

Министерство образования Российской Федерации
Воронежский государственный университет



КОМПЬЮТЕРНЫЙ ПРАКТИКУМ

ЭКОНОМЕТРИКА СЛОЖНЫХ ЭКОНОМИЧЕСКИХ ПРОЦЕССОВ

УЧЕБНОЕ ПОСОБИЕ
для студентов, обучающихся
по специальностям
060200 «Экономика труда»,
060600 «Мировая экономика»,
061800 «Математические методы
в экономике»

Воронеж – 2004

Утверждено научно-методическим советом
экономического факультета,
протокол № 1 от 29.01.2004г.

Составители: Давнис В.В.,
Тинякова В.И.,
Мокшина С.И.,
Воищева О.С.,
Щекунских С.С.

Учебное пособие подготовлено на кафедре информационных технологий и математических методов в экономике экономического факультета Воронежского государственного университета. Рекомендуется для студентов 3 курса дневного и вечернего отделений экономического факультета, а также магистров первого года обучения.

ПРЕДИСЛОВИЕ

Компьютерное моделирование экономических процессов становится не только обязательным, но и наиболее востребованным элементом подготовки современного экономиста. Данное пособие целиком посвящено этому вопросу. Главным образом оно ориентировано на формирование у студентов навыков практического выполнения достаточно сложного комплекса расчетов по построению эконометрических моделей и проведению с ними вычислительных экспериментов.

В практикум включены задания по темам продвинутого курса эконометрики, изучение которых предусматривается во втором семестре годового курса или магистерскими программами. В начале каждой темы приводится сводка необходимых для проведения расчетов формул. За сводкой формул следуют примеры решения типовых задач. Заданиями предусматривается не только проведение расчетов необходимых для построения эконометрических моделей, но и содержательная интерпретация результатов моделирования. Для проверки знаний и закрепления навыков в пособии предусмотрены задания для самостоятельной работы.

Решения задач практически по всем темам, кроме одной («Модели бинарного выбора»), выполнены в Microsoft Excel. Однако это не исключает возможность выполнения этих же заданий в любом статистическом (STATISTIKA, SPSS, SAS) или эконометрическом (EViews, STATA) пакетах.

Ориентация авторов на Excel обусловлена следующими моментами. Во-первых, это очень мощный, достаточно универсальный табличный процессор, включающий в себя надстройку «Пакет анализа» и библиотеку из множества функций. Кроме того, он является тем самым программным продуктом, в котором современный специалист проводит основную массу своих расчетов. Во-вторых, Excel предоставляет студентам возможность «прочувствовать» все детали и тонкости изучаемых методов, что естественным образом повышает уровень усвояемости учебного материала.

1. МУЛЬТИКОЛЛИНЕАРНОСТЬ

1.1. Расчетные формулы

i. 1.1.1. Ридж-оценки вектора коэффициентов регрессии

$$\hat{\mathbf{b}} = (\mathbf{X}'\mathbf{X} + a \mathbf{I})^{-1} \mathbf{X}'\mathbf{y},$$

где $a \in [\underline{a}; \bar{a}]$ (как правило, $\underline{a} = 0,1$; $\bar{a} = 0,4$).

1.1.2. Стандартная ошибка $S_{\hat{b}_k}$ k -ой ридж-оценки коэффициента регрессии, равная корню квадратному из соответствующего диагонального элемента ковариационной матрицы векторной оценки

$$S_{\hat{b}} = \hat{S} \sqrt{(\mathbf{X}'\mathbf{X} + a \mathbf{I})^{-1}},$$

где $\hat{S}^2 = \frac{\mathbf{e}'\mathbf{e}}{n - m - 1}$ рассчитывается по остаткам $\mathbf{e} = \mathbf{y} - \mathbf{X}\hat{\mathbf{b}}$.

