4	40	120	8	190
12	242	2	35	155	8	192	65	15	130
13	190	5	100	180	6	120	60	16	60
14	150	6	50	163	6	125	85	11	195
15	125	9	130	135	7	124	80	13	86

№	Вариант 28 Audi 100			Вариант 29 Audi 80			Вариант 30 Audi A4		
	y	x_1	x_2	y	x_1	x_2	y	x_1	x_2
1	78	26	190	180	19	350	270	11	130
2	70	27	150	110	23	160	220	15	178
3	100	26	200	75	24	244	530	7	180
4	160	22	110	80	23	300	330	11	170
5	75	23	150	130	22	300	350	9	135
6	150	17	275	190	17	224	650	5	60
7	65	24	350	180	20	60	444	8	106
8	120	21	260	140	20	200	460	7	144
9	175	19	250	120	25	292	490	6	120
10	75	23	150	70	22	400	530	8	105
11	150	17	275	90	22	150	450	7	140
12	120	21	260	210	16	192	690	5	107
13	105	25	270	90	21	235	410	9	110
14	185	18	277	120	23	180	235	14	198
15	160	18	400	120	19	260	440	8	155

Приложение 3

Временные ряды заданы по вариантам

	Варианты									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
t	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**
1	6,1	3,6	5,2	4,9	5,4	5,2	4,2	4,9	4,2	3,1
2	2,8	2,1	2,9	2,8	2,6	2,8	1,6	1,9	2,0	1,4
3	8,4	7,5	8,7	8,6	8,8	8,7	8,2	8,0	8,4	8,6
4	0,1	0,3	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,4
5	8,1	8,1	8,5	7,9	7,9	8,2	8,5	7,9	8,4	8,9
6	5,6	6,3	5,6	5,2	5,2	5,4	6,1	5,3	6,3	7,4
7	11,0	13,1	10,8	11,3	10,4	11,8	12,7	10,6	12,4	13,3
8	2,8	4,9	2,9	2,3	2,5	3,0	4,2	2,7	4,0	5,5
9	10,8	13,6	10,0	10,1	9,6	10,8	12,3	9,9	12,0	13,5
10	8,1	11,7	8,2	8,0	7,7	7,7	10,5	8,2	10,1	12,4
11	14,2	17,1	13,8	12,8	13,4	14,0	16,5	13,7	16,1	18,4
12	4,9	9,9	5,0	5,3	4,6	5,5	8,8	5,6	8,3	11,3
13	12,7	18,4	13,0	12,6	12,4	13,0	17,3	12,7	16,5	18,9
14	10,4	15,9	10,7	10,2	10,1	10,6	15,0	10,9	13,9	17,8
15	16,6	22,7	16,3	15,9	15,7	15,8	21,3	16,0	20,1	23,5
16	7,9	15,0	8,0	7,2	7,6	7,8	13,5	7,7	12,2	16,1

	Варианты									
	11	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
t	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**
1	5,9	4,4	4,8	6,2	4,0	4,4	4,5	4,5	5,4	4,6
2	3,5	2,1	3,0	4,6	1,4	2,9	2,6	3,0	3,0	2,5
3	8,4	8,4	9,5	10,7	8,6	9,1	8,1	8,3	8,2	8,9
4	0,1	0,9	2,4	3,1	0,4	2,0	0,1	0,1	0,1	1,6
5	7,4	8,7	10,9	11,4	9,0	10,6	7,7	8,1	7,4	9,0
6	4,9	7,3	8,6	10,3	7,1	9,1	4,9	4,9	4,5	7,5
7	11,0	14,0	15,4	16,4	12,9	15,4	10,6	11,5	10,2	13,6
8	2,2	6,2	7,4	8,6	5,7	7,6	2,1	2,6	1,2	5,8
9	9,6	14,6	15,8	17,2	13,7	15,6	10,4	10,9	9,0	14,1
10	7,3	12,1	14,4	15,6	12,4	14,4	8,1	8,7	6,3	11,4
11	12,2	18,7	20,5	22,1	19,0	20,5	13,2	14,1	11,7	17,7
12	4,4	11,4	13,3	14,2	10,4	13,5	5,0	5,5	3,3	9,9
13	11,6	19,6	22,1	22,7	19,1	22,0	12,7	13,5	10,0	18,2
14	8,8	18,3	19,9	20,8	17,4	19,7	10,4	10,6	7,4	16,6
15	14,4	24,4	26,1	27,8	23,4	26,4	15,9	16,7	13,3	22,0
16	6,5	16,1	19,2	20,0	16,4	18,3	8,2	8,3	4,9	14,0

	Варианты									
	21	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**
t	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**
1	6,1	4,3	6,5	4,1	4,3	5,4	4,8	4,6	6,1	5,2
2	4,6	2,1	3,1	2,4	3,4	2,4	2,6	2,5	3,5	3,6
3	11,7	7,8	8,1	7,8	10,1	8,6	8,8	8,8	8,6	9,2
4	3,8	0,9	0,1	0,1	2,3	0,1	0,1	0,1	0,1	1,5
5	12,3	8,5	7,1	7,6	10,2	7,2	7,8	8,3	7,4	9,9
6	10,7	6,6	3,6	6,0	8,9	5,3	6,2	6,1	4,8	8,8
7	17,1	13,4	9,1	12,1	15,5	10,4	11,4	11,4	9,7	14,9
8	10,0	4,8	0,6	3,9	7,7	2,4	3,4	3,1	1,4	7,8
9	18,2	13,6	7,8	11,2	16,1	9,8	10,9	11,2	8,4	15,8
10	16,9	11,3	4,5	10,0	14,1	7,0	9,1	8,8	5,9	14,1
11	23,3	18,2	10,2	15,1	20,7	13,2	14,7	15,0	11,6	20,4
12	15,4	9,7	1,6	7,2	12,8	5,1	6,3	6,5	2,6	13,1
13	24,0	18,5	8,2	15,5	21,8	12,1	14,7	14,5	10,3	22,1
14	22,8	15,9	5,5	13,1	19,7	9,9	11,8	11,8	7,1	20,0
15	29,6	22,6	10,6	18,6	26,1	15,9	17,7	17,7	12,5	26,3
16	22,1	15,2	1,8	10,7	18,8	7,0	10,3	9,5	4,2	18,4

