Министерство образования и науки
Российской Федерации
Рубцовский индустриальный институт (филиал)
ФГБОУ ВПО «Алтайский государственный технический
университет им. И.И. Ползунова»

Н.В. Рассказова

ЭКОНОМЕТРИКА. НАЧАЛЬНЫЙ КУРС

Методическое пособие
к лабораторным и контрольным работам
для студентов направления подготовки «Экономика»
заочной формы обучения

Рубцовск 2015

УДК 519.2 + 330.43

Рассказова Н.В. Эконометрика. Начальный курс: Методическое пособие к лабораторным и контрольным работам для студентов направления подготовки «Экономика» заочной формы обучения / Рубцовский индустриальный институт – Рубцовск, 2015. – 43 с.

Методическое пособие содержит задания для лабораторных работ и необходимые расчетные соотношения для их выполнения. Основное внимание уделяется реализации этих соотношений в табличном процессоре Microsoft Office Excel. Приведены задания для контрольных работ и вопросы для подготовки к зачету.

Рассмотрены и одобрены
на заседании НМС РИИ
Протокол № 1 от 19.02.15

Рецензент: доцент каф. высш. матем., физ. и хим., к.п.н. И.И. Кулешова

© Рубцовский индустриальный институт, 2015

СОДЕРЖАНИЕ

ВВЕДЕНИЕ

Построение эконометрических моделей обуславливает большой объем вычислений. Удобной универсальной вычислительной средой для решения задач эконометрики является *табличный процессор Excel*. В данном методическом пособии изложены (в форме лабораторных работ) численные методики решения основных задач парного и множественного регрессионного анализа в вычислительной среде табличного процессора Excel.

Каждая лабораторная работа посвящена решению определенной задачи (или подзадачи) эконометрики. Для каждой работы приводится алгоритм решения рассматриваемой в работе задачи (т.е. формулы или расчетные соотношения), а затем дается фрагмент документа Excel (версия XP), реализующий алгоритм решения задачи. При этом алгоритм решения может быть реализован путем программирования арифметических или логических выражений в ячейках электронной таблицы или путем обращения к «стандартным» функциям или модулям Excel. Поэтому предполагается, что читатель знаком с адресацией ячеек (относительной, абсолютной и смешанной), арифметическими операциями и программированием простейших выражений в ячейках Excel.

ОБЩИЕ УКАЗАНИЯ К ВЫПОЛНЕНИЮ КОНТРОЛЬНЫХ РАБОТ

1. Контрольную работу следует выполнять на листах формата А4.
2. На титульном листе работы должен быть указан номер варианта, который определяется по последней цифре шифра зачетной книжки (**1** – 1-й вариант, **2** – 2-й вариант,…, **9** – 9-й вариант, **0** – 10-й вариант).
3. Все используемые эконометрические формулы и расчеты должны сопровождаться комментариями к ним и интерпретацией полученных результатов.
4. Полученные расчетные таблицы Excel необходимо привести в тексте работы или оформить в качестве приложения.

Тема 1. ЛИНЕЙНАЯ ПАРНАЯ РЕГРЕССИЯ

Эта тема включает выполнение шести лабораторных работ, посвященных построению и исследованию уравнения линейной регрессии вида:

$$\tilde{y} = b_0 + b_1 x. \tag{1.1}$$

Выборка для построения данного уравнения взята из следующего примера:

Пример 1.1. Для анализа зависимости объема потребления Y (у.е.) домохозяйства от располагаемого дохода X (у.е.) отобрана выборка объема $N = 12$ (помесячно в течение года), результаты которой приведены в таблице 1.1.

Таблица 1.1

i	1	2	3	4	5	6	7	8	9	10	11	12
X	107	109	110	113	120	122	123	128	136	140	145	150
Y	102	105	108	110	115	117	119	125	132	130	141	144

Лабораторная работа №1.1.
«Коэффициенты уравнения линейной парной регрессии»

Цель работы. Вычисление коэффициентов уравнения линейной парной регрессии по выборке из примера 1.1 (см. таблицу 1.1).

Расчетные соотношения. Для нахождения коэффициентов уравнения применим метод наименьших квадратов (МНК). Для этого необходимо минимизировать функцию:

$Z = \sum_{i=1}^{N}(y_i - \tilde{y}_i)^2 = \sum_{i=1}^{N}(y_i - b_0 - b_1 x_i)^2 \to \min$, где N – число наблюдений в выборке, $e_i = y_i - \tilde{y}_i$ – отклонения (остатки).

Необходимым условием существования минимума функции является равенство нулю частных производных данной функции по параметрам b_0 и b_1. Найдем эти частные производные:

$$\begin{cases} \dfrac{\partial Z}{\partial b_0} = -2\sum_{i=1}^{N}(y_i - b_0 - b_1 x_i) = 0; \\ \dfrac{\partial Z}{\partial b_1} = -2\sum_{i=1}^{N}(y_i - b_0 - b_1 x_i)\cdot x_i = 0. \end{cases}$$

Преобразуем полученную систему уравнений:

$$\begin{cases} \sum_{i=1}^{N} y_i - b_0\sum_{i=1}^{N} 1 - b_1\sum_{i=1}^{N} x_i = 0; \\ \sum_{i=1}^{N} y_i x_i - b_0\sum_{i=1}^{N} x_i - b_1\sum_{i=1}^{N} x_i x_i = 0. \end{cases}$$

$$\begin{cases} b_0 N + b_1\sum_{i=1}^{N} x_i = \sum_{i=1}^{N} y_i; \\ b_0\sum_{i=1}^{N} x_i + b_1\sum_{i=1}^{N} x_i^2 = \sum_{i=1}^{N} y_i x_i. \end{cases}$$

Разделив оба уравнения на N, получим систему нормальных уравнений:

$$\begin{cases} b_0 + b_1\bar{x} = \bar{y}; \\ b_0\bar{x} + b_1\overline{x^2} = \overline{xy}, \end{cases} \tag{1.2}$$

где $\bar{x} = \dfrac{1}{N}\sum_{i=1}^{N} x_i$, $\bar{y} = \dfrac{1}{N}\sum_{i=1}^{N} y_i$, $\overline{x^2} = \dfrac{1}{N}\sum_{i=1}^{N} x_i^2$, $\overline{xy} = \dfrac{1}{N}\sum_{i=1}^{N} x_i y_i$. (1.3)

Выразим из первого уравнения b_0:

$$b_0 = \bar{y} - b_1\bar{x}. \tag{1.4}$$

Подставив выражение для b_0 во второе уравнение, получим формулу для вычисления b_1:

$$b_1 = \frac{\overline{xy} - \bar{x}\cdot\bar{y}}{\overline{x^2} - (\bar{x})^2} = \frac{S_{xy}}{S_x^2}, \tag{1.5}$$

где S_{xy} – выборочная ковариация, определяемая по формуле:

$$S_{xy} = \overline{xy} - \bar{x} \cdot \bar{y}\,; \qquad (1.6)$$

S_x^2 – выборочная дисперсия величины X, определяемая по формуле:

$$S_x^2 = \overline{x^2} - (\bar{x})^2\,. \qquad (1.7)$$

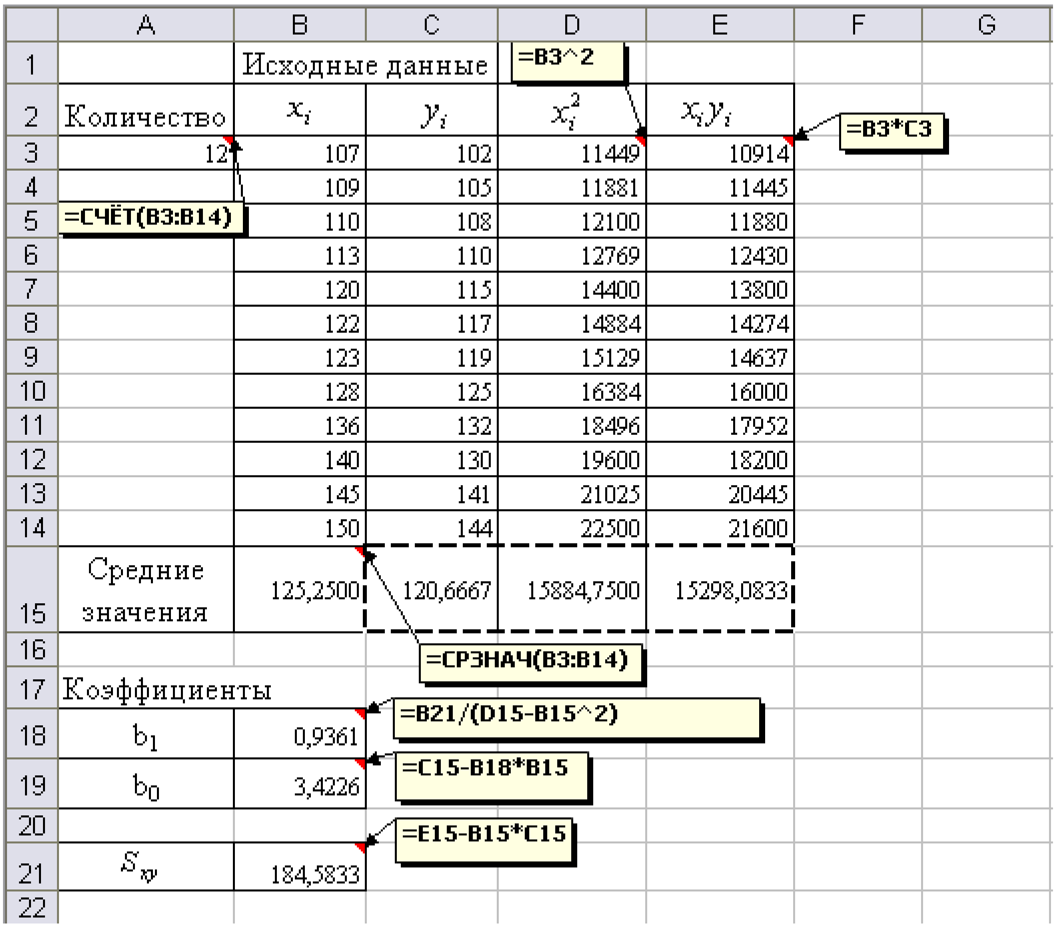
Решение. Вычислим коэффициенты b_0 и b_1, используя табличный процессор Excel. На рис. 1.1 показан фрагмент листа Excel, в котором:

– размещены данные таблицы 1.1;

– составлены формулы для вычисления величин x_i^2, $x_i y_i$;

– вычислены средние значения с помощью функции Excel СРЗНАЧ(*диапазон ячеек*);

– вычислена ковариация S_{xy} по формуле (1.6), коэффициенты b_0, b_1 по формулам (1.4), (1.5) соответственно.



	A	B	C	D	E	F	G
1		Исходные данные					
2	Количество	x_i	y_i	x_i^2	$x_i y_i$		
3	12	107	102	11449	10914		
4		109	105	11881	11445		
5		110	108	12100	11880		
6		113	110	12769	12430		
7		120	115	14400	13800		
8		122	117	14884	14274		
9		123	119	15129	14637		
10		128	125	16384	16000		
11		136	132	18496	17952		
12		140	130	19600	18200		
13		145	141	21025	20445		
14		150	144	22500	21600		
15	Средние значения	125,2500	120,6667	15884,7500	15298,0833		
16							
17	Коэффициенты						
18	b_1	0,9361					
19	b_0	3,4226					
20							
21	S_{xy}	184,5833					
22							

Рис. 1.1. Вычисление коэффициентов парной линейной регрессии

Замечание: Для расчета в таблице достаточно вычислять только значения в третьей строке, для остальных ячеек таблицы (по 14 строку) и для выделенных пунктиром ячеек применять автозаполнение.

В результате выполнения вычислений получаем:

$$S_{xy} = 15298{,}0833 - 125{,}25 \cdot 120{,}6667 = 184{,}5833\,,$$

$$b_1 = \frac{184,5833}{15884,75 - 125,25^2} = 0,9361,$$

$$b_0 = 120,6667 - 0,9361 \cdot 125,25 = 3,4226.$$

Тогда уравнение регрессии (1.1) примет вид:

$$\tilde{y} = 3,4226 + 0,9361x. \qquad (1.8)$$

Дадим интерпретацию оценкам коэффициентов регрессии.

Коэффициент $b_1 = 0,9361$ может трактоваться как предельная склонность к потреблению. Фактически он показывает, на какую величину изменится объем потребления, если располагаемый доход возрастает на одну единицу.

Свободный член b_0 уравнения регрессии формально означает значение *Y* при величине располагаемого дохода *X*, равной нулю (т.е. автономное потребление). Следует иметь в виду, что интерпретация оценки свободного члена модели может иметь или не иметь реального смысла в зависимости от конкретной задачи. В нашем случае значение $b_0 = 3,4226$ говорит о том, что при нулевом располагаемом доходе расходы на потребление составят в среднем 3,4226 у.е. Этот факт можно объяснить для отдельного домохозяйства (оно может тратить накопленные или одолженные средства), но для совокупности домохозяйств он теряет смысл.

Вопросы:

– Что такое выборочная и генеральная совокупность наблюдений случайных величин?
– Как найти систему нормальных уравнений?
– Приведите формулы расчета коэффициентов b_0 и b_1 линейного уравнения регрессии.

Лабораторная работа №1.2.

«Коэффициент корреляции и его значимость»

Цель работы. Вычисление выборочного коэффициента корреляции, проверка его значимости при уровне значимости $\alpha = 0,05$ по выборке из примера 1.1 (см. таблицу 1.1).

Расчетные соотношения. Уравнение (1.1) означает, что между случайными величинами *X* и *Y* имеется корреляционная зависимость, в данном случае линейная. Теснота линейной связи между величинами *X* и *Y* определяется с помощью коэффициента корреляции:

$$r_{xy} = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sqrt{\overline{x^2} - \left(\bar{x}\right)^2} \cdot \sqrt{\overline{y^2} - \left(\bar{y}\right)^2}} = \frac{S_{xy}}{S_x \cdot S_y}, \qquad (1.9)$$

где $S_x = \sqrt{\overline{x^2} - (\bar{x})^2}$, $S_y = \sqrt{\overline{y^2} - (\bar{y})^2}$, $\overline{y^2} = \frac{1}{N}\sum_{i=1}^{N} y_i^2$. (1.10)

S_x, S_y – стандартные отклонения случайных величин *X* и *Y* соответственно.

Выборочный коэффициент корреляции обладает следующими свойствами:

1. Если между случайными величинами X и Y существует положительная (прямая) линейная зависимость, то $r_{xy} > 0$, т.е. при увеличении x увеличивается и y; если отрицательная (обратная), то $r_{xy} < 0$, т.е. при увеличении x значение y уменьшается.
2. Выборочный коэффициент корреляции r_{xy} является безразмерной величиной.
3. $-1 \le r_{xy} \le 1$.
4. Если $r_{xy} = 0$, то между случайными величинами X и Y отсутствует линейная связь.
5. Чем ближе r_{xy} по модулю к 1, тем сильнее линейная связь между X и Y.

Замечание: Близкая к нулю величина коэффициента корреляции говорит об отсутствии линейной связи переменных, но не об отсутствии связи между ними вообще.

Близкое к 1 значение коэффициента корреляции свидетельствует о том, что данные наблюдений хорошо согласуются с представлением их в виде линейной регрессионной модели.