1.2. Решение типовых задач

Задание 1.2.1. Руководство одного из крупнейших коммерческих банков ОАО «Александрит» ищет пути уменьшения расходов, связанных с основным видом деятельности – размещением вкладов физических лиц. С этой целью было решено проанализировать, в какой мере среднеквартальное число клиентов (y) определяется величиной затрат на рекламу (x_1) и суммой расходов на связи с общественностью (x_2). Данные об этих показателях за последние 20 кварталов представлены в табл. 1.2.1.

Таблица 1.2.1

t	y	x_1	x_2	t	y	x_1	x_2
1	21,324	11,670	42,012	11	114,123	63,330	227,988
2	27,765	15,000	54,000	12	123,759	68,330	245,988
3	30,954	16,670	60,012	13	138,355	76,680	276,048
4	45,125	25,000	90,000	14	138,124	76,680	276,048
5	58,123	32,220	115,992	15	141,786	78,330	281,988
6	60,473	33,330	119,988	16	135,668	75,000	270,000
7	69,965	38,330	137,988	17	146,771	81,120	292,032
8	75,456	41,670	150,012	18	149,768	82,780	298,008
9	92,667	51,110	183,996	19	158,234	87,780	316,008
10	105,457	58,330	209,988	20	167,463	92,770	333,972

Решение с помощью табличного процессора Excel

1. Ввод исходных данных.
2. Построение регрессионного уравнения с использованием «Пакета анализа» (рис. 1.2.1).