Приложение 4

Таблица значений F-критерия Фишера при уровне значимости $\alpha = 0,05$

k_2 \ k_1	1	2	3	4	5	6	8	12	24	∞
1	161,5	199,5	215,7	224,6	230,2	233,9	238,9	243,9	249,0	254,3
2	18,51	19,00	19,16	19,25	19,30	19,33	19,37	19,41	19,45	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,84	8,74	8,64	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,04	5,91	5,77	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,82	4,68	4,53	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4,00	3,84	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,73	3,57	3,41	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,28	3,12	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,23	3,07	2,90	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,07	2,91	2,74	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	2,95	2,79	2,61	2,40
12	4,75	3,88	3,49	3,26	3,11	3,00	2,85	2,69	2,50	2,30
13	4,67	3,80	3,41	3,18	3,02	2,92	2,77	2,60	2,42	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,70	2,53	2,35	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,64	2,48	2,29	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,59	2,42	2,24	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,55	2,38	2,19	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,51	2,34	2,15	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,48	2,31	2,11	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,45	2,28	2,08	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,42	2,25	2,05	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,40	2,23	2,03	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,38	2,20	2,00	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,36	2,18	1,98	1,73
25	4,24	3,38	2,99	2,76	2,60	2,49	2,34	2,16	1,96	1,71
26	4,22	3,37	2,98	2,74	2,59	2,47	2,32	2,15	1,95	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,30	2,13	1,93	1,67
28	4,20	3,34	2,95	2,71	2,56	2,44	2,29	2,12	1,91	1,65
29	4,18	3,33	2,93	2,70	2,54	2,43	2,28	2,10	1,90	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,27	2,09	1,89	1,62
35	4,12	3,26	2,87	2,64	2,48	2,37	2,22	2,04	1,83	1,57
40	4,08	3,23	2,84	2,61	2,45	2,34	2,18	2,00	1,79	1,51
45	4,06	3,21	2,81	2,58	2,42	2,31	2,15	1,97	1,76	1,48
50	4,03	3,18	2,79	2,56	2,40	2,29	2,13	1,95	1,74	1,44
60	4,00	3,15	2,76	2,52	2,37	2,25	2,10	1,92	1,70	1,39
70	3,98	3,13	2,74	2,50	2,35	2,23	2,07	1,89	1,67	1,35
80	3,96	3,11	2,72	2,49	2,33	2,21	2,06	1,88	1,65	1,31
90	3,95	3,10	2,71	2,47	2,32	2,20	2,04	1,86	1,64	1,28
100	3,94	3,09	2,70	2,46	2,30	2,19	2,03	1,85	1,63	1,26

Приложение 5

Критические значения t-критерия Стьюдента при уровне значимости 0,10, 0,05, 0,01 (двухсторонний)

Число степеней свободы d.f.	α			Число степеней свободы d.f.	α		
	00,10	0,05	0,01		00,10	0,05	0,01
1	6,3138	12,706	63,657	18	1,7341	2,1009	2,8784
2	2,9200	4,3027	9,9248	19	1,7291	2,0930	2,8609
3	2,3534	3,1825	5,8409	20	1,7247	2,0860	2,8453
4	2,1318	2,7764	4,5041	21	1,7207	2,0796	2,8314
5	2,0150	2,5706	4,0321	22	1,7171	2,0739	2,8188
6	1,9432	2,4469	3,7074	23	1,7139	2,0687	2,8073
7	1,8946	2,3646	3,4995	24	1,7109	2,0639	2,7969
8	1,8595	2,3060	3,3554	25	1,7081	2,0595	2,7874
9	1,8331	2,2622	3,2498	26	1,7056	2,0555	2,7787
10	1,8125	2,2281	3,1693	27	1,7033	2,0518	2,7707
11	1,7959	2,2010	3,1058	28	1,7011	2,0484	2,7633
12	1,7823	2,1788	3,0545	29	1,6991	2,0452	2,7564
13	1,7709	2,1604	3,0123	30	1,6973	2,0423	2,7500
14	1,7613	2,1448	2,9768	40	1,6839	2,0211	2,7045
15	1,7530	2,1315	2,9467	60	1,6707	2,0003	2,6603
16	1,7459	2,1199	2,9208	120	1,6577	1,9799	2,6174
17	1,7396	2,1098	2,8982	∞	1,6449	1,9600	2,5758

Деркач Дмитрий Васильевич

ЭКОНОМЕТРИКА: задания и методические рекомендации по выполнению самостоятельных работ для студентов ОЗО

Подписано в печать: 10.09.2010 г. Формат 60×84/16
Гарнитура Times New Roman. Тираж 100 экз. Усл.печ.л. 2,6. Уч.изд.л. 2,7.
Отпечатано на множительной технике математического факультета АГПУ

Армавирский государственный педагогический университет
Математический факультет
Кафедра алгебры, геометрии и методики преподавания математики
352900, г. Армавир, ул. Р.Люксембург, 159