Для установления наличия значимой линейной связи между случайными величинами X и Y, следует проверить гипотезу о статистической значимости коэффициента корреляции при помощи t-критерия Стьюдента. В этом случае используется следующая гипотеза:

$H_0 : r_{xy} = 0;$

$H_1 : r_{xy} \ne 0.$

Для проверки H_0 рассчитывается t-статистика:

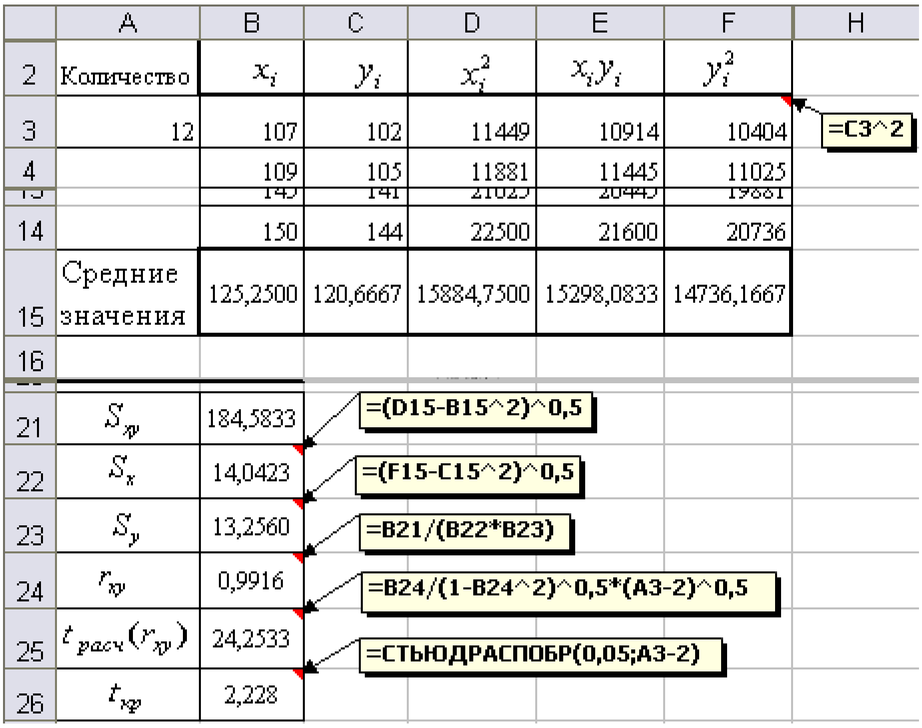
$$t_{расч}(r_{xy}) = \frac{r_{xy}}{\sqrt{1 - r_{xy}^{\ 2}}} \sqrt{N-2}\,. \qquad (1.11)$$

При справедливости H_0 t-статистика имеет распределение Стьюдента с $\nu = N - 2$ степенями свободы. По односторонней таблице распределения Стьюдента для заданного уровня значимости α и числа степеней свободы $N-2$ определяется критическое значение $t_{кр} = t\left(\frac{\alpha}{2}, N-2\right)$. Данное значение можно получить с помощью функции Excel (категория – *Статистические*):

$$t_{кр} = t\left(\frac{\alpha}{2}, N-2\right) = СТЬЮДРАСПОБР(\alpha; N-2)\,. \qquad (1.12)$$

Если выполняется неравенство $\left|t_{расч}\right| \le t_{кр}$, то нет оснований для отклонения гипотезы H_0, т.е. линейной связи между случайными величинами X и Y нет (коэффициент корреляции незначим). Если $\left|t_{расч}\right| > t_{кр}$, то нулевая гипотеза отклоняется в пользу альтернативной H_1, т.е. между случайными величинами X и Y есть линейная связь (коэффициент корреляции значим).

Решение. Фрагмент листа Excel с вычислениями приведен на рис. 1.2.



	A	B	C	D	E	F	H
2	Количество	x_i	y_i	x_i^2	$x_i y_i$	y_i^2	
3	12	107	102	11449	10914	10404	=C3^2
4		109	105	11881	11445	11025	
13		145	141	21025	20445	19881	
14		150	144	22500	21600	20736	
15	Средние значения	125,2500	120,6667	15884,7500	15298,0833	14736,1667	
16							
21	S_{xy}	184,5833	=(D15-B15^2)^0,5				
22	S_x	14,0423	=(F15-C15^2)^0,5				
23	S_y	13,2560	=B21/(B22*B23)				
24	r_{xy}	0,9916	=B24/(1-B24^2)^0,5*(A3-2)^0,5				
25	$t_{расч}(r_{xy})$	24,2533	=СТЬЮДРАСПОБР(0,05;A3-2)				
26	$t_{кр}$	2,228					

Рис. 1.2. Вычисление коэффициента корреляции
(*Замечание:* некоторые ячейки скрыты)

В результате выполнения вычислений получаем значение стандартных отклонений:

$$S_x = \sqrt{15884{,}75 - 125{,}25^2} = 14{,}0423,$$

$$S_y = \sqrt{14736{,}1667 - 120{,}6667^2} = 13{,}256;$$

значение коэффициента корреляции:

$$r_{xy} = \frac{184{,}5833}{14{,}0423 \cdot 13{,}256} = 0{,}9916.$$

Положительное значение коэффициента корреляции указывает на прямую зависимость между случайными величинами *X* и *Y*, т.е. при увеличении располагаемого дохода *X* увеличивается и объем потребления *Y*.

$$t_{расч}(r_{xy}) = \frac{0{,}9916}{\sqrt{1 - 0{,}9916^2}} \sqrt{12 - 2} = 24{,}2533;$$

$$t_{кр} = СТЬЮДРАСПОБР(0{,}05;\ 10) = 2{,}228.$$

Поскольку $|t_{расч}| > t_{кр}$, то гипотеза об отсутствии линейной зависимости отвергается, т.е. коэффициент корреляции статистически значим и между располагаемым доходом и объемом потребления есть линейная связь.

Вопросы:

– Для чего используется коэффициент корреляции?

– Какие значения может принимать коэффициент корреляции и что это означает?
– Как проверить статистическую значимость коэффициента корреляции?

Лабораторная работа №1.3.
«Значимость коэффициентов уравнения и их доверительные интервалы»

Цель работы. Определение статистической значимости коэффициентов уравнения и доверительных интервалов с помощью стандартных ошибок коэффициентов по выборке из примера 1.1 (см. таблицу 1.1).

Расчетные соотношения. Оценка значимости коэффициентов регрессии проводится с целью установления несущественных факторов: фактор, коэффициент при котором в уравнении линейной регрессии статистически незначим, оказывает несущественное влияние и должен быть исключен из модели.

Для оценки значимости коэффициента регрессии его величину сравнивают с его стандартной ошибкой S_{b_j}, т.е. определяют расчетное значение t-критерия Стьюдента:

$$t_{расч}(b_j) = \frac{b_j}{S_{b_j}}, \tag{1.13}$$

где $j = 0,...,m$, m – число объясняющих переменных.

Полученное значение сравнивают с табличным (критическим) и делают вывод о значимости или незначимости коэффициента регрессии аналогично выводу о значимости коэффициента корреляции.

Стандартные ошибки коэффициентов регрессии:

$$S_{b_1} = \sqrt{S_{b_1}^2}, \quad S_{b_0} = \sqrt{S_{b_0}^2}, \tag{1.14}$$

где $S_{b_1}^2$, $S_{b_0}^2$ – дисперсии коэффициентов b_1 и b_0 соответственно, которые вычисляются по формулам:

$$S_{b_1}^2 = \frac{S^2}{\sum_{i=1}^{N}(x_i - \bar{x})^2}, \quad S_{b_0}^2 = \frac{S^2 \cdot \overline{x^2}}{\sum_{i=1}^{N}(x_i - \bar{x})^2} = S_{b_1}^2 \cdot \overline{x^2}, \tag{1.15}$$

где S^2 – остаточная дисперсия, характеризующая степень рассеяния фактических значений y относительно расчетных значений $\tilde{y}$:

$$S^2 = \frac{\sum_{i=1}^{N}(y_i - \tilde{y}_i)^2}{N - m - 1}. \tag{1.16}$$

Отметим, что корень из необъясненной дисперсии S называется *стандартной ошибкой регрессии*:

$$S = \sqrt{\frac{\sum_{i=1}^{N}(y_i - \tilde{y}_i)^2}{N - m - 1}}. \tag{1.17}$$

Чтобы найти интервал, который накрывает истинное значение коэффициента с вероятностью $100(1-\alpha)\%$, нужно определить $100(1-\alpha)\%$-ный доверительный интервал:

$$\left(b_j - t_{кр} S_{b_j};\; b_j + t_{кр} S_{b_j}\right),\; j = 0, \ldots, m \tag{1.18}$$

Критическое значение $t_{кр}$:

$$t_{кр} = t\left(\frac{\alpha}{2}, N-m-1\right) = СТЬЮДРАСПОБР(\alpha;\, N-m-1)\,. \tag{1.19}$$

Решение. Фрагменты с вычислениями приведены на рис. 1.3 – 1.4.

	A	B	C	D	E	F	G	H	J	K
1	Исходные данные									
2	Количество	x_i	y_i	x_i^2	$x_i y_i$	y_i^2	$\tilde{y}_i$	$x_i - \bar{x}$	$e_i = y_i - \tilde{y}_i$	
3	12	107	102	11449	10914	10404	103,5832	-18,2500	-1,5832	
4		109	105	11881	11445	11025	105,4554	-16,2500	-0,4554	
5		110	108	12100	11880	11664	106,3914	-15,2500	1,6086	
6		113	110	12769	12430	12100	109,1997	-12,2500	0,8003	
7		120	115	14400	13800	13225	115,7522	-5,2500	-0,7522	
8		122	117	14884	14274	13689	117,6244	-3,2500	-0,6244	
9		123	119	15129	14637	14161	118,5605	-2,2500	0,4395	
10		128	125	16384	16000	15625	123,2409	2,7500	1,7591	
11		136	132	18496	17952	17424	130,7295	10,7500	1,2705	
12		140	130	19600	18200	16900	134,4739	14,7500	[illegible]	
13		145	141	21025	20445	19881	139,1543	19,7500	[illegible]	
14		150	144	22500	21600	20736	143,8347	24,7500	0,1653	
15	Средние значения	125,2500	120,6667	15884,7500	15298,0833	14736,1667	Сумма квадратов	2366,2500	35,2488	
16										
17	Коэффициенты			S^2	3,5249		S	1,8775		
18	b_1	0,9361		$S_{b_1}^2$	0,0015		S_{b_1}	0,0386	$t_{расч}(b_1)$	24,2533
19	b_0	3,4226		$S_{b_0}^2$	23,6627		S_{b_0}	4,8644	$t_{расч}(b_0)$	0,7036

=\$B\$19+\$B\$18*B3

=B3-\$B\$15

=C3-G3

=СУММКВ(H3:H14)

=J15/(A3-2)

=КОРЕНЬ(E17)

=B18/H18

=E17/H15

=E18*D15

Рис. 1.3. Значимость коэффициентов уравнения и доверительные интервалы (*Замечание:* некоторые ячейки скрыты)

В результате выполнения вычислений получим:

Стандартная ошибка регрессии $S = \sqrt{\dfrac{35{,}2488}{12-2}} = \sqrt{3{,}5249} = 1{,}8775$.

Стандартные ошибки коэффициентов регрессии:

$$S_{b_1} = \sqrt{\frac{3{,}5249}{2366{,}25}} = \sqrt{0{,}0015} = 0{,}0386\,,\quad S_{b_0} = 0{,}0386 \cdot \sqrt{15884{,}75} = \sqrt{23{,}6627} = 4{,}8644\,.$$

Тогда расчетные значения t-критерия Стьюдента:

$$t_{расч}(b_1) = \frac{b_1}{S_{b_1}} = \frac{0{,}9361}{0{,}0386} = 24{,}2533\,,$$

$t_{расч}(b_0) = \dfrac{b_0}{S_{b_0}} = \dfrac{3{,}4226}{4{,}8644} = 0{,}7036$;

$t_{кр} = t(0{,}025;12-1-1) = СТЬЮДРАСПОБР(0{,}05;10) = 2{,}228$.

Поскольку для коэффициента b_1 $|t_{расч}(b_1)| = 24{,}2533 > 2{,}228 = t_{кр}$, то нулевая гипотеза (о равенстве нулю коэффициента b_1) должна быть отвергнута при уровне значимости $\alpha = 0{,}05$. Это подтверждает статистическую значимость коэффициента b_1.

Аналогично проверяется статистическая значимость коэффициента b_0: так как $|t_{расч}(b_0)| = 0{,}7036 < 2{,}228 = t_{кр}$, то гипотеза о статистической незначимости коэффициента b_0 не отклоняется при уровне значимости $\alpha = 0{,}05$. Это означает, что в данном случае свободным членом уравнения регрессии можно пренебречь, рассматривая регрессию как $\tilde{y} = b_1 x$, однако наличие свободного члена может лишь уточнить вид зависимости, поэтому, если нет серьезных причин для удаления свободного члена из уравнения регрессии, то лучше его использовать в модели.

95%-е доверительные интервалы (см. рис 1.4.):

для b_1: $(0{,}9361 - 2{,}228 \cdot 0{,}0386;\ 0{,}9361 + 2{,}228 \cdot 0{,}0386)$ или $(0{,}8501;\ 1{,}0221)$;

для b_0: $(3{,}4226 - 2{,}228 \cdot 4{,}8644;\ 3{,}4226 + 2{,}228 \cdot 4{,}8644)$ или $(-7{,}416;\ 14{,}2612)$.

Таким образом, истинное значение коэффициента b_1 с вероятностью 95% находится в интервале $(0{,}8501;\ 1{,}0221)$.

Доверительный интервал для b_0 содержит противоречивые результаты, так как найденный интервал $(-7{,}416;\ 14{,}2612)$ показывает, что истинное значение коэффициента b_0 одновременно содержит положительные, отрицательные величины и даже нуль, чего не может быть. Данная интервальная оценка также подтверждает незначимость коэффициента b_0.

Вопросы:

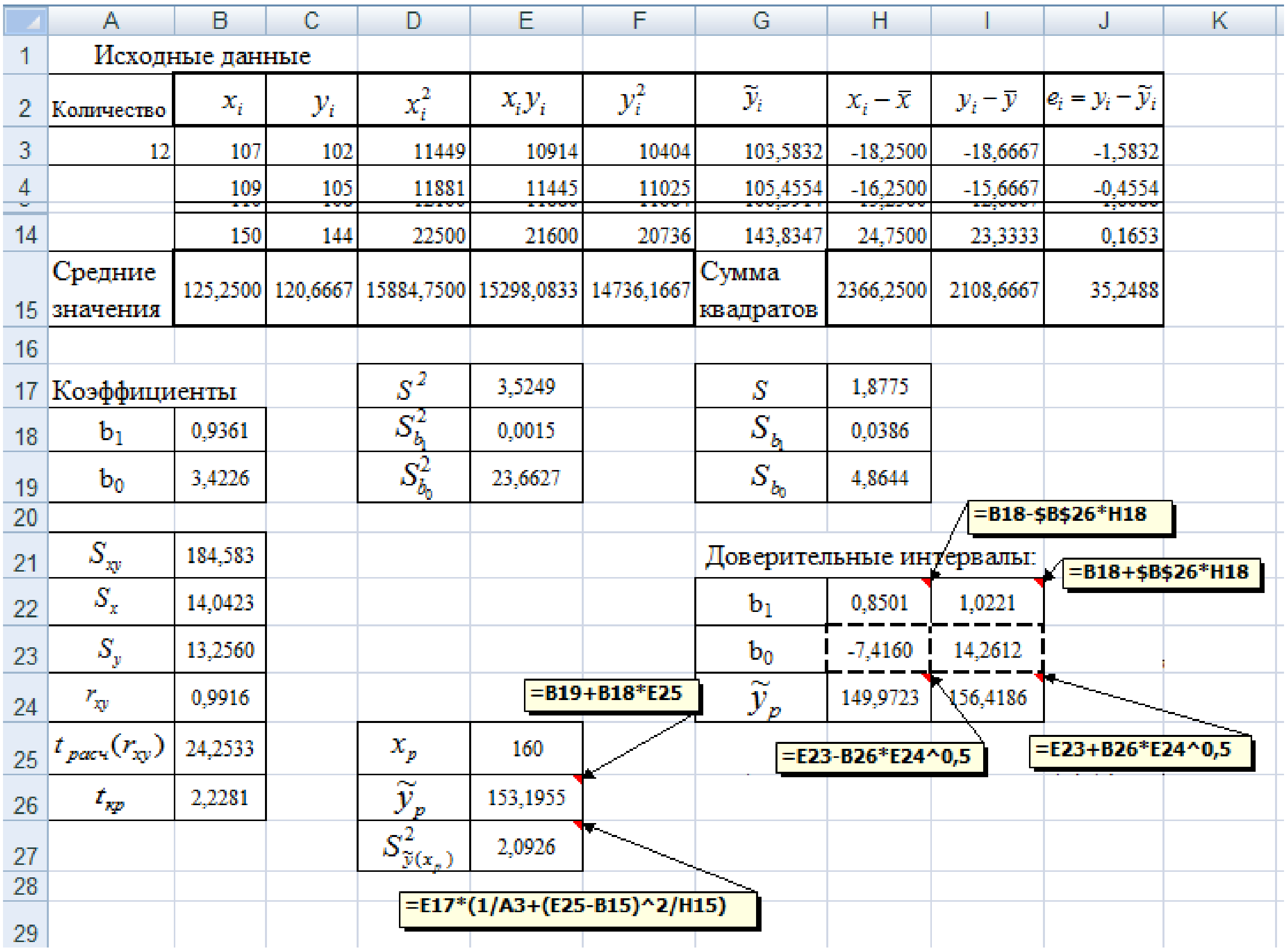
- Что показывает остаточная дисперсия и как она рассчитывается?
- Как определить стандартные ошибки коэффициентов регрессии?
- Для чего определяется статистическая значимость коэффициентов регрессии?
- Как определить интервальные оценки коэффициентов регрессии?