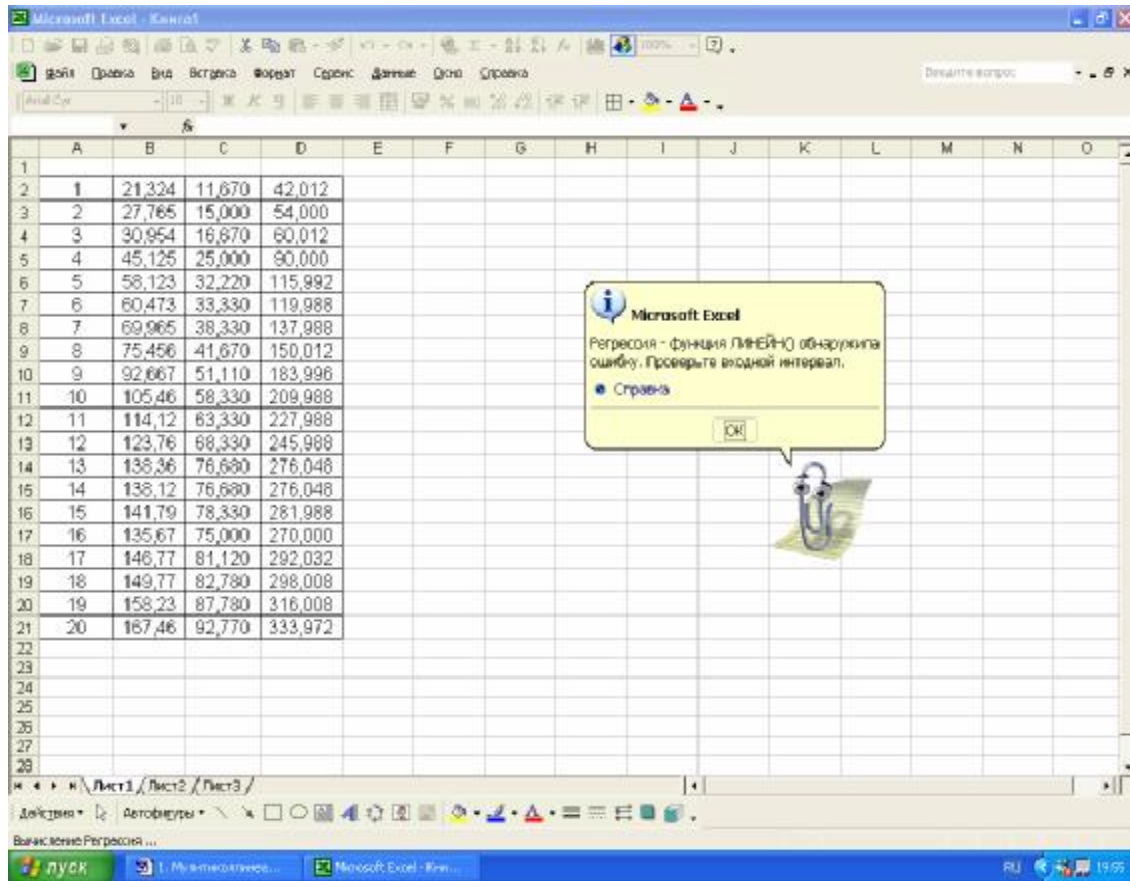


Рис. 1.2.1.

3. Полученный результат позволяет выдвинуть гипотезу о наличии эффекта мультиколлинеарности.
 - 3.1. Формирование матрицы $(X'X)$ с помощью функций **ТРАНСП** и **МУМНОЖ**

74577,79	268480
268480	966528,2

- 3.2. Вычисление определителя матрицы $(X'X)$ с помощью функции **МОПР** и по формуле

$$\Delta = 74577,79 \cdot 966528,2 - 268480 \cdot 268480 = 0.$$

Равенство определителя нулю говорит о наличии явления мультиколлинеарности в строгом смысле. Следовательно, необходимо исключить один из факторов и перестроить модель заново.

4. Построение регрессионной модели с единственным фактором – расходами на связи с общественностью (см. Вывод итогов 1.1).

ВЫВОД ИТОГОВ 1.1

<i>Регрессионная статистика</i>	
Множественный R	0,999982
R-квадрат	0,999964
Нормированный R-квадрат	0,999962
Стандартная ошибка	0,293376
Наблюдения	20

<i>Дисперсионный анализ</i>					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>
Регрессия	1	43402,52	43402,52	504271,3	1,74E-41
Остаток	18	1,549256	0,08607		
Итого	19	43404,07			

	<i>Коэффициенты</i>	<i>Стандартная ошибка</i>	<i>t-статистика</i>	<i>P-значение</i>	<i>Нижние 95%</i>	<i>Верхние 95%</i>
Y-пересечение	0,537742	0,154752	3,47486	0,002703	0,21262	0,862864
Переменная X 1	0,499892	0,000704	710,1206	1,74E-41	0,498413	0,501371

Таким образом, построенная модель имеет вид

$$\hat{y}_t = 0,5377 + 0,4999 x_{1t}.$$

Высокий коэффициент корреляции свидетельствует о существенной взаимосвязи моделируемого показателя с фактором. Сравнение расчетного значения F -критерия с табличным $F_{0,95}(1, 19) = 4,38$ позволяет сделать вывод об адекватности построенной модели. Сравнение расчетных значений t -статистик с табличным $t_{0,95}(19) = 2,093$ говорит о значимости включенного в модель фактора x_1 .

Задание 1.2.2. Известно, что стоимость выпуска газеты в значительной степени определяется величиной типографских расходов. Для того,

чтобы иметь возможность воздействовать на эту стоимость, издатели наиболее популярных газет решили изучить факторы, определяющие сумму годовых затрат на печать газет, и оценить степень их влияния. С этой целью для 20 городов России были собраны данные о годовых расходах на печать (y , млн. руб.), объемах розничной продажи газеты в городе (млн. руб.) и количества семей в городе. Заметим, что для факторов были взяты их логарифмы (x_1 и x_2 , соответственно) с целью уменьшения разброса данных, а следовательно, и упрощения их обработки. Все эти данные представлены в табл. 1.2.2. Постройте модель множественной регрессии, отражающую зависимость среднегодовых расходов на издание газеты от соответствующих факторов.

Таблица 1.2.2

№ п.п.	y	x_1	x_2	№ п.п.	y	x_1	x_2
1.	21,452	2,1371	1,0682	11.	20,055	2,1473	1,0746
2.	16,935	1,8620	0,9315	12.	21,535	2,3686	1,1847
3.	18,150	1,9946	0,9974	13.	24,352	2,6759	1,3378
4.	19,950	2,1912	1,0968	14.	19,550	2,1483	1,0747
5.	19,350	2,1265	1,0639	15.	14,150	1,5547	0,7754
6.	17,350	1,9065	0,9531	16.	14,550	1,5982	0,7998
7.	25,150	2,7656	1,3819	17.	16,550	1,8184	0,9099
8.	15,371	1,6863	0,8438	18.	16,950	1,8624	0,9317
9.	13,957	1,5329	0,7662	19.	26,381	2,8951	1,4472
10.	27,150	2,9837	1,4914	20.	16,580	1,8189	0,9098

Решение с помощью табличного процессора Excel

1. Ввод исходных данных.
2. Построение регрессионного уравнения с использованием «Пакета анализа» (см. Вывод итогов 1.2) и анализ полученных результатов.

Анализ вывода итогов 1.2 свидетельствует о том, что полученные стандартные ошибки значительно больше самих расчетных коэффициентов. Коэффициенты при факторах x_1 и x_2 незначимы, так как для них Р-значение больше 0,05. В то же время сравнение расчетного значения F - критерия с табличным $F_{0,95}(2, 17) = 3,59$ позволяет сделать вывод об адекватности рассматриваемой модели. Коэффициенты корреляции (см. табл. 1.2.3) говорят о существенной взаимосвязи модели-

руемого показателя с факторами. Одной из причин противоречивости результатов модели является тесная взаимосвязь между факторами. Все эти факты говорят о том, что изучаемая модель требует более детального анализа.

Таблица 1.2.3

	y	x_1	x_2
y	1	0,993536978	0,993500936
x_1	0,993536978	1	0,999994612
x_2	0,993500936	0,999994612	1

ВЫВОД ИТОГОВ 1.2

Регрессионная статистика	
Множественный R	0,99358
R-квадрат	0,98720
Нормированный R-квадрат	0,98570
Стандартная ошибка	0,48029
Наблюдения	20

Дисперсионный анализ					
	df	SS	MS	F	Значимость F
Регрессия	2	302,52769	151,26385	655,72475	0,00000
Остаток	17	3,92159	0,23068		
Итого	19	306,44928			

	Коэффициенты	Стандартная ошибка	t-статистика	P-значение	Нижние 95%	Верхние 95%
Y-пересечение	0,12398	0,55176	0,22469	0,82490	-1,04014	1,28810
Переменная X 1	35,26429	76,72580	0,45961	0,65162	-126,61323	197,14181
Переменная X 2	-52,31648	153,53595	-0,34074	0,73747	-376,24947	271,61650

3. Проверка условия невырожденности матрицы $(X'X)$.

3.1. Формирование матрицы $(X'X)$ с помощью функций ТРАНСП и МУМНОЖ.

92,14842	46,07888
46,07888	23,04179

3.2. Вычисление определителя матрицы $(X'X)$ с помощью функции МОПР.

$$\Delta = 0,000939286.$$

Близость определителя к нулю, а также проведенный выше анализ позволяют сделать вывод о наличии частичной мультиколлинеарности.

4. Устранение эффекта мультиколлинеарности с помощью ридж-оценивания.

4.1. Включение в модель дополнительной переменной x_0 , принимающей единственное значение, равное 1.

4.2. Расчет коэффициентов регрессии с использованием матричных функций Excel.

4.2.1. Формирование матрицы, обратной к матрице системы нормальных уравнений $(X'X + aI)^{-1}$ с помощью функций **ТРАНСП**, **МУМНОЖ** и **МОБР** при $a \in [0,1; 0,4]$.