Лабораторная работа №1.4.

«Точечный и интервальный прогнозы»

Цель работы. Используя уравнение (1.8), спрогнозировать потребление, если располагаемый доход равен $x_p = 160$ у. е. Построить интервальную оценку для данного значения дохода с надежностью $\gamma = 0{,}95$.

Расчетные соотношения. Прогноз на основе построенной модели может быть точечным и интервальным. Точечный прогноз $\tilde{y}_p$ определяется путем подстановки соответствующего значения x_p в уравнение регрессии $\tilde{y}_p = b_0 + b_1 x_p$.



	A	B	C	D	E	F	G	H	I	J
1	Исходные данные									
2	Количество	x_i	y_i	x_i^2	$x_i y_i$	y_i^2	$\tilde{y}_i$	$x_i - \bar{x}$	$y_i - \bar{y}$	$e_i = y_i - \tilde{y}_i$
3	12	107	102	11449	10914	10404	103,5832	-18,2500	-18,6667	-1,5832
4		109	105	11881	11445	11025	105,4554	-16,2500	-15,6667	-0,4554
14		150	144	22500	21600	20736	143,8347	24,7500	23,3333	0,1653
15	Средние значения	125,2500	120,6667	15884,7500	15298,0833	14736,1667	Сумма квадратов	2366,2500	2108,6667	35,2488
16										
17	Коэффициенты			S^2	3,5249		S	1,8775		
18	b_1	0,9361		$S^2_{b_1}$	0,0015		S_{b_1}	0,0386		
19	b_0	3,4226		$S^2_{b_0}$	23,6627		S_{b_0}	4,8644		
20										
21	S_{xy}	184,583					Доверительные интервалы:			
22	S_x	14,0423					b_1	0,8501	1,0221	
23	S_y	13,2560					b_0	-7,4160	14,2612	
24	r_{xy}	0,9916					$\tilde{y}_p$	149,9723	156,4186	
25	$t_{расч}(r_{xy})$	24,2533		x_p	160					
26	$t_{кр}$	2,2281		$\tilde{y}_p$	153,1955					
27				$S^2_{\tilde{y}(x_p)}$	2,0926					
28										
29										

Рис. 1.4. Построение интервальной оценки
(*Замечание:* некоторые ячейки скрыты)

Интервальная оценка (доверительный интервал для зависимой переменной) с надежностью γ определяется выражением:

$$\left(\tilde{y}_p - t_{кр} S_{\tilde{y}(x_p)};\ \tilde{y}_p + t_{кр} S_{\tilde{y}(x_p)}\right), \tag{1.20}$$

где $S_{\tilde{y}(x_p)}$ – стандартная ошибка прогнозируемого по линии регрессии значения.

Оценка $S^2_{\tilde{y}(x_p)}$ для дисперсии функции $\tilde{y}(x)$:

$$S^2_{\tilde{y}(x_p)} = S^2 \left[\frac{1}{N} + \frac{(x_p - \bar{x})^2}{\sum_{i=1}^{N} (x_i - \bar{x})^2} \right], \tag{1.21}$$

где $S^2 = \dfrac{\sum_{i=1}^{N}(y_i - \tilde{y}_i)^2}{N-m-1}$.

Критическое значение $t_{кр}$ определяется как для коэффициентов уравнения с учетом того, что $\gamma = 1-\alpha$:

$$t_{кр} = t\left(\frac{1-\gamma}{2}; N-m-1\right) = t\left(\frac{\alpha}{2}; N-m-1\right) = СТЬЮДРАСПОБР(\alpha; N-m-1).$$

Решение. Фрагмент листа Excel с вычислениями приведен на рис. 1.4.

В результате выполнения вычислений получим:

$$\tilde{y}_p = b_0 + b_1 x_p = 3{,}4226 + 0{,}9361 \cdot 160 = 153{,}1955;$$

$$S^2_{\tilde{y}(x_p)} = 3{,}5249 \cdot \left[\frac{1}{12} + \frac{(160-125{,}25)^2}{2366{,}25}\right] = 2{,}0926;$$

$$t_{кр} = t\left(\frac{1-0{,}95}{2}; 12-2\right) = t\left(\frac{0{,}05}{2}; 10\right) = СТЬЮДРАСПОБР(0{,}05; 10) = 2{,}228.$$

Тогда значения нижней $y_p^н$ и верхней $y_p^в$ границ интервала (1.20):

$$\left(153{,}1955 - 2{,}228 \cdot \sqrt{2{,}0926};\ 153{,}1955 + 2{,}228 \cdot \sqrt{2{,}0926}\right) \text{ или } (149{,}9723;\ 156{,}4186).$$

Таким образом, среднее потребление при доходе 160 у.е. с вероятностью 95% будет находиться в интервале $(149{,}9723;\ 156{,}4186)$.

Вопросы:

– Как использовать уравнение регрессии для прогнозирования в эконометрической модели?
– Для чего определяется доверительный интервал для зависимой переменной?

Лабораторная работа №1.5.
«Определение автокорреляции остатков»

Цель работы. Определение автокорреляции остатков с помощью критерия Дарбина-Уотсона по выборке из примера 1.1 (см. таблицу 1.1).

Расчетные соотношения. Автокорреляция – это коррелированность (зависимость) между любыми отклонениями e_i, в частности, между соседними. Автокорреляция остатков может быть вызвана несколькими причинами:

– наличие ошибок измерения в значениях исходных данных;
– невключение в модель фактора, оказывающего существенное влияние на результат, влияние которого отражается в остатках, вследствие чего последние могут оказаться автокоррелированными;
– неправильная спецификация функциональной формы модели (например, должна быть линейная, а взята гиперболическая или другая модель).

Наиболее распространенный метод определения автокорреляции – критерий Дарбина-Уотсона, который заключается в вычислении статистики:

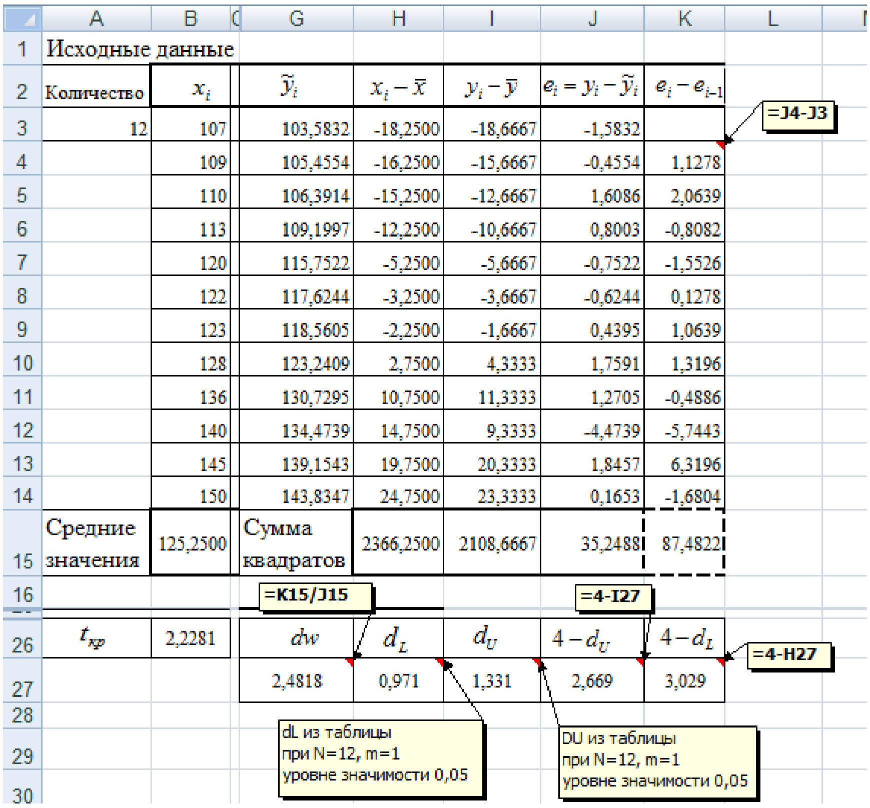
$$dw = \frac{\sum_{i=2}^{N}(e_i - e_{i-1})^2}{\sum_{i=1}^{N} e_i^{\,2}}. \qquad (1.22)$$

Выдвигается гипотеза H_0 об отсутствии автокорреляции остатков. Альтернативные гипотезы H_1 и H_1^* состоят, соответственно, в наличии положительной или отрицательной автокорреляции. Далее по специальным таблицам (см. приложение 4) определяются критические значения критерия Дарбина-Уотсона d_L и d_U для заданного числа наблюдений N, числа независимых переменных m и уровня значимости α.

Есть положительная автокорреляция. H_0 отклоняется. С вероятностью $p=1-\alpha$ принимается H_1	Область неопределенности	Нет оснований отклонять H_0 (автокорреляция остатков отсутствует)	Область неопределенности	Есть отрицательная автокорреляция. H_0 отклоняется. С вероятностью $p=1-\alpha$ принимается H_1^*

0 $\quad d_L \quad d_U \quad 2 \quad 4-d_U \quad 4-d_L \quad 4$

Если фактическое значение критерия Дарбина-Уотсона попадает в область неопределенности, то на практике предполагают наличие автокорреляции остатков и отклоняют гипотезу H_0.



	A	B	G	H	I	J	K
1	Исходные данные						
2	Количество	x_i	$\tilde{y}_i$	$x_i-\bar{x}$	$y_i-\bar{y}$	$e_i=y_i-\tilde{y}_i$	e_i-e_{i-1}
3	12	107	103,5832	-18,2500	-18,6667	-1,5832	
4		109	105,4554	-16,2500	-15,6667	-0,4554	1,1278
5		110	106,3914	-15,2500	-12,6667	1,6086	2,0639
6		113	109,1997	-12,2500	-10,6667	0,8003	-0,8082
7		120	115,7522	-5,2500	-5,6667	-0,7522	-1,5526
8		122	117,6244	-3,2500	-3,6667	-0,6244	0,1278
9		123	118,5605	-2,2500	-1,6667	0,4395	1,0639
10		128	123,2409	2,7500	4,3333	1,7591	1,3196
11		136	130,7295	10,7500	11,3333	1,2705	-0,4886
12		140	134,4739	14,7500	9,3333	-4,4739	-5,7443
13		145	139,1543	19,7500	20,3333	1,8457	6,3196
14		150	143,8347	24,7500	23,3333	0,1653	-1,6804
15	Средние значения	125,2500	Сумма квадратов	2366,2500	2108,6667	35,2488	87,4822
16							
26	$t_{кр}$	2,2281	dw	d_L	d_U	$4-d_U$	$4-d_L$
27			2,4818	0,971	1,331	2,669	3,029
28							
29							
30							

Рис. 1.5. Определение автокорреляции остатков
(*Замечание:* некоторые ячейки скрыты)

При наличии автокорреляции остатков полученное уравнение регрессии обычно считается неудовлетворительным.

Решение. На рис. 1.5 показан фрагмент листа Excel, в котором:

- составлены формулы для вычисления величин e_i , $e_i - e_{i-1}$;
- вычислены суммы квадратов с помощью функции Excel СУММКВ(*диапазон ячеек*);
- вычислена статистика dw по формуле (1.22).

В результате выполнения вычислений получим:

$$dw = \frac{\sum_{i=2}^{N}(e_i - e_{i-1})^2}{\sum_{i=1}^{N} e_i{}^2} = \frac{87{,}4822}{35{,}2488} = 2{,}4818.$$

По таблицам распределения Дарбина-Уотсона для заданного числа наблюдений $N = 12$, числа объясняющих переменных *m=1* и уровня значимости $\alpha = 0{,}05$ определяем критические значения $d_L = 0{,}971$ и $d_U = 1{,}331$. Определим величины $4 - d_U = 4 - 1{,}331 = 2{,}669$, $4 - d_L = 4 - 0{,}971 = 3{,}029$.

Поскольку статистика Дарбина-Уотсона $dw = 2{,}4818$ попадает в интервал $d_U = 1{,}331 < dw < 2{,}669 = 4 - d_U$, то нет оснований отклонять H_0, т.е. автокорреляция остатков отсутствует, что является одним из важных параметров для построения качественной модели.

Вопросы:

- Чем может быть вызвана автокорреляция остатков?
- Как определить наличие или отсутствие автокорреляции остатков?
- Для чего применяется исследование автокорреляции остатков?

Лабораторная работа №1.6.

«Общее качество уравнения регрессии»

Цель работы. Для уравнения регрессии, построенного по выборке из примера 1.1, для уровня значимости $\alpha = 0{,}05$ оценить значимость коэффициента детерминации, проверить прогнозные качества модели, сделать вывод об общем качестве модели.

Расчетные соотношения. Для проверки общего качества уравнения используется коэффициент детерминации:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \tilde{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2}. \tag{1.23}$$

В общем случае справедливо соотношение $0 \le R^2 \le 1$ (R^2 может оказаться отрицательным для уравнений, в которых отсутствует свободный член). Чем ближе R^2 к единице, тем теснее линейная связь между X и Y. Чем слабее такая связь, тем R^2 ближе к нулю. Коэффициент детерминации рассматривается лишь как один из ряда показателей, который нужно проанализировать, чтобы уточнить строящуюся модель, так как могут быть случаи неправильно специфицированных моделей, имеющих высокий коэффициент детерминации.

Для определения значимости коэффициента детерминации проверяют гипотезу:

$H_0 : R^2 = 0;$

$H_1 : R^2 > 0.$

Для проверки H_0 применяют критерий Фишера. Для этого рассчитывается F-статистика:

$$F_{расч} = \frac{R^2}{1-R^2} \cdot \frac{N-m-1}{m}. \tag{1.24}$$

При справедливости H_0 F-статистика имеет распределение Фишера с числами степеней свободы $k_r = m$ и $k_e = N - m - 1$. По таблицам распределения Фишера для заданного уровня значимости α определяется критическое значение $F_{кр} = F(\alpha; m; N-m-1)$. Данное значение можно получить с помощью функции Excel (категория – *Статистические*):

$$F_{кр} = F(\alpha; m; N-m-1) = FРАСПОБР(\alpha; m; N-m-1). \tag{1.25}$$

Если $F_{расч} > F_{кр}$, то нулевая гипотеза отклоняется, следовательно, уравнение регрессии достаточно качественно отражает динамику изменения зависимой переменной. Это равносильно тому, что коэффициент детерминации значим. Если $F_{расч} < F_{кр}$, то нет оснований для отклонения H_0 и общее качество модели невысоко (коэффициент детерминации незначим).