При $a = 0,1$

1,268975	0,124448	-1,40753
0,124448	10,05558	-20,1794
-1,40753	-20,1794	41,5964

При $a = 0,2$

1,268965	0,117394	-1,39341
0,117394	5,055277	-10,1733
-1,39341	-10,1733	21,57345

При $a = 0,3$

1,268962	0,115042	-1,38871
0,115042	3,388073	-6,83709
-1,38871	-6,83709	14,89738

При $a = 0,35$

1,268961	0,11437	-1,38736
0,11437	2,911689	-5,88381
-1,38736	-5,88381	12,98978

При $a = 0,4$

1,26896	0,113866	-1,38635
0,113866	2,554389	-5,16882
-1,38635	-5,16882	11,55902

4.2.2. Получение вектора оценок коэффициентов регрессии путем умножения обратной матрицы на матрицы $(X'X + aI)^{-1}$ и $X'y$ при различных значениях a . Оформление результатов в виде табл. 1.2.4.

Таблица 1.2.4

	$a = 0,1$	$a = 0,2$	$a = 0,3$	$a = 0,35$	$a = 0,4$
\hat{b}_0	0,100097	0,100092	0,100091	0,10009	0,10009
\hat{b}_1	18,33722	18,3339	18,33279	18,33248	18,33224
\hat{b}_2	-18,4439	-18,4373	-18,4351	-18,4345	-18,4340

5. Расчет стандартных ошибок коэффициентов регрессии.

5.1. Вычисление остаточной дисперсии при различных значениях a и оформление результатов расчетов в виде табл. 1.2.5.

Таблица 1.2.5

<i>Квадраты отклонений расчетных от фактических значений</i>				
$a = 0,1$	$a = 0,2$	$a = 0,3$	$a = 0,35$	$a = 0,4$
3,4792082	3,4792200	3,4792239	3,4792250	3,4792258
0,0165009	0,0165014	0,0165016	0,0165017	0,0165017
0,0167740	0,0167739	0,0167739	0,0167739	0,0167739
0,0102582	0,0102597	0,0102602	0,0102603	0,0102604
0,0148038	0,0148047	0,0148049	0,0148050	0,0148051
0,0171801	0,0171795	0,0171793	0,0171792	0,0171792
0,0309093	0,0309073	0,0309066	0,0309064	0,0309063
0,0077689	0,0077694	0,0077695	0,0077696	0,0077696
0,0145113	0,0145104	0,0145101	0,0145100	0,0145100
0,0241942	0,0241935	0,0241933	0,0241932	0,0241931
0,1594127	0,1594083	0,1594069	0,1594065	0,1594061
0,0219283	0,0219290	0,0219292	0,0219292	0,0219293
0,0202607	0,0202605	0,0202604	0,0202604	0,0202603
0,0149396	0,0149403	0,0149405	0,0149406	0,0149406
0,0248156	0,0248109	0,0248094	0,0248089	0,0248086
0,0110599	0,0110605	0,0110607	0,0110608	0,0110608
0,0126209	0,0126216	0,0126218	0,0126219	0,0126220
0,0137129	0,0137134	0,0137135	0,0137136	0,0137136
0,0132457	0,0132453	0,0132452	0,0132451	0,0132451
0,0087153	0,0087154	0,0087155	0,0087155	0,0087155
<i>Сумма квадратов отклонений расчетных от фактических значений</i>				
3,9328205	3,9328249	3,9328264	3,9328268	3,9328271
<i>Остаточная дисперсия</i>				
0,23134238	0,23134264	0,23134273	0,23134275	0,23134277

5.2. Получение стандартных ошибок в виде корня квадратного из произведения диагональных элементов обратной матрицы на остаточную дисперсию при различных значениях a . Оформление результатов расчетов в виде табл. 1.2.6.

Таблица 1.2.6

	$a = 0,1$	$a = 0,2$	$a = 0,3$	$a = 0,35$	$a = 0,4$
$S_{\hat{b}_0}$	0,541819	0,541817	0,541817	0,541816	0,541816
$S_{\hat{b}_1}$	1,525216	1,081435	0,885328	0,82073	0,768726
$S_{\hat{b}_2}$	3,102098	2,234023	1,856449	1,73352	1,635267

Таким образом, наименьшая стандартная ошибка получена при $a = 0,4$. Следовательно, построенная модель может быть записана в виде

$$y = 0,10009 + 18,33224 x_1 - 18,434 x_2.$$

6. Проверка значимости полученных коэффициентов регрессии

$$t_{\hat{b}_0} = \frac{\hat{b}_0}{S_{\hat{b}_0}} = \frac{0,10009}{0,541816} = 0,18473; \quad t_{\hat{b}_1} = \frac{\hat{b}_1}{S_{\hat{b}_1}} = \frac{18,33224}{0,768726} = 23,8475;$$

$$t_{\hat{b}_2} = \frac{\hat{b}_2}{S_{\hat{b}_2}} = \frac{-18,43400}{1,635267} = -11,2728.$$

Сравнение расчетных значений t -статистик с табличным $t_{0,95}(17) = 2,110$ свидетельствует о значимости включенных в модель факторов x_1 и x_2 . Таким образом, построенную модель можно использовать для целей анализа и прогнозирования.

1.3. Задания для самостоятельной работы

Задание 1.3.1. Владельцы интернет-аукциона «Э-Слава» составляют бизнес-план своей деятельности на следующие два месяца. Их прежде всего интересует вопрос, каким образом можно увеличить объем реализации в натуральном выражении. В ходе исследования было выявлено, что на количество совершаемых в среднем за месяц покупок (y) влияют такие факторы, как затраты на баннерную рекламу (тыс. руб., x_1), расходы на мероприятия, осуществляемые с целью привлечения интернет-пользователей на сайт аукциона (тыс. руб., x_2), и число зарегистрированных пользователей сайта (x_3). Поэтому было решено построить модель множественной регрессии

рессии, отражающую зависимость количества покупок от указанных факторов. Данные об этих показателях за последние 20 месяцев представлены в табл. 1.3.1.

Таблица 1.3.1

t	y	x_1	x_2	x_3	t	y	x_1	x_2	x_3
1	535	8,39	30,31	985	11	561	6,58	23,78	772
2	515	6,83	24,68	802	12	388	4,48	16,17	525
3	382	5,54	20,00	650	13	630	7,41	26,76	869
4	721	8,47	30,59	994	14	769	8,55	30,88	1003
5	276	6,13	22,13	719	15	470	5,52	19,95	648
6	513	5,77	20,85	677	16	511	6,01	21,7	705
7	664	7,80	28,18	915	17	549	6,19	22,36	726
8	409	4,80	17,35	563	18	531	8,85	31,96	1038
9	537	5,42	19,57	636	19	499	7,21	26,05	846
10	794	9,31	33,62	1092	20	503	5,99	21,65	703

Задание 1.3.2. Фирма «Ваше очарование» довольно успешно осуществляет торговлю косметическими товарами на российском рынке. Ее успех определяется, в частности, человеческим фактором. С целью изучения его влияния на среднеквартальный объем продаж (млн. руб., y) через такие показатели, как фонд оплаты труда (млн. руб., x_1) и численность работников фирмы (чел., x_2), была сформирована табл. 1.3.2. В этой таблице приведены данные по этим показателям за последние 18 кварталов. Постройте двухфакторную регрессионную модель, отражающую зависимость объема продаж от указанных факторов.

Таблица 1.3.2

t	y	x_1	x_2	t	y	x_1	x_2
1	39832	3975	1986	10	43671	4344	2171
2	33527	3337	1668	11	39518	3895	1946
3	36181	3586	1792	12	47639	4941	2469
4	42873	4065	2031	13	39518	3895	1946
5	28279	2859	1429	14	29376	2919	1458
6	52256	5001	2499	15	28215	2809	1404
7	36425	3611	1805	16	33193	3298	1648
8	27734	2760	1379	17	33927	3377	1688
9	54547	5419	2708	18	57736	5250	2624

2. АВТОКОРРЕЛИРОВАННОСТЬ ОСТАТКОВ

2.1. Расчетные формулы

2.1.1. Критерий Дарбина – Уотсона

$$\bar{d} = \frac{\hat{\mathbf{e}}' \mathbf{A} \hat{\mathbf{e}}}{\hat{\mathbf{e}}' \hat{\mathbf{e}}},$$