Прогнозные качества модели проверяют при помощи средней относительной ошибки аппроксимации:

$$\overline{E}_{отн} = \frac{100}{N} \sum_{i=1}^{N} \left| \frac{y_i - \tilde{y}_i}{y_i} \right|. \tag{1.26}$$

Ошибка аппроксимации не более 10% свидетельствует о хорошем подборе модели к исходным данным.

При определении общего качества модели обычно анализируются следующие параметры:

- t-статистики;
- коэффициент детерминации;
- статистика Дарбина-Уотсона;
- прогнозные качества модели.

Решение. Фрагмент листа Excel с вычислениями приведен на рис. 1.6.

В результате выполнения вычислений получим:

$$R^2 = 1 - \frac{35{,}2488}{2108{,}6667} = 0{,}98328,$$

$$F_{расч} = \frac{0{,}98328}{1-0{,}98328} \cdot \frac{12-1-1}{1} = 588{,}2235,$$

$$F_{кр} = F(0{,}05; 1; 12-1-1) = FРАСПОБР(0{,}05; 1; 10) = 4{,}9646.$$

Поскольку $F_{расч} = 588{,}2235 > 4{,}9646 = F_{кр}$, то коэффициент детерминации статистически значим.

	A	B	C	D	E	F	G	H	I	J	K	L
1	Исходные данные											
2	Количество	x_i	y_i	x_i^2	$x_i y_i$	y_i^2	$\tilde{y}_i$	$x_i-\bar{x}$	$y_i-\bar{y}$	$e_i = y_i - \tilde{y}_i$	$e_i - e_{i-1}$	$\left\lvert\frac{y_i-\tilde{y}_i}{y_i}\right\rvert$
3	12	107	102	11449	10914	10404	103,5832	-18,2500	-18,6667	-1,5832		0,01552
4		109	105	11881	11445	11025	105,4554	-16,2500	-15,6667	-0,4554	1,1278	0,00434
5		110	108	12100	11880	11664	106,3914	-15,2500	-12,6667	1,6086	2,0639	0,01489
6		113	110	12769	12430	12100	109,1997	-12,2500	-10,6667	0,8003	-0,8082	0,00728
7		120	115	14400	13800	13225	115,7522	-5,2500	-5,6667	-0,7522	-1,5526	0,00654
8		122	117	14884	14274	13689	117,6244	-3,2500	-3,6667	-0,6244	0,1278	0,00534
9		123	119	15129	14637	14161	118,5605	-2,2500	-1,6667	0,4395	1,0639	0,00369
10		128	125	16384	16000	15625	123,2409	2,7500	4,3333	1,7591	1,3196	0,01407
11		136	132	18496	17952	17424	130,7295	10,7500	11,3333	1,2705	-0,4886	0,00962
12		140	130	19600	18200	16900	134,4739	14,7500	9,3333	-4,4739	-5,7443	0,03441
13		145	141	21025	20445	19881	139,1543	19,7500	20,3333	1,8457	6,3196	0,01309
14		150	144	22500	21600	20736	143,8347	24,7500	23,3333	0,1653	-1,6804	0,00115
15	Средние значения	125,2500	120,6667	15884,7500	15298,0833	14736,1667	Сумма квадратов	2366,2500	2108,6667	35,2488	87,4822	
16											Сумма	0,12995
20												
21	S_{xy}	184,583		R^2	0,98328		Доверительные интервалы:				Ошибка аппроксимации:	
22	S_x	14,0423		Fрасч	588,2235		[illegible]	[illegible]501	1,0221			
23	S_y	13,2560		Fкр	4,9646			[illegible]	14,2612		$\bar{E}_{отн}$	1,0829

=ABS(J3/C3)

=C3-C15

=СУММ(L3:L14)

=1-J15/I15

=E21/(1-E21)*(A3-2)/1

=ФРАСПОБР(0,05; 1; A3-2)

=100/A3*L16

Рис. 1.6. Оценка общего качества и прогнозных качеств уравнения (*Замечание:* некоторые ячейки скрыты)

По коэффициенту детерминации для линейной зависимости можно сказать, что 98,328% изменения среднего объема потребления домохозяйства обусловлено зависимостью от дохода, остальные 1,672% приходятся на долю прочих факторов, не учтенных уравнением линейной регрессии.

Средняя относительная ошибка аппроксимации:

$$\bar{E}_{отн} = \frac{100}{12} \cdot 0{,}12995 = 1{,}0829\%.$$

Ошибка аппроксимации не более 10% свидетельствует о хороших прогнозных качествах модели.

Вывод: По критерию Стьюдента коэффициент b_1 статистически значим, а коэффициент b_0 статистически незначим, но поскольку для парной регрессии более важным является анализ статистической значимости коэффициента b_1, так как именно в нем скрыто влияние объясняющей переменной X на Y, то незначимость коэффициента b_0 на общее качество модели не оказывает особого влияния. Коэффициент корреляции и коэффициент детерминации статистически значимы, автокорреляция остатков отсутствует. Ошибка аппроксимации менее 10% подтверждает высокие прогнозные качества модели. Данные характеристики говорят о высоком общем качестве уравнения регрессии. Таким образом, линейная модель очень достоверно отражает зависимость объема потребления от располагаемого дохода.

Вопросы:

- Какая связь между коэффициентом детерминации и коэффициентом корреляции для парной регрессии?
- Какие значения может принимать коэффициент детерминации?
- Для чего рассчитывают среднюю относительную ошибку аппроксимации?
- По каким параметрам обычно оценивают общее качество уравнения регрессии?

Тема 2. НЕЛИНЕЙНАЯ ПАРНАЯ РЕГРЕССИЯ

При аппроксимировании данных нелинейными моделями следует иметь в виду нелинейность по переменным и нелинейность по параметрам модели. Так, модель гиперболическая $\tilde{y} = b_0 + \frac{b_1}{x}$ линейна по параметрам b_0 и b_1 и нелинейна по *x*; для построения и анализа следует использовать алгоритм, аналогичный построению линейной модели с заменой x_i на $\frac{1}{x_i}$.

Нелинейные по параметрам модели можно с помощью логарифмирования или иных преобразований свести к линейным.

Данная тема включает выполнение двух лабораторных работ, посвященных построению уравнения нелинейной (в данном случае степенной) парной регрессии вида:

$$\tilde{y} = b_0 x^{b_1}. \qquad (2.1)$$

Выборка для построения этого уравнения взята из примера 1.1.

Лабораторная работа №2.1.
«Нелинейная регрессия»

Цель работы. По данным примера 1.1 построить уравнение регрессии для степенной зависимости $\tilde{y} = b_0 x^{b_1}$ (см. таблицу 1.1).

Расчетные соотношения. Преобразуем уравнение $\tilde{y} = b_0 x^{b_1}$ в линейное путем логарифмирования: $\ln y = \ln b_0 + b_1 \ln x$. Если обозначить $y' = \ln y$, $x' = \ln x$, $b_0' = \ln b_0$, то будем иметь линейную по параметрам модель $y' = b_0' + b_1 x'$.

Для определения оценок b_0' и b_1 используем формулы, аналогичные формулам (1.4), (1.5), для линейной регрессии:

$$b_1 = \frac{\overline{x'y'} - \overline{x'} \cdot \overline{y'}}{\overline{x'^2} - \left(\overline{x'}\right)^2}, \quad b_0' = \overline{y'} - b_1 \overline{x'}. \tag{2.2}$$

Решение. Фрагмент листа Excel с вычислениями приведен на рис. 2.1.

	A	B	C	D	E	F	G	H
1		Исходные данные		=ln(B3)		=D3^2		=D3*E3
2	Количество	x_i	y_i	x_i'	y_i'	$x_i'^2$	$x_i'y_i'$	
3	12	107	102	4,6728	4,6250	21,8353	21,6117	
4		109	105	4,6913	4,6540	22,0087	21,8333	
5		110	108	4,7005	4,6821	22,0945	22,0083	
6		113	110	4,7274	4,7005	22,3482	22,2210	
7		120	115	4,7875	4,7449	22,9201	22,7163	
8		122	117	4,8040	4,7622	23,0786	22,8776	
9		123	119	4,8122	4,7791	23,1571	22,9980	
10		128	125	4,8520	4,8283	23,5422	23,4271	
11		136	132	4,9127	4,8828	24,1342	23,9875	
12		140	130	4,9416	4,8675	24,4198	24,0536	
13		145	141	4,9767	4,9488	24,7679	24,6287	
14		150	144	5,0106	4,9698	25,1065	24,9019	
15	Средние значения	125,2500	120,6667	4,8241	4,7871	23,2844	23,1054	
16			=(G15-D15*E15)/(F15-D15^2)					
17	b_1	0,9727						
18	b_0'	0,0946	=E15-B17*D15					
19	b_0	1,0992	=EXP(B18)					
20								

Рис. 2.1. Коэффициенты для степенной зависимости

В результате выполнения вычислений получим:

$$b_1 = \frac{23{,}1054 - 4{,}8241 \cdot 4{,}7871}{23{,}2844 - (4{,}8241)^2} = 0{,}9727;$$

$$b_0' = 4{,}7871 - 0{,}9727 \cdot 4{,}8241 = 0{,}0946.$$

Так как $b_0' = \ln b_0$, то оценкой коэффициента b_0 будет $b_0 = e^{b_0'}$ или, используя функцию Excel (категория – *Математические*):

$$b_0 = EXP(b_0') = 1{,}0992.$$

Найденные оценки коэффициентов позволяют записать искомое уравнение регрессии в виде $\tilde{y} = 1{,}0992 \cdot x^{0{,}9727}$.

Вопросы:

– Как рассчитать коэффициенты нелинейной модели?

– В каком случае нужно применять нелинейные модели?

Лабораторная работа №2.2.

«Исследование нелинейной регрессии»

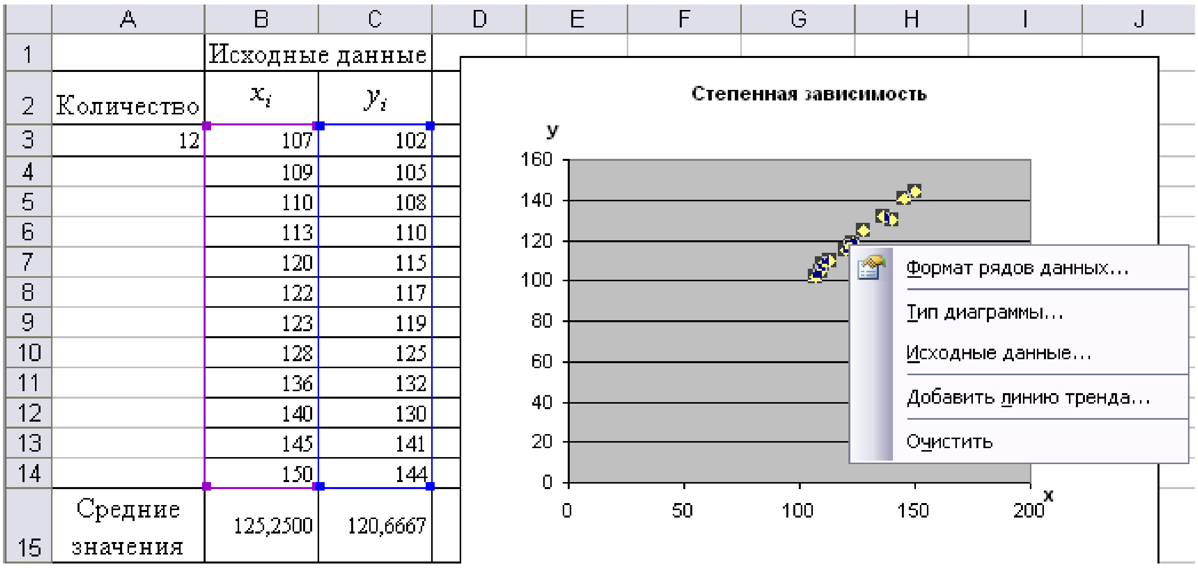
Цель работы. По данным примера 1.1 построить график степенной зависимости, рассчитать индекс корреляции, сравнить с линейной зависимостью коэффициенты детерминации.

Команда «Добавить линию тренда». Эта команда позволяет построить следующие уравнения регрессии:

- линейную $\tilde{y} = b_0 + b_1 x$
- полиноминальную $\tilde{y} = b_0 + b_1 x + ... + b_k x^k$ $(k \le 6)$;
- логарифмическую $\tilde{y} = b_0 + b_1 \ln x$
- степенную $\tilde{y} = b_0 x^{b_1}$;
- экспоненциальную $\tilde{y} = b_0 e^{b_1 x}$.

Для построения одной из перечисленных регрессий необходимо выполнить следующие шаги:

1. Выделить исходные данные $\{x_i, y_i\}$, $i = 1,2,..., N$, введенные в столбец (рис. 2.2).
2. По этим данным построить диаграмму (например, точечную). Для этого в меню Вставка выбрать пункт Диаграмма, далее – Точечная и т.д.
3. Установить курсор на построенном графике, сделать щелчок правой кнопкой и в появившемся контекстном меню выполнить команду *Добавить линию тренда* (см. рис. 2.2).



	A	B	C
1		Исходные данные	
2	Количество	x_i	y_i
3	12	107	102
4		109	105
5		110	108
6		113	110
7		120	115
8		122	117
9		123	119
10		128	125
11		136	132
12		140	130
13		145	141
14		150	144
15	Средние значения	125,2500	120,6667

Рис. 2.2. Построение графика по исходным данным

4. В появившемся диалоговом окне (см. рис. 2.3) на вкладке *«Тип»* выбрать нужное уравнение регрессии.