где
$$\mathbf{A} = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ -1 & 2 & -1 & \mathbf{L} & 0 \\ 0 & -1 & 2 & \mathbf{L} & 0 \\ \mathbf{M} & \mathbf{M} & \mathbf{M} & \mathbf{M} & \mathbf{M} \\ 0 & 0 & 0 & \mathbf{L} & 0 \end{pmatrix}.$$

2.1.2. Коэффициент автокорреляции

$$\hat{r} = \frac{\sum \hat{e}_t \hat{e}_{t-1}}{\sum \hat{e}_{t-1}^2}.$$

2.1.3. Преобразование исходных данных для устранения автокорреляции в остатках

$$\mathbf{y}^* = \hat{\mathbf{P}} \mathbf{y}, \quad \mathbf{X}^* = \hat{\mathbf{P}} \mathbf{X},$$

где $\hat{\mathbf{P}}$ такое, что $\hat{\mathbf{P}}' \hat{\mathbf{P}} = \hat{\Sigma}_0^{-1}$. Матрица $\hat{\mathbf{P}}$ представляет собой корень квадратный из матрицы, обратной к ковариационной матрице остатков $\hat{\Sigma}_0^{-1}$, и имеет вид

$$\hat{\mathbf{P}} = \begin{pmatrix} \sqrt{1 - \hat{r}^2} & 0 & 0 & \mathbf{L} & 0 & 0 \\ -\hat{r} & 1 & 0 & \mathbf{L} & 0 & 0 \\ 0 & -\hat{r} & 1 & \mathbf{L} & 0 & 0 \\ \mathbf{M} & \mathbf{M} & \mathbf{M} & \mathbf{L} & \mathbf{M} & \mathbf{M} \\ 0 & 0 & 0 & \mathbf{L} & -\hat{r} & 1 \end{pmatrix}.$$

2.2. Решение типовой задачи

Задание 2.2.1. Руководству крупной компании «ПластКо», производящей различные товары из пластмассы, в том числе и чайные кружки, с помощью конфиденциальных источников удалось получить информацию о

некоторых показателях работы конкурирующей фирмы. Часть этой информации представлена в табл. 2.2.1.

Таблица 2.2.1

t	y	x_1	x_2	t	y	x_1	x_2
1	42,08	14,53	16,74	11	52,47	20,77	19,33
2	41,49	15,30	16,81	12	50,68	21,17	17,04
3	39,06	15,92	19,50	13	51,64	21,34	16,74
4	45,09	17,41	22,12	14	56,19	22,91	19,81
5	51,67	18,37	22,34	15	66,22	22,96	31,92
6	51,18	18,83	17,47	16	63,23	23,69	26,31
7	54,78	18,84	20,24	17	68,96	24,82	25,93
8	60,33	19,71	20,37	18	64,26	25,54	21,96
9	49,76	20,01	12,71	19	63,75	25,63	24,05
10	55,46	20,26	22,98	20	69,68	28,73	25,66

В этой таблице через t обозначено время (месяц), y – стоимость одной пластмассовой кружки для чая (руб.), x_1 – стоимость материальных ресурсов, израсходованных на производство одной кружки (руб.), x_2 – стоимость рабочей силы (зарботной платы), затраченной на производство одной кружки (руб.). Кроме того, известно, что эта фирма-конкурент ведет работу по снижению себестоимости выпускаемой продукции, и величины ее затрат на материальные и человеческие ресурсы, необходимые для производства одной кружки, в следующем месяце, скорее всего, составят по 20 руб. Совет директоров компании решил поручить своему консультанту по экономическим вопросам построить модель, отражающую зависимость стоимости кружки от соответствующих факторов, с целью получения возможности прогнозировать цены, по которым конкурирующая фирма сможет отпускать кружки оптовым покупателям.

Решение с помощью табличного процессора Excel

1. Ввод исходных данных с включением в модель дополнительной переменной x_0 , принимающей единственное значение, равное 1.
2. Нахождение параметров регрессии МНК с использованием матричных функций Excel.