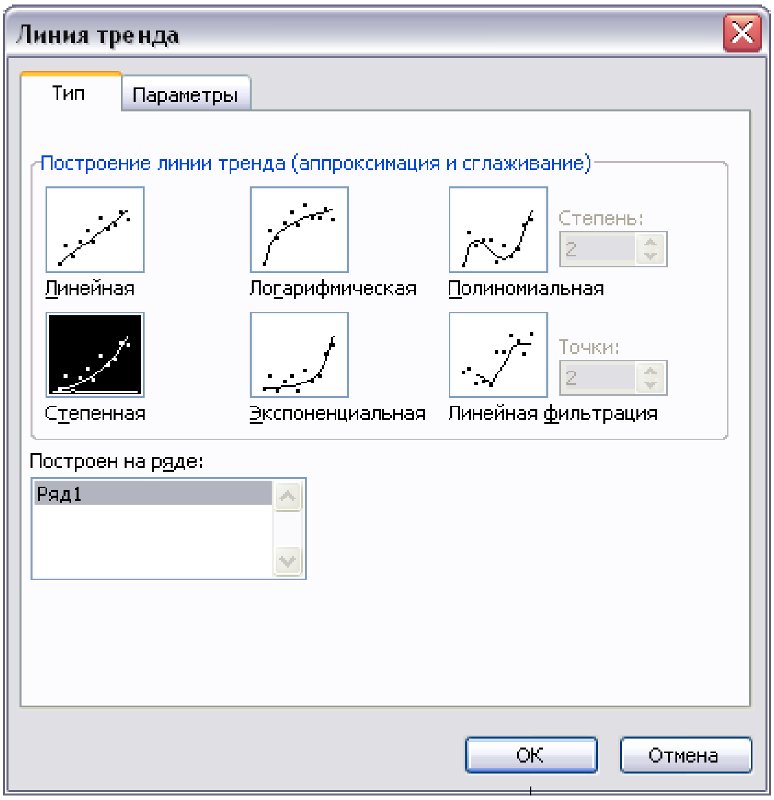


Рис. 2.3. Выбор вида уравнения регрессии

5. На вкладке *«Параметры»* (см. рис. 2.4) отметить опцию *«Показывать уравнение на диаграмме»* и щелкнуть *Ok.*

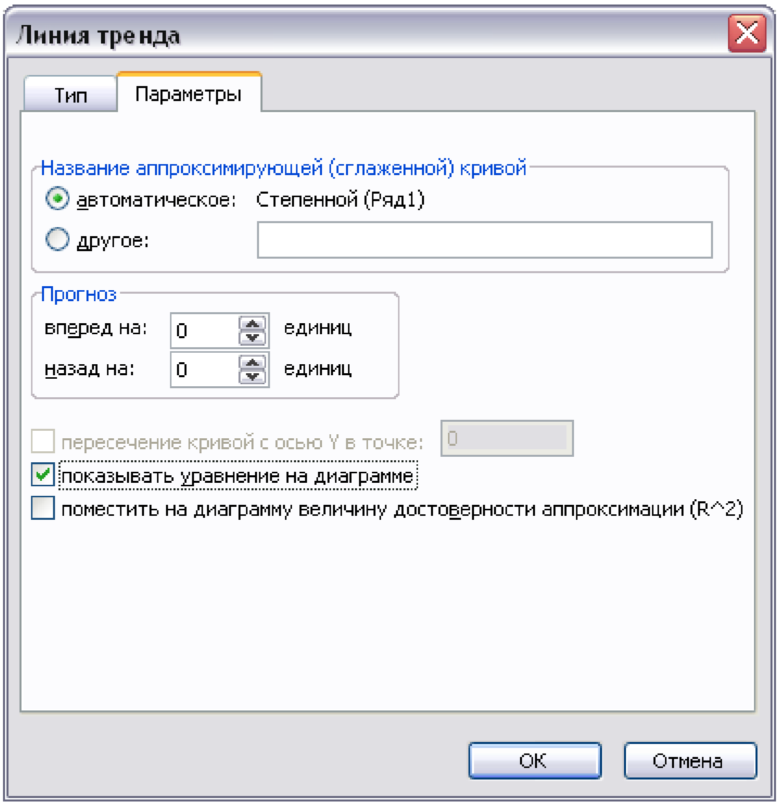


Рис. 2.4. Задание опций вывода информации

После выполнения данных действий получим изображение исходных данных в виде точечной диаграммы, степенную линию тренда и соответствующий вид уравнения регрессии (рис. 2.5).

Рис. 2.5. График и уравнение построенной регрессии

Определение тесноты связи между случайными величинами X на Y при нелинейной зависимости определяет индекс корреляции:

$$R=\sqrt{1-\frac{\sum_{i=1}^{N}(y_i-\tilde{y}_i)^2}{\sum_{i=1}^{N}(y_i-\bar{y})^2}}\,. \qquad (2.3)$$

Коэффициент детерминации и квадрат индекса корреляции вычисляются одинаково, поэтому данные показатели можно сравнить у линейной и нелинейных зависимостей.

Решение. Фрагменты листа Excel с этапами построения графика приведены на рис. 2.2 – 2.5, с вычислениями на рис. 2.6.

Уравнение на диаграмме соответствует построенному ранее уравнению для степенной зависимости $\tilde{y}=1{,}0992\cdot x^{0{,}9727}$.

Индекс корреляции равен:

$$R=\sqrt{1-\frac{35{,}2152}{2108{,}6667}}=0{,}9916\,.$$

Близость индекса корреляции к единице свидетельствует о сильной связи степенного типа между переменными.

=B19*B3^B17

=C3-C15

=C3-H3

	A	B	C	D	E	F	G	H	I	J
1		Исходные данные								
2	Количество	x_i	y_i	x'_i	y'_i	$x_i'^2$	$x'_i y'_i$	$\tilde{y}_i$	$y_i - \bar{y}$	$e_i = y_i - \tilde{y}_i$
3	12	107	102	4,6728	4,6250	21,8353	21,6117	103,5362	-18,6667	-1,5362
4		109	105	4,6913	4,6540	22,0087	21,8333	105,4181	-15,6667	-0,4181
5		110	108	4,7005	4,6821	22,0945	22,0083	106,3588	-12,6667	1,6412
6		113	110	4,7274	4,7005	22,3482	22,2210	109,1793	-10,6667	0,8207
7		120	115	4,7875	4,7449	22,9201	22,7163	115,7525	-5,6667	-0,7525
8		122	117	4,8040	4,7622	23,0786	22,8776	117,6287	-3,6667	-0,6287
9		123	119	4,8122	4,7791	23,1571	22,9980	118,5664	-1,6667	0,4336
10		128	125	4,8520	4,8283	23,5422	23,4271	123,2521	4,3333	1,7479
11		136	132	4,9127	4,8828	24,1342	23,9875	130,7388	11,3333	1,2612
12		140	130	4,9416	4,8675	24,4198	24,0536	134,4777	9,3333	-4,4777
13		145	141	4,9767	4,9488	24,7679	24,6287	139,1471	20,3333	1,8529
14		150	144	5,0106	4,9698	25,1065	24,9019	143,8122	23,3333	0,1878
15	Средние значения	125,2500	120,6667	4,8241	4,7871	23,2844	23,1054	Сумма квадратов	2108,6667	35,2152
16										
17	b_1	0,9727		R	0,9916					
18	b'_0	0,0946		R^2	0,9833					
19	b_0	1,0992								
20										

=(1-J15/I15)^0,5

=E17^2

Рис. 2.6. Определение индекса корреляции и коэффициента детерминации

Коэффициент детерминации для степенной зависимости равен $R^2 = 0{,}9916^2 = 0{,}9833$. Сравнивая коэффициенты детерминации для линейной ($R^2 = 0{,}98328 \approx 0{,}9833$) и степенной зависимости ($R^2 = 0{,}9833$), можно сделать вывод, что обе зависимости одинаково аппроксимируют данные задачи.

Коэффициент детерминации для степенной зависимости показывает, что 98,33% изменения среднего объема потребления домохозяйства обусловлено зависимостью от дохода, остальные 1,67% приходятся на долю прочих факторов, не учтенных в степенной модели.

Вопросы:

– Как построить график для нелинейной зависимости?
– Какая величина определяет тесноту связи при нелинейной зависимости? Как она связана с коэффициентов детерминации?
– Какие значения может принимать R и что это означает?

Тема 3. ЛИНЕЙНАЯ МНОЖЕСТВЕННАЯ РЕГРЕССИЯ

Эта тема включает выполнение двух лабораторных работ, посвященных построению и исследованию уравнения линейной множественной регрессии вида:

$$\tilde{y} = b_0 + b_1 x_1 + b_2 x_2 . \tag{3.1}$$

Выборка для построения этого уравнения взята из следующего примера:

Пример 3.1. Исследуется зависимость рентабельности предприятия *Y* (%) от производительности труда X_1 (у.е.) и среднего возраста оборудования X_2 (в годах). Результаты наблюдений представлены таблицей 3.1.

Таблица 3.1

Рента-бельность, %	Производи-тельность труда	Средний возраст оборуд-я (г)	Рента-бельность, %	Производи-тельность труда	Средний возраст оборуд-я (г)
8	11	18	16	18	11
11	13	15	18	20	10
11	14	14	17	20	10
16	17	10	19	22	9
15	18	11	20	23	8
19	21	9	11	12	15

Лабораторная работа №3.1.

«Коэффициенты уравнения линейной множественной регрессии»

Цель работы. Вычисление коэффициентов уравнения линейной множественной регрессии по выборке из примера 3.1 (см. таблицу 3.1).

Расчетные соотношения. Применив к уравнению (3.1) метод наименьших квадратов, получим вектор $B = \begin{pmatrix} b_0 \\ b_1 \\ b_2 \end{pmatrix}$ коэффициентов уравнения, который рассчитывается следующим образом:

$$B = (X^T X)^{-1} \cdot X^T Y , \tag{3.2}$$

где $X = \begin{pmatrix} 1 & 11 & 18 \\ 1 & 13 & 15 \\ ... & ... & ... \\ 1 & 12 & 15 \end{pmatrix}$ – матрица размером 12×3 ($N \times (m+1)$ – в общем случае), первый столбец которой составлен из единиц, а другие два столбца составлены из значений x_{1i}, x_{2i};

$Y = \begin{pmatrix} 8 \\ 11 \\ ... \\ 11 \end{pmatrix}$ – вектор, составленный из 12 значений (N – в общем случае);

$$X^T = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 11 & 13 & 14 & 17 & 18 & 21 & 18 & 20 & 20 & 22 & 23 & 12 \\ 18 & 15 & 14 & 10 & 11 & 9 & 11 & 10 & 10 & 9 & 8 & 15 \end{pmatrix}$$ – матрица, полученная транспонированием матрицы X;

$(X^T X)^{-1}$ – обратная матрица, полученная обращением произведения матриц $X^T X$.

Матричные функции Excel.

Транспонирование матрицы осуществляется с помощью функции ТРАНСП (категория функций – *Ссылки и массивы*). Обращение к функции имеет вид:

ТРАНСП (*диапазон ячеек*),

где параметр *диапазон ячеек* задает все элементы транспонируемой матрицы (или вектора).

Умножение матриц осуществляется с помощью функции МУМНОЖ (категория функций – *Математические*). Обращение к функции имеет вид:

МУМНОЖ(*диапазон_1; диапазон_2*),

где параметр *диапазон_1* задает элементы первой из перемножаемых матриц, а параметр *диапазон_2* – элементы второй матрицы. При этом перемножаемые матрицы должны иметь соответствующие размеры (если первая матрица $n \times k$, вторая – $k \times m$, то результатом будет матрица $n \times m$).

Обращение матрицы (вычисление обратной матрицы) осуществляется с помощью функции МОБР (категория функций – *Математические).* Обращение к функции имеет вид:

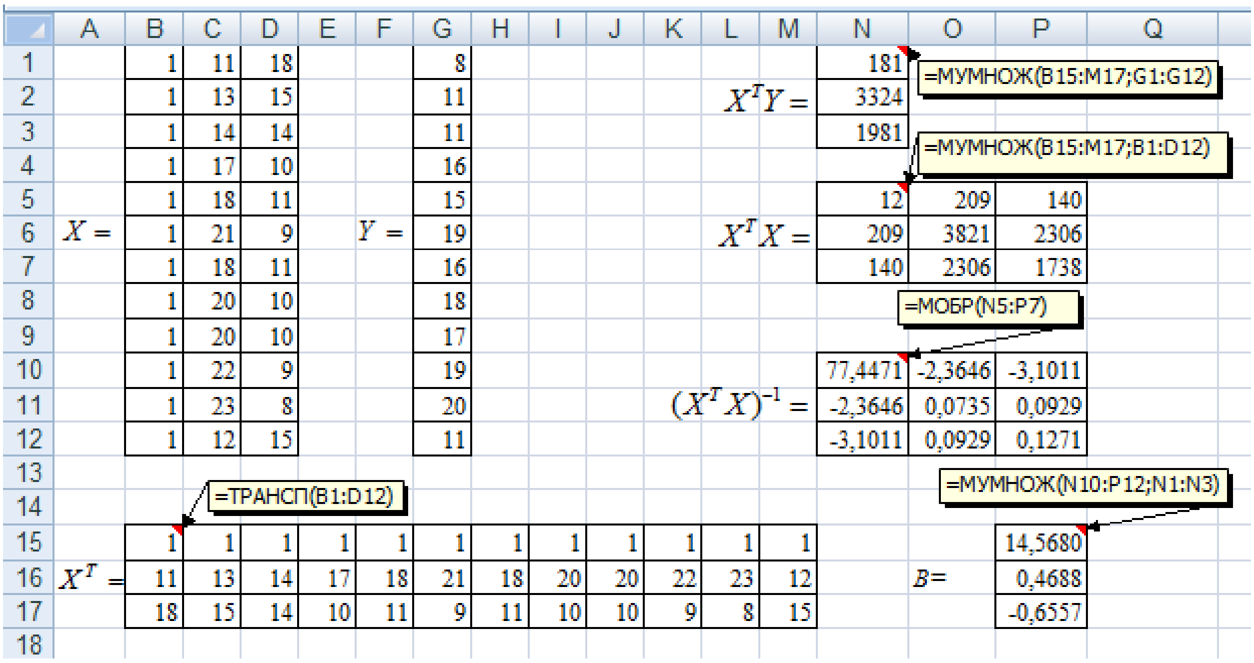
МОБР (*диапазон ячеек*),

где параметр *диапазон ячеек* задает все элементы обращаемой матрицы, которая должна быть квадратной и невырожденной.

При использовании этих функций ***необходимо соблюдать следующий порядок действий:***

- в ячейку ввести арифметическое выражение, содержащее обращение к матричным функциям Excel (в результате будет получено только одно значение);
- выделить диапазон ячеек, в которые будет занесен результат выполнения матричных функций, начиная с ячейки, в которой было получено первое значение (при этом надо учитывать размеры исходных матриц); нажать клавишу F2;
- одновременно нажать клавиши [Ctrl]+[Shift]+[Enter] для получения значений остальных элементов результирующей матрицы или вектора.

Решение. Сформируем матрицу X и вектор Y (рис. 3.1).



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1		1	11	18			8							181	=МУМНОЖ(B15:M17;G1:G12)		
2		1	13	15			11						$X^TY =$	3324			
3		1	14	14			11							1981	=МУМНОЖ(B15:M17;B1:D12)		
4		1	17	10			16										
5		1	18	11			15							12	209	140	
6	$X =$	1	21	9		$Y =$	19						$X^TX =$	209	3821	2306	
7		1	18	11			16							140	2306	1738	
8		1	20	10			18								=МОБР(N5:P7)		
9		1	20	10			17										
10		1	22	9			19							77,4471	-2,3646	-3,1011	
11		1	23	8			20						$(X^TX)^{-1} =$	-2,3646	0,0735	0,0929	
12		1	12	15			11							-3,1011	0,0929	0,1271	
13															=МУМНОЖ(N10:P12;N1:N3)		
14			=ТРАНСП(B1:D12)														
15		1	1	1	1	1	1	1	1	1	1	1	1		14,5680		
16	$X^T =$	11	13	14	17	18	21	18	20	20	22	23	12	$B=$	0,4688		
17		18	15	14	10	11	9	11	10	10	9	8	15		-0,6557		
18																	

Рис. 3.1. Вычисление коэффициентов множественной регрессии

Затем выполним формирование матриц X^T, $X^T X$, $(X^T X)^{-1}$, вектора $X^T Y$ и вычисление вектора B по формуле (3.2).

Получим вектор коэффициентов $B = \begin{pmatrix} 14{,}568 \\ 0{,}4688 \\ -0{,}6557 \end{pmatrix}$, и тогда уравнение регрессии (3.1) примет вид:

$$\tilde{y}(x_1, x_2) = 14{,}568 + 0{,}4688 x_1 - 0{,}6557 x_2 . \qquad (3.3)$$

Дадим интерпретацию оценкам коэффициентов регрессии.

Коэффициент $b_1 = 0{,}4688$ означает, что при увеличении производительности труда на одну единицу рентабельность возрастет на 0,4688% при неизменном значении x_2; коэффициент $b_2 = -0{,}6557$ показывает, что если средний возраст оборудования возрастет на один год, то рентабельность снизится на 0,6557% при неизменном значении x_1.

Вопросы:

- Как применить МНК для построения множественного линейного уравнения регрессии?
- Как определить коэффициенты для множественной регрессии?

Лабораторная работа №3.2.

«Определение параметров модели в режиме Регрессия»

Цель работы. Используя пространственную выборку таблицы 3.1 и применяя режим ***Регрессия***, оценить значимость коэффициентов уравнения регрессии $\tilde{y} = b_0 + b_1 x_1 + b_2 x_2$ и значимость коэффициента детерминации.

Режим ***Регрессия*** *модуля* ***Анализ данных*** в табличном процессоре Excel осуществляет вычисление коэффициентов линейной регрессии с m переменными, расчетных значений для проверки значимости коэффициентов регрессии и детерминации, построение доверительных интервалов и др.

Для вызова режима *Регрессия* модуля *Анализ данных* необходимо:

– обратиться к пункту меню ***Сервис***;
– в появившемся меню выбрать команду ***Анализ данных;***
– в списке режимов работы модуля ***Анализ данных*** выбрать режим *Регрессия* и щелкнуть на кнопке ***Ok***.

Замечание: Если в меню ***Сервис*** нет пункта ***Анализ данных***, то выбрать пункт ***Надстройки*** и отметить надстройку ***Пакет анализа.*** Повторить вызов режима *Регрессия*.

После вызова режима *Регрессия* на экране появляется диалоговое окно (рис. 3.2), в котором задаются следующие параметры:

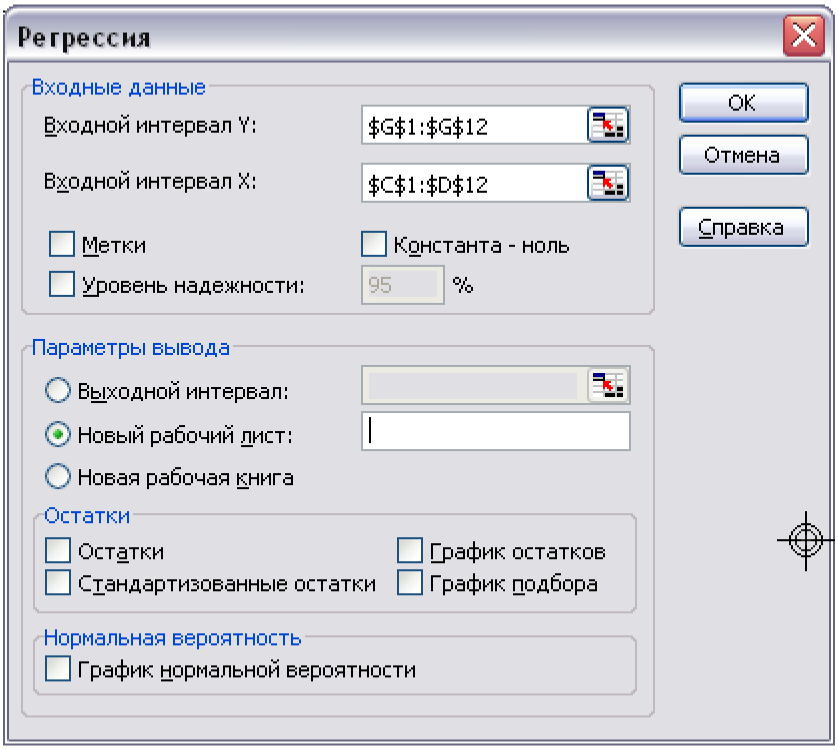


Рис. 3.2. Диалоговое окно режима *Регрессия*

1. *Входной интервал Y* – вводится диапазон адресов ячеек, содержащих значения y_i (ячейки должны составлять один столбец).
2. *Входной интервал X* – вводится диапазон адресов ячеек, содержащих значения независимых переменных. Значения каждой переменной представляются одним или несколькими столбцами, расположенными рядом (непрерывный диапазон). Количество переменных не более 16 (т.е. $m \le 16$).
3. *Метки* – включается, если первая строка во входном диапазоне содержит заголовок. В этом случае автоматически будут созданы стандартные названия.

4. *Уровень надежности* – при включении этого параметра задается надежность $\gamma = 1 - \alpha$ (по умолчанию $\gamma = 95\%$).
5. *Константа - ноль* – при включении этого параметра коэффициент $b_0 = 0$.
6. *Выходной интервал* – при включении активизируется поле, в которое необходимо ввести адрес левой верхней ячейки выходного диапазона, который будет содержать ячейки с результатами вычислений режима *Регрессия.*
7. *Новый рабочий лист* – при включении этого параметра открывается новый лист, в который начиная с ячейки *A1* вставляются результаты работы режима *Регрессия.*
8. *Новая рабочая книга* – при включении этого параметра открывается новая рабочая книга, на первом листе которой, начиная с ячейки *A1*, вставляются результаты работы режима *Регрессия.*
9. *Остатки* – при включении вычисляется столбец, содержащий отклонения $e_i = y_i - \tilde{y}_i$, $i = 1,...,N$.

Также можно задать вывод стандартизированных остатков и графиков.

	A	B	C	D	E	F	G
1	ВЫВОД ИТОГОВ						
2							
3	*Регрессионная статистика*						
4	Множественный R	0,99167					
5	R-квадрат	0,98341					
6	Нормированный R-квадрат	0,97973					
7	Стандартная ошибка	0,55794					
8	Наблюдения	12					
9							
10	Дисперсионный анализ						
11		*df*	*SS*	*MS*	*F*	*Значимость F*	
12	Регрессия	2	166,11502	83,05751	266,81329	0,0000000097	
13	Остаток	9	2,80165	0,31129			
14	Итого	11	168,91667				
15							
16		*Коэфф-ты*	*Стандартная ошибка*	*t-статистика*	*P-Значение*	*Нижние 95%*	*Верхние 95%*
17	Y-пересечение	14,56795	4,91008	2,96695	0,01578	3,46058	25,67532
18	Переменная X 1	0,46880	0,15127	3,09907	0,01274	0,12660	0,81101
19	Переменная X 2	-0,65568	0,19888	-3,29684	0,00928	-1,10559	-0,20578

Рис. 3.3. Результаты работы режима *Регрессия*

Решение. Пусть данные введены, как показано на рис. 3.1.

Вызовем режим *Регрессия* и в диалоговом окне зададим необходимые параметры (см. рис. 3.2). Результаты работы приводятся на рис. 3.3.

Дадим краткую интерпретацию показателям, значения которых вычисляются в режиме *Регрессия.*

Рассмотрим показатели, объединенные названием *Регрессионная статистика* (см. рис. 3.3).

Множественный R – корень квадратный из коэффициента детерминации.

R-квадрат – коэффициент детерминации R^2.

Нормированный R-квадрат – скорректированный коэффициент детерминации $\overline{R}^2$.

Стандартная ошибка – ошибка регрессии S.

Наблюдения – число наблюдений *N*.

Рассмотрим показатели, объединенные названием *Дисперсионный анализ* (рис. 3.3).

df – число степеней свободы. Для строки *Регрессия* показатель равен числу независимых переменных $k_r = m$; для строки *Остаток* – равен $k_e = N - m - 1$; для строки *Итого* – равен $k_r + k_e$.

SS – сумма квадратов отклонений. Для строки *Регрессия* показатель равен величине $SS_r = Q_r = \sum_{i=1}^{N}(y_i - \bar{y})^2$; для строки *Остаток* – равен величине $SS_e = Q_e = \sum_{i=1}^{N}(y_i - \tilde{y}_i)^2$; для строки *Итого* – равен $Q = Q_r + Q_e$.

MS – дисперсии, вычисленные по формуле $MS = \dfrac{SS}{df}$, т.е. дисперсия на одну степень свободы.

F – значение, равное *F*-критерию Фишера, вычисленного по формуле:

$$F_{расч} = \frac{SS_r / k_r}{SS_e / k_e} = \frac{R^2}{1 - R^2} \frac{N - m - 1}{m}.$$

Значимость F – значение уровня значимости, соответствующее вычисленной величине *F*-критерия и равное вероятности $P(F(k_r, k_e) \ge F_{расч})$, где $F(k_r, k_e)$ – случайная величина, подчиняющаяся распределению Фишера с k_r и k_e степенями свободы. Эту вероятность можно также определить с помощью функции Excel (категория – *Статистические*) FРАСП$(F_{расч}; k_r; k_e)$. Если вероятность меньше уровня значимости α, то построенная регрессия является значимой.

Коэффициенты – вычисленные значения коэффициентов $b_0, b_1, ..., b_m$.

Стандартная ошибка – значения S_{b_j}, $j = 0, ..., m$, вычисленные по формуле

$$S_{b_j} = \sqrt{S^2 \left\{ \left(X^T X \right)^{-1} \right\}_{jj}},$$

где $\left\{ \left(X^T X \right)^{-1} \right\}_{jj}$ – диагональный элемент матрицы $\left(X^T X \right)^{-1}$.

t*–статистика* – значения статистик $t_{расч}(b_j) = \dfrac{b_j}{S_{b_j}}$.

P–значение – содержит вероятности случайных событий $P\big(t(N - m - 1) \ge t_{расч}(b_j)\big)$, где $t(N - m - 1)$ – случайная величина, подчиняющаяся распределению Стьюдента с $N - m - 1$ степенями свободы. Эту вероятность можно также определить с помощью функции Excel (категория – *Статистические*) СТЬЮДРАСП$\left(t_{расч}; N-m-1; 2\right)$. Если эта вероятность меньше уровня значимости α, то принимается гипотеза о значимости соответствующего коэффициента регрессии.

Из рис. 3.3 видно, что все коэффициенты является значимыми.

Нижние 95% и Верхние 95% – обозначают соответственно нижние и верхние доверительные интервалы для оцениваемых коэффициентов b_j.

Определим статистическую значимость коэффициентов по критерию Стьюдента, взяв некоторые расчетные значения, полученные в режиме Регрессия (см. рис. 3.3).

Расчетные значения *t*-критерия Стьюдента:

$t_{расч}(b_0) = 2{,}9669$, $t_{расч}(b_1) = 3{,}0991$, $t_{расч}(b_2) = -3{,}2968$.

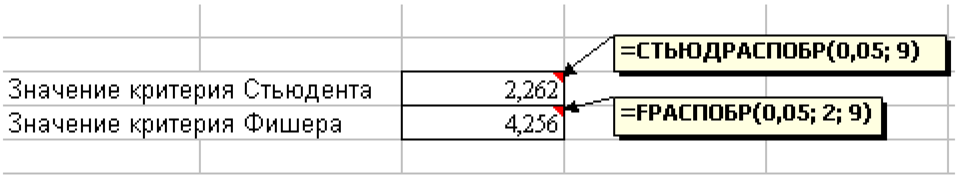


Рис. 3.4. Критические значения для оценки значимости

Критическое значение определяется как для парной регрессии, только в данном случае число объясняющих переменных *m=2*:

$t_{кр} = СТЬЮДРАСПОБР(0{,}05; 12-2-1) = СТЬЮДРАСПОБР(0{,}05; 9) = 2{,}262$.

Поскольку для всех коэффициентов $\left|t_{расч}\right| > 2{,}262 = t_{кр}$, то все коэффициенты статистически значимы на уровне значимости $\alpha = 0{,}05$.

Определим статистическую значимость коэффициента детерминации по критерию Фишера по результатам режима Регрессия.

Коэффициент детерминации $R^2 = 0{,}9834$,

F-статистика $F_{расч} = \dfrac{R^2}{1-R^2} \cdot \dfrac{N-m-1}{m} = \dfrac{0{,}9834}{1-0{,}9834} \cdot \dfrac{12-2-1}{2} = 266{,}8133$.

Найдем критическое значение $F_{кр} = FРАСПОБР(0{,}05; 2; 12-2-1) = 4{,}256$.

Поскольку $F_{расч} = 266{,}8133 > 4{,}256 = F_{кр}$, то коэффициент детерминации статистически значим.

Статистическая значимость коэффициентов уравнения и коэффициента детерминации говорят о высоком качестве модели множественной регрессии; наблюдается тесная линейная зависимость рентабельности предприятия от производительности труда и среднего возраста оборудования.

Вопросы:

– Можно ли применять режим Регрессия к парной линейной регрессии?
– Как определить значимость коэффициентов для множественной регрессии?

СПИСОК ЛИТЕРАТУРЫ

1. Айвазян С.А., Мхитарян В.С. Прикладная статистика. Основы эконометрики. Т.2. М.: ЮНИТИ, 2001.
2. Бородич С.А. Эконометрика: Учебное пособие. Мн.: Новое знание, 2001. – 408 с. – (Экономическое образование)
3. Гмурман В.Е. Теория вероятностей и математическая статистика. М.: Высшая школа, 1999.
4. Доугерти К. Введение в эконометрику. Пер. с англ. М.: Инфра-М, 1997.
5. Замков О.О., Черемных Ю.А., Толстопятенко А.В. Математические методы в экономике. М.: Дело и Сервис, 1999.
6. Катышев П.К., Магнус Я.Р., Пересецкий А.А. Сборник задач к начальному курсу эконометрики. 3-е изд., испр. М.: Дело, 2003.
7. Кремер Н.Ш., Путко Б.А. Эконометрика: Учебник для вузов. Под ред. проф. Н.Ш. Кремера. – М.: ЮНИТИ-ДАНА, 2002. –311 с.
8. Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика. Начальный курс: Учебник. – 7-е изд., испр. – М.: Дело, 2005.
9. Практикум по эконометрике: Учебное пособие / И.И. Елисеева, С.В. Курышева, Н.М. Гордеенко и др.; Под ред. И.И. Елисеевой. – М.: Финансы и статистика, 2003. –192 с.: ил.
10. Чавкин А.М. Методы и модели рационального управления в рыночной экономике: разработка управленческих решений: Учеб. пособие. М.: Финансы и статистика, 2001.
11. Эконометрика: Учебник / Елисеева И.И., Курышева С.В., Костеева Т.В. и др.; Под ред. Елисеевой И.И. – 2-е изд., перераб. и доп. – М.: Финансы и статистика, 2006.

ПРИЛОЖЕНИЕ 1

Задания для выполнения контрольных работ

Задание 1:

1. Построить линейную парную регрессию (регрессию вида $\tilde{y} = b_0 + b_1 x$). Вычисление коэффициентов b_0, b_1 выполнить методом наименьших квадратов, дать интерпретацию.
2. Вычислить выборочный коэффициент корреляции и проверить гипотезу о его значимости.
3. Проверить значимость коэффициентов регрессии.
4. Используя построенное уравнение, спрогнозировать значение $\tilde{y}_p$ при $x_p = 1{,}2 \cdot \bar{x}$.
5. Определить, есть или нет автокорреляция остатков с помощью критерия Дарбина-Уотсона.
6. Вычислить коэффициент детерминации и проверить его значимость.
7. Оценить прогнозные качества модели.
8. Сделать общий вывод по качеству построенной модели.

Задание 2:

1. Построить уравнение нелинейной парной регрессии:
   – для нечетных вариантов (1, 3, …) логарифмическую зависимость $\tilde{y} = b_0 + b_1 \ln x$;
   – для четных вариантов (2, 4, …) экспоненциальную зависимость $\tilde{y} = b_0 e^{b_1 x}$.
2. Построить график для найденной зависимости.
3. Рассчитать индекс корреляции, сделать вывод о тесноте нелинейной связи между переменными.

Условия задач для заданий 1, 2.

Вариант 1. Торговое предприятие имеет несколько филиалов. Исследуется зависимость годового товарооборота отдельного филиала от размера торговой площади:

Торговая площадь (X)	0,31	0,98	1,21	1,29	1,12	1,49	0,78	0,94	1,29	0,48	0,24
Товарооборот	2,93	5,27	6,85	7,01	7,02	8,35	4,33	5,77	7,68	3,16	1,52

Вариант 2. Торговое предприятие имеет несколько филиалов. Исследуется зависимость годового товарооборота отдельного филиала от среднедневной интенсивности потока покупателей:

Интенсивность потока покупателей (X)	9,24	10,51	10,81	11,89	12,72	13,92	12,54	12,36	12,27	11,01	8,25
Товарооборот	2,93	5,27	6,85	7,01	7,02	8,35	4,33	5,77	7,68	3,16	1,52

Вариант 3. Исследуется зависимость себестоимости 1т литья У (руб.) от выработки литья на одного работающего Х (т) по 11 литейным цехам заводов:

X	14,6	13,5	21,5	17,4	44,8	111,9	20,1	28,1	22,3	25,3	56,0
Y	239	254	262	251	158	101	259	186	204	198	170

Вариант 4. Исследуется зависимость себестоимости 1т литья У (руб.) от брака литья Х (т) по 11 литейным цехам заводов:

X	4,2	5,5	6,7	7,7	1,2	2,2	8,4	6,4	4,2	3,2	3,1
Y	239	254	262	251	158	101	259	186	204	198	170

Вариант 5. В таблице приведены данные о годовом потреблении свинины У на душу населения в США (в фунтах) и оптовых ценах на свинину Х (в долларах за фунт) за период с 1948 по 1957 год:

X	0,54	0,47	0,46	0,45	0,43	0,5	0,52	0,53	0,51	0,55
Y	67,8	67,7	69,2	71,9	72,4	68,5	67	66,6	64,5	61,5

Вариант 6. В таблице приведены данные о размерах совокупного располагаемого дохода X и совокупных расходах на личное потребление Y в США в период с 1979 по 1988 год. Обе величины выражены в текущих долларах США.

X	695	751	810	914	998	1096	1194	1313	1474	1650
Y	621	672	737	811	887	976	1084	1204	1346	1506

Вариант 7. Приведены данные по располагаемому доходу домохозяйств (X) и затратам домохозяйств на розничные покупки (Y) за 10 лет:

X	9,09	9,14	9,1	9,28	9,23	9,35	9,53	9,76	10,28	10,67
Y	5,49	5,54	5,31	5,51	5,42	5,32	5,54	5,69	5,87	6,16

Вариант 8. По имеющимся статистическим данным исследуется зависимость между темпом прироста заработной платы (X) и уровнем безработицы (Y) (в %):

Год	79	80	81	82	83	84	85	86	87	88
X	2,67	2,73	2,8	2,93	3,02	3,15	3,27	3,45	3,6	3,8
Y	1,8	1,9	1,8	1,85	1,7	1,5	1,25	1,4	1,3	1,2

Вариант 9. Анализируется зависимость между инфляцией (X) и безработицей (Y). Используются статистические данные за 10 лет:

X	5,12	4,08	3,07	2,54	3,38	2,2	0,7	1,1	0,93	1
Y	9,55	9,1	8,5	6,53	4,63	6,5	3,69	3,71	3,2	3,6

Вариант 10. В таблице приведены данные об изменении потребительского спроса на куриные яйца семейного хозяйства Y (шт.) в зависимости от цены X (руб.) на этот продукт в течение 10 недель:

X	17	17,5	18	19	19	20	21	22	23	23
Y	10	11	12	10	11	12	12	14	13	14

Задание 3:

Имеются статистические данные, описывающие зависимость производительности труда в некоторой отрасли производства (переменная *Y*) от удельного веса рабочих с технической подготовкой (объясняющая переменная X_1) и удельного веса механизированных работ (объясняющая переменная X_2).

№ завода	Удельный вес рабочих с технической подготовкой, %	Удельный вес механизированных работ, %	Производительность труда
1	64+*V*	84+*V*	4350
2	61+*V*	83+*V*	4100
3	47+*V*	67+*V*	2950
4	46+*V*	63+*V*	2900
5	49+*V*	69+*V*	2950
6	54+*V*	70+*V*	3350
7	53+*V*	73+*V*	3410
8	61+*V*	81+*V*	4050
9	57+*V*	77+*V*	3650
10	54+*V*	72+*V*	3450
11	60+*V*	80+*V*	3950
12	67+*V*	83+*V*	4400
13	63+*V*	85+*V*	4220

В таблице значение *V* – последняя цифра в зачетной книжке.

1. Найти коэффициенты множественной регрессии $\tilde{y} = b_0 + b_1x_1 + b_2x_2$ в Excel.
2. Вызвать режим ***Регрессия***, проверить правильность коэффициентов, рассчитанных с помощью функций Excel.

ПРИЛОЖЕНИЕ 2

Вопросы для подготовки к зачету

1. Понятие эконометрики, цели и задачи дисциплины. Числовые характеристики случайной величины (СВ). Законы распределения СВ.
2. Ковариация, коэффициент корреляции. Свойства коэффициента корреляции. Оценка значимости коэффициента корреляции.
3. Линейная модель парной регрессии. Оценка параметров модели с помощью метода наименьших квадратов (МНК).
4. Предпосылки МНК.
5. Нелинейные модели регрессии и их линеаризация, индекс корреляции.
6. Оценка значимости коэффициентов парной регрессии (для линейной и нелинейных моделей).
7. Автокорреляция остатков. Критерий знаков. Критерий Дарбина-Уотсона.
8. Спецификация парной регрессии.
9. Линейная модель множественной регрессии. Оценка параметров модели с помощью МНК.
10. Парная и частная корреляция для множественной регрессии.
11. Коэффициент детерминации, проверка значимости. Скорректированный коэффициент детерминации.
12. Оценка значимости коэффициентов линейной модели множественной регрессии.
13. Мультиколлинеарность, способы ее устранения.
14. Построение точечных и интервальных прогнозов с помощью модели парной регрессии.
15. Регрессионные модели с переменной структурой, фиктивные переменные.
16. Обобщенный метод наименьших квадратов.
17. Гетероскедастичность.
18. Модели стационарных и нестационарных временных рядов. Их идентификация.
19. Системы одновременных уравнений. Общий вид. Модель спроса и предложения.
20. Системы одновременных уравнений. Проблема идентифицируемости.
21. Метод инструментальных переменных.
22. Косвенный, двухшаговый и трехшаговый МНК.

ПРИЛОЖЕНИЕ 3

Распределение Стьюдента (t-распределение)

ν – число степеней свободы,

α – уровень значимости.

Пример: $t_{\alpha;\ \nu} = t_{0,05;\ 20} = 1{,}725$

ν	α								
	0,4	**0,25**	**0,1**	**0,05**	**0,025**	**0,01**	**0,005**	**0,001**	**0,0005**
1	0,325	1,000	3,078	6,314	12,706	31,821	63,656	318,289	636,578
2	0,289	0,816	1,886	2,920	4,303	6,965	9,925	22,328	31,600
3	0,277	0,765	1,638	2,353	3,182	4,541	5,841	10,214	12,924
4	0,271	0,741	1,533	2,132	2,776	3,747	4,604	7,173	8,610
5	0,267	0,727	1,476	2,015	2,571	3,365	4,032	5,894	6,869
6	0,265	0,718	1,440	1,943	2,447	3,143	3,707	5,208	5,959
7	0,263	0,711	1,415	1,895	2,365	2,998	3,499	4,785	5,408
8	0,262	0,706	1,397	1,860	2,306	2,896	3,355	4,501	5,041
9	0,261	0,703	1,383	1,833	2,262	2,821	3,250	4,297	4,781
10	0,260	0,700	1,372	1,812	2,228	2,764	3,169	4,144	4,587
11	0,260	0,697	1,363	1,796	2,201	2,718	3,106	4,025	4,437
12	0,259	0,695	1,356	1,782	2,179	2,681	3,055	3,930	4,318
13	0,259	0,694	1,350	1,771	2,160	2,650	3,012	3,852	4,221
14	0,258	0,692	1,345	1,761	2,145	2,624	2,977	3,787	4,140
15	0,258	0,691	1,341	1,753	2,131	2,602	2,947	3,733	4,073
16	0,258	0,690	1,337	1,746	2,120	2,583	2,921	3,686	4,015
17	0,257	0,689	1,333	1,740	2,110	2,567	2,898	3,646	3,965
18	0,257	0,688	1,330	1,734	2,101	2,552	2,878	3,610	3,922
19	0,257	0,688	1,328	1,729	2,093	2,539	2,861	3,579	3,883
20	0,257	0,687	1,325	1,725	2,086	2,528	2,845	3,552	3,850
21	0,257	0,686	1,323	1,721	2,080	2,518	2,831	3,527	3,819
22	0,256	0,686	1,321	1,717	2,074	2,508	2,819	3,505	3,792
23	0,256	0,685	1,319	1,714	2,069	2,500	2,807	3,485	3,768
24	0,256	0,685	1,318	1,711	2,064	2,492	2,797	3,467	3,745
25	0,256	0,684	1,316	1,708	2,060	2,485	2,787	3,450	3,725
26	0,256	0,684	1,315	1,706	2,056	2,479	2,779	3,435	3,707
27	0,256	0,684	1,314	1,703	2,052	2,473	2,771	3,421	3,689
28	0,256	0,683	1,313	1,701	2,048	2,467	2,763	3,408	3,674
29	0,256	0,683	1,311	1,699	2,045	2,462	2,756	3,396	3,660
30	0,256	0,683	1,310	1,697	2,042	2,457	2,750	3,385	3,646
40	0,255	0,681	1,303	1,684	2,021	2,423	2,704	3,307	3,551
50	0,255	0,679	1,299	1,676	2,009	2,403	2,678	3,261	3,496
60	0,254	0,679	1,296	1,671	2,000	2,390	2,660	3,232	3,460
80	0,254	0,678	1,292	1,664	1,990	2,374	2,639	3,195	3,416
100	0,254	0,677	1,290	1,660	1,984	2,364	2,626	3,174	3,390
120	0,254	0,677	1,289	1,658	1,980	2,358	2,617	3,160	3,373
200	0,254	0,676	1,286	1,653	1,972	2,345	2,601	3,131	3,340
500	0,253	0,675	1,283	1,648	1,965	2,334	2,586	3,107	3,310
∞	0,253	0,674	1,282	1,645	1,960	2,326	2,576	3,090	3,290

ПРИЛОЖЕНИЕ 4

Распределение Дарбина–Уотсона

Критические точки d_L и d_U при уровне значимости $\alpha = 0{,}05$ (N – объем выборки, m – число объясняющих переменных в уравнении регрессии)

N	$m=1$		$m=2$		$m=3$		$m=4$		$m=5$		$m=6$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
6	0,610	1,400										
7	0,700	1,356	0,467	1,896								
8	0,763	1,332	0,359	1,777	0,368	2,287						
9	0,824	1,320	0,629	1,699	0,435	2,128	0,296	2,388				
10	0,879	1,320	0,697	1,641	0,525	2,016	0,376	2,414	0,243	2,822		
11	0,927	1,324	0,658	1,604	0,595	1,928	0,444	2,283	0,316	2,645	0,203	3,005
12	0,971	1,331	0,812	1,579	0,658	1,864	0,512	2,177	0,379	2,506	0,268	2,932
13	1,010	1,340	0,861	1,562	0,715	1,816	0,574	2,094	0,445	2,390	0,328	2,692
14	1,045	1,330	0,905	1,551	0,767	1,779	0,632	2,030	0,505	2,296	0,389	2,572
15	1,077	1,361	0,946	1,543	0,814	1,750	0,685	1,977	0,562	2,220	0,447	2,472
16	1,106	1,371	0,982	1,539	0,857	1,728	0,734	1,935	0,615	2,157	0,502	2,388
17	1,133	1,381	1,015	1,536	0,897	1,710	0,779	1,900	0,664	2,104	0,554	2,318
18	1,158	1,391	1,046	1,535	0,933	1,696	0,820	1,872	0,710	2,060	0,603	2,257
19	1,180	1,401	1,074	1,536	0,967	1,685	0,859	1,848	0,752	2,023	0,649	2,206
20	1,201	1,411	1,100	1,537	0,998	1,676	0,894	1,828	0,792	1,991	0,692	2,162
21	1,221	1,420	1,125	1,538	1,026	1,669	0,927	1,812	0,829	1,964	0,732	2,124
22	1,239	1,429	1,147	1,541	1,053	1,664	0,958	1,797	0,863	1,940	0,769	2,090
23	1,257	1,437	1,168	1,543	1,078	1,660	0,986	1,785	0,895	1,920	0,804	2,061
24	1,273	1,446	1,188	1,546	1,101	1,656	1,013	1,775	0,925	1,902	0,837	2,035
25	1,288	1,454	1,206	1,550	1,123	1,654	1,038	1,767	0,953	1,886	0,868	2,012
26	1,302	1,461	1,224	1,553	1,143	1,652	1,062	1,759	0,979	1,873	0,897	1,992
27	1,316	1,469	1,240	1,556	1,162	1,651	1,084	1,753	1,104	1,861	0,925	1,974
28	1,328	1,476	1,255	1,560	1,181	1,650	1,104	1,747	1,028	1,850	0,951	1,958
29	1,341	1,483	1,270	1,563	1,198	1,650	1,124	1,743	1,050	1,841	0,975	1,944
30	1,352	1,489	1,284	1,567	1,214	1,650	1,143	1,739	1,071	1,833	0,998	1,931
31	1,363	1,496	1,297	1,570	1,229	1,650	1,160	1,735	1,090	1,825	1,020	1,920
32	1,373	1,502	1,309	1,574	1,244	1,650	1,177	1,732	1,109	1,819	1,041	1,909
33	1,383	1,508	1,321	1,577	1,258	1,651	1,193	1,730	1,127	1,813	1,061	1,900
34	1,393	1,514	1,333	1,580	1,271	1,652	1,208	1,728	1,144	1,808	1,080	1,891
35	1,402	1,519	1,343	1,584	1,283	1,653	1,222	1,726	1,160	1,803	1,097	1,884
36	1,411	1,525	1,354	1,587	1,295	1,654	1,236	1,724	1,175	1,799	1,114	1,877
37	1,419	1,530	1,364	1,590	1,307	1,655	1,249	1,723	1,190	1,795	1,131	1,870
38	1,427	1,535	1,373	1,594	1,318	1,656	1,261	1,722	1,204	1,792	1,146	1,864
39	1,435	1,540	1,382	1,597	1,328	1,658	1,273	1,722	1,218	1,789	1,161	1,859
40	1,442	1,544	1,391	1,600	1,338	1,659	1,285	1,721	1,230	1,786	1,175	1,854
45	1,475	1,566	1,430	1,615	1,383	1,668	1,336	1,720	1,287	1,776	1,238	1,835
50	1,503	1,585	1,462	1,628	1,421	1,674	1,378	1,721	1,335	1,771	1,291	1,822
55	1,528	1,601	1,490	1,641	1,452	1,681	1,414	1,724	1,374	1,768	1,334	1,814
60	1,549	1,616	1,514	1,652	1,480	1,689	1,444	1,727	1,408	1,767	1,372	1,808
65	1,567	1,629	1,536	1,662	1,503	1,696	1,471	1,731	1,438	1,767	1,404	1,805
70	1,583	1,641	1,554	1,672	1,525	1,703	1,494	1,735	1,464	1,768	1,433	1,802
75	1,598	1,650	1,571	1,680	1,543	1,709	1,515	1,739	1,487	1,770	1,458	1,801
80	1,611	1,662	1,586	1,688	1,560	1,715	1,534	1,743	1,507	1,772	1,480	1,801
85	1,624	1,671	1,600	1,696	1,575	1,721	1,550	1,747	1,525	1,774	1,500	1,801
90	1,636	1,679	1,612	1,703	1,589	1,726	1,566	1,751	1,542	1,776	1,518	1,801
95	1,645	1,687	1,623	1,709	1,602	1,732	1,579	1,755	1,557	1,778	1,535	1,802
100	1,654	1,694	1,634	1,715	1,613	1,736	1,592	1,758	1,571	1,780	1,550	1,803
150	1,720	1,746	1,706	1,760	1,693	1,774	1,679	1,788	1,665	1,802	1,651	1,817
200	1,758	1,778	1,748	1,789	1,738	1,799	1,728	1,810	1,718	1,820	1,707	1,831

ПРИЛОЖЕНИЕ 5

Распределение Фишера (F -распределение)

Уровень значимости $\alpha = 0{,}05$

ν_2	ν_1 (число степеней свободы)										
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
1	161,5	199,50	215,71	224,58	230,16	233,99	236,77	238,88	240,54	241,88	242,98
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,40
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,76
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,94
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,70
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,03
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,60
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,31
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,10
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,94
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,82
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,72
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,63
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,57
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,51
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,46
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,41
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,37
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,34
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,31
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,28
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,26
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,24
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,22
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,20
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,18
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20	2,17
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	2,15
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18	2,14
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,13
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2,04
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,95
80	3,96	3,11	2,72	2,49	2,33	2,21	2,13	2,06	2,00	1,95	1,91
100	3,94	3,09	2,70	2,46	2,31	2,19	2,10	2,03	1,97	1,93	1,89
120	3,92	3,07	2,68	2,45	2,29	2,18	2,09	2,02	1,96	1,91	1,87
200	3,89	3,04	2,65	2,42	2,26	2,14	2,06	1,98	1,93	1,88	1,84
∞	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83	1,79

ν_2	ν_1 (число степеней свободы)										
	12	**15**	**20**	**24**	**30**	**40**	**50**	**60**	**100**	**120**	**200**	∞
1	243,90	246	248	249	250	251	252	252,2	253	253,3	253,68	254
2	19,41	19,43	19,45	19,45	19,46	19,47	19,48	19,48	19,49	19,49	19,49	19,50
3	8,74	8,70	8,66	8,64	8,62	8,59	8,58	8,57	8,55	8,55	8,54	8,53
4	5,91	5,86	5,80	5,77	5,75	5,72	5,70	5,69	5,66	5,66	5,65	5,63
5	4,68	4,62	4,56	4,53	4,50	4,46	4,44	4,43	4,41	4,40	4,39	4,36
6	4,00	3,94	3,87	3,84	3,81	3,77	3,75	3,74	3,71	3,70	3,69	3,67
7	3,57	3,51	3,44	3,41	3,38	3,34	3,32	3,30	3,27	3,27	3,25	3,23
8	3,28	3,22	3,15	3,12	3,08	3,04	3,02	3,01	2,97	2,97	2,95	2,93
9	3,07	3,01	2,94	2,90	2,86	2,83	2,80	2,79	2,76	2,75	2,73	2,71
10	2,91	2,85	2,77	2,74	2,70	2,66	2,64	2,62	2,59	2,58	2,56	2,54
11	2,79	2,72	2,65	2,61	2,57	2,53	2,51	2,49	2,46	2,45	2,43	2,40
12	2,69	2,62	2,54	2,51	2,47	2,43	2,40	2,38	2,35	2,34	2,32	2,30
13	2,60	2,53	2,46	2,42	2,38	2,34	2,31	2,30	2,26	2,25	2,23	2,21
14	2,53	2,46	2,39	2,35	2,31	2,27	2,24	2,22	2,19	2,18	2,16	2,13
15	2,48	2,40	2,33	2,29	2,25	2,20	2,18	2,16	2,12	2,11	2,10	2,07
16	2,42	2,35	2,28	2,24	2,19	2,15	2,12	2,11	2,07	2,06	2,04	2,01
17	2,38	2,31	2,23	2,19	2,15	2,10	2,08	2,06	2,02	2,01	1,99	1,96
18	2,34	2,27	2,19	2,15	2,11	2,06	2,04	2,02	1,98	1,97	1,95	1,92
19	2,31	2,23	2,16	2,11	2,07	2,03	2,00	1,98	1,94	1,93	1,91	1,88
20	2,28	2,20	2,12	2,08	2,04	1,99	1,97	1,95	1,91	1,90	1,88	1,84
21	2,25	2,18	2,10	2,05	2,01	1,96	1,94	1,92	1,88	1,87	1,84	1,81
22	2,23	2,15	2,07	2,03	1,98	1,94	1,91	1,89	1,85	1,84	1,82	1,78
23	2,20	2,13	2,05	2,01	1,96	1,91	1,88	1,86	1,82	1,81	1,79	1,76
24	2,18	2,11	2,03	1,98	1,94	1,89	1,86	1,84	1,80	1,79	1,77	1,73
25	2,16	2,09	2,01	1,96	1,92	1,87	1,84	1,82	1,78	1,77	1,75	1,71
26	2,15	2,07	1,99	1,95	1,90	1,85	1,82	1,80	1,76	1,75	1,73	1,69
27	2,13	2,06	1,97	1,93	1,88	1,84	1,81	1,79	1,74	1,73	1,71	1,67
28	2,12	2,04	1,96	1,91	1,87	1,82	1,79	1,77	1,73	1,71	1,69	1,65
29	2,10	2,03	1,94	1,90	1,85	1,81	1,77	1,75	1,71	1,70	1,67	1,64
30	2,09	2,01	1,93	1,89	1,84	1,79	1,76	1,74	1,70	1,68	1,66	1,62
40	2,00	1,92	1,84	1,79	1,74	1,69	1,66	1,64	1,59	1,58	1,55	1,51
60	1,92	1,84	1,75	1,70	1,65	1,59	1,56	1,53	1,48	1,47	1,44	1,39
80	1,88	1,79	1,70	1,65	1,60	1,54	1,51	1,48	1,43	1,41	1,38	1,32
100	1,85	1,77	1,68	1,63	1,57	1,52	1,48	1,45	1,39	1,38	1,34	1,28
120	1,83	1,75	1,66	1,61	1,55	1,50	1,46	1,43	1,37	1,35	1,32	1,25
200	1,80	1,72	1,62	1,57	1,52	1,46	1,41	1,39	1,32	1,30	1,26	1,19
∞	1,75	1,67	1,57	1,52	1,46	1,39	1,35	1,32	1,24	1,22	1,17	1

ПРИЛОЖЕНИЕ 6

Функция Лапласа (стандартизированное нормальное распределение)

$$\Phi(u) = \frac{1}{\sqrt{2\pi}} \int_0^u e^{-\frac{t^2}{2}} dt$$

Пример: $\Phi(1{,}65) = P(0 \le U \le 1{,}65) = 0{,}4505$

u	Сотые доли									
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0200	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2703	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3883	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4779	0,4783	0,4788	0,4793	0,4798	0,4803	0,4807	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4879	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4983	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990

3,1 0,49903 **3,2** 0,49931 **3,3** 0,49952 **3,4** 0,49966 **3,5** 0,49977 **3,6** 0,49984
3,7 0,49989 **3,8** 0,49993 **3,9** 0,49995 **4,0** 0,499968 **4,5** 0,49999 **5,0** 0,49999997

ПРИЛОЖЕНИЕ 7

χ^2 - распределение

ν	α												
	0,995	**0,990**	**0,975**	**0,950**	**0,900**	**0,750**	**0,500**	**0,250**	**0,100**	**0,050**	**0,025**	**0,010**	**0,005**
1	$0{,}4\cdot10^{-6}$	$0{,}2\cdot10^{-5}$	10^{-5}	$4\cdot10^{-4}$	0,02	0,10	0,45	1,32	2,71	3,84	5,02	6,63	7,88
2	0,01	0,02	0,05	0,10	0,21	0,58	1,39	2,77	4,61	5,99	7,38	9,21	10,60
3	0,07	0,11	0,22	0,35	0,58	1,21	2,37	4,11	6,25	7,81	9,35	11,34	12,84
4	0,21	0,30	0,48	0,71	1,06	1,92	3,36	5,39	7,78	9,49	11,14	13,28	14,86
5	0,41	0,55	0,83	1,15	1,61	2,67	4,35	6,63	9,24	11,07	12,83	15,09	16,75
6	0,68	0,87	1,24	1,64	2,20	3,45	5,35	7,84	10,64	12,59	14,45	16,81	18,55
7	0,99	1,24	1,69	2,17	2,83	4,25	6,35	9,04	12,02	14,07	16,01	18,48	20,28
8	1,34	1,65	2,18	2,73	3,49	5,07	7,34	10,22	13,36	15,51	17,53	20,09	21,95
9	1,73	2,09	2,70	3,33	4,17	5,90	8,34	11,39	14,68	16,92	19,02	21,67	23,59
10	2,16	2,56	3,25	3,94	4,87	6,74	9,34	12,55	15,99	18,31	20,48	23,21	25,19
11	2,60	3,05	3,82	4,57	5,58	7,58	10,34	13,70	17,28	19,68	21,92	24,73	26,76
12	3,07	3,57	4,40	5,23	6,30	8,44	11,34	14,85	18,55	21,03	23,34	26,22	28,30
13	3,57	4,11	5,01	5,89	7,04	9,30	12,34	15,98	19,81	22,36	24,74	27,69	29,82
14	4,07	4,66	5,63	6,57	7,79	10,17	13,34	17,12	21,06	23,68	26,12	29,14	31,32
15	4,60	5,23	6,26	7,26	8,55	11,04	14,34	18,25	22,31	25,00	27,49	30,58	32,80
16	5,14	5,81	6,91	7,96	9,31	11,91	15,34	19,37	23,54	26,30	28,85	32,00	34,27
17	5,70	6,41	7,56	8,67	10,09	12,79	16,34	20,49	24,77	27,59	30,19	33,41	35,72
18	6,26	7,01	8,23	9,39	10,86	13,68	17,34	21,60	25,99	28,87	31,53	34,81	37,16
19	6,84	7,63	8,91	10,12	11,65	14,56	18,34	22,72	27,20	30,14	32,85	36,19	38,58
20	7,43	8,26	9,59	10,85	12,44	15,45	19,34	23,83	28,41	31,41	34,17	37,57	40,00
21	8,03	8,90	10,28	11,59	13,24	16,34	20,34	24,93	29,62	32,67	35,48	38,93	41,40
22	8,64	9,54	10,98	12,34	14,04	17,24	21,34	26,04	30,81	33,92	36,78	40,29	42,80
23	9,26	10,20	11,69	13,09	14,85	18,14	22,34	27,14	32,01	35,17	38,08	41,64	44,18
24	9,89	10,86	12,40	13,85	15,66	19,04	23,34	28,24	33,20	36,42	39,36	42,98	45,56
25	10,52	11,52	13,12	14,61	16,47	19,94	24,34	29,34	34,38	37,65	40,65	44,31	46,93
26	11,16	12,20	13,84	15,38	17,29	20,84	25,34	30,43	35,56	38,89	41,92	45,64	48,29
27	11,81	12,88	14,57	16,15	18,11	21,75	26,34	31,53	36,74	40,11	43,19	46,96	49,65
28	12,46	13,56	15,31	16,93	18,94	22,66	27,34	32,62	37,92	41,34	44,46	48,28	50,99
29	13,12	14,26	16,05	17,71	19,77	23,57	28,34	33,71	39,09	42,56	45,72	49,59	52,34
30	13,79	14,95	16,79	18,49	20,60	24,48	29,34	34,80	40,26	43,77	46,98	50,89	53,67
40	20,71	22,16	24,43	26,51	29,05	33,66	39,34	45,62	51,81	55,76	59,34	63,69	66,77
50	27,99	29,71	32,36	34,76	37,69	42,94	49,33	56,33	63,17	67,50	71,42	76,15	79,49
60	35,53	37,48	40,48	43,19	46,46	52,29	59,33	66,98	74,40	79,08	83,30	88,38	91,95
70	43,28	45,44	48,76	51,74	55,33	61,70	69,33	77,58	85,53	90,53	95,02	100,43	104,21
80	51,17	53,54	57,15	60,39	64,28	71,14	79,33	88,13	96,58	101,88	106,63	112,33	116,32
90	59,20	61,75	65,65	69,13	73,29	80,62	89,33	98,65	107,57	113,15	118,14	124,12	128,30
100	67,33	70,06	74,22	77,93	82,36	90,13	99,33	109,14	118,50	124,34	129,56	135,81	140,17

ПРИЛОЖЕНИЕ 8

Критические значения количества рядов для определения наличия автокорреляции по методу рядов $(\alpha = 0{,}05)$

Нижняя граница K_1

N_1	N_2																		
	2	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
2											2	2	2	2	2	2	2	2	2
3					2	2	2	2	2	2	2	2	2	3	3	3	3	3	3
4				2	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4
5			2	2	3	3	3	3	3	4	4	4	4	4	4	4	5	5	5
6		2	2	3	3	3	3	4	4	4	4	о	5	5	5	5	5	6	6
7		2	2	3	3	3	4	4	5	5	5	5	5	6	6	6	6	6	6
8		2	3	3	3	4	4	5	5	5	6	6	6	6	6	7	7	7	7
9		2	3	3	4	4	5	5	5	6	6	6	7	7	7	7	8	8	8
10		2	3	3	4	5	5	5	6	6	7	7	7	7	8	8	8	8	9
11		2	3	4	4	5	5	6	6	7	7	7	8	8	8	9	9	9	9
12	2	2	3	4	4	5	6	6	7	7	7	8	8	8	9	9	9	10	10
13	2	2	3	4	5	5	6	6	7	7	8	8	9	9	9	10	10	10	10
14	2	2	3	4	5	5	6	6	7	7	8	8	9	9	10	10	10	11	11
15	2	3	3	4	5	6	6	7	7	8	8	9	9	10	10	11	11	11	12
16	2	3	4	4	5	6	6	7	8	8	9	9	10	10	11	11	11	12	12
17	2	3	4	4	5	6	7	7	8	9	9	10	10	11	11	11	12	12	13
18	2	3	4	5	5	6	7	8	8	9	9	10	10	11	11	12	12	13	13
19	2	3	4	5	6	6	7	8	8	9	10	10	11	11	12	12	13	13	13
20	2	3	4	5	6	6	7	8	9	9	10	10	11	12	12	13	13	13	14

Верхняя граница K_2

N_1	N_2																		
	2	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
4				9	9														
5			9	10	10	11	11												
6			9	10	11	12	12	13	13	13	13								
7				11	12	13	13	14	14	14	14	15	15	15					
8				11	12	13	14	14	15	15	16	16	16	16	17	17	17	17	17
9					13	14	14	15	16	16	16	17	17	18	18	18	18	18	18
10					13	14	15	16	16	17	17	18	18	18	19	19	19	20	20
11					13	14	15	16	17	17	18	19	19	19	20	20	20	21	21
12					13	14	16	16	17	18	19	19	20	20	21	21	21	22	22
13						15	16	17	18	19	19	20	20	21	21	22	22	23	23
14						15	16	17	18	19	20	20	21	22	22	23	23	23	24
15						15	16	18	18	19	20	21	22	22	23	23	24	24	25
16							17	18	19	20	21	21	22	23	23	24	25	25	25
17							17	18	19	20	21	22	23	23	24	25	25	26	26
18							17	18	19	20	21	22	23	24	25	25	26	26	27
19							17	18	20	21	22	23	23	24	25	26	26	27	27
20							17	18	20	21	22	23	24	25	25	26	27	27	28

Пример: пусть при $N = 20$ будет 11 знаков «+» $(= N_1)$ и 9 знаков «–» $(= N_2)$. Тогда при $\alpha = 0{,}05$ нижняя граница $K_1 = 6$, верхняя граница $K_2 = 16$. Если $K_{набл} \le 6$ или $K_{набл} \ge 16$, то гипотеза об отсутствии автокорреляции должна быть отклонена.

Рассказова Наталья Владимировна

ЭКОНОМЕТРИКА. НАЧАЛЬНЫЙ КУРС

Методическое пособие
к лабораторным и контрольным работам
для студентов направления подготовки «Экономика»
заочной формы обучения

Редактор Е.Ф. Изотова

Подписано к печати 25.02.15. Формат 60х84 /16.
Усл. печ. л. 2,69. Тираж 50 экз. Заказ 151381. Рег. № 12.

Отпечатано в ИТО Рубцовского индустриального института
658207, Рубцовск, ул. Тракторная, 2/6.