2.1. Оценка вектора коэффициентов регрессии $\hat{\mathbf{b}} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y}$ с помощью функций **ТРАНСП**, **МУМНОЖ** и **МОБР**.

2.1.1. Формирование обратной матрицы к матрице системы нормальных уравнений $(\mathbf{X}'\mathbf{X})^{-1}$

1,9071	-0,0603	-0,0286
-0,0603	0,0055	-0,0026
-0,0286	-0,0026	0,0039

2.1.2. Получение вектора оценок коэффициентов регрессии

3,8470
1,8108
0,6343

2.2. Расчет остатков $\hat{\mathbf{e}} = \mathbf{y} - \mathbf{X}\hat{\mathbf{b}} = \mathbf{y} - \hat{\mathbf{y}}$.

2.2.1. Нахождение вектора расчетных значений $\hat{\mathbf{y}} = \mathbf{X}\hat{\mathbf{b}}$ с помощью функции **МУМНОЖ**.

2.2.2. Вычисление разностей $\hat{\mathbf{e}} = \mathbf{y} - \hat{\mathbf{y}}$ и оформление промежуточных результатов в виде табл. 2.2.2.

Таблица 2.2.2

t	y	\hat{y}	\hat{e}	t	y	\hat{y}	\hat{e}
1	42,08	40,77	1,31	11	52,47	53,72	-1,25
2	41,49	42,21	-0,73	12	50,68	52,99	-2,31
3	39,06	45,04	-5,99	13	51,64	53,11	-1,46
4	45,09	49,40	-4,31	14	56,19	57,90	-1,71
5	51,67	51,28	0,39	15	66,22	65,67	0,55
6	51,18	49,02	2,16	16	63,23	63,43	-0,21
7	54,78	50,80	3,98	17	68,96	65,24	3,73
8	60,33	52,46	7,88	18	64,26	64,02	0,24
9	49,76	48,14	1,61	19	63,75	65,51	-1,76
10	55,46	55,11	0,35	20	69,68	72,15	-2,46

2.3. Нахождение оценки ковариационной матрицы Σ для вектора $\hat{\mathbf{b}}$ по формуле $\hat{\Sigma} = \hat{\mathcal{S}}^2 (\mathbf{X}'\mathbf{X})^{-1}$.

2.3.1. Нахождение произведения векторов $\hat{\mathbf{e}}'\mathbf{e}$ с помощью функций **ТРАНСП** и **МУМНОЖ**

$$\hat{\mathbf{e}}'\mathbf{e} = 177,3741.$$

2.3.2. Вычисление величины $\hat{\mathcal{S}}^2 = (\hat{\mathbf{e}}'\mathbf{e}) / (n - m - 1)$

$$\hat{\mathcal{S}}^2 = 177,3741 / 17 = 10,4348.$$

2.3.3. Нахождение $\hat{\Sigma}$ как произведение $\hat{\mathcal{S}}^2$ и обратной матрицы $(\mathbf{X}'\mathbf{X})^{-1}$

19,8979	-0,6292	-0,2983
-0,6292	0,0574	-0,0270
-0,2983	-0,0270	0,0410

3. Проверка гипотезы о наличии автокорреляции в остатках с использованием критерия Дарбина – Уотсона.

3.1. Вычисление фактического значения критерия Дарбина – Уотсона с использованием матричных операций Excel

$$\bar{d} = \frac{\hat{\mathbf{e}}' \mathbf{A} \hat{\mathbf{e}}}{\hat{\mathbf{e}}' \hat{\mathbf{e}}} = \frac{165,78}{177,37} = 0,9346 ,$$

3.2. Формулирование гипотез:

H_0 – в остатках нет автокорреляции;

H_1 – в остатках есть положительная автокорреляция;

H_1^* – в остатках есть отрицательная автокорреляция.

3.3. Определение по таблицам значений критерия Дарбина – Уотсона критических значений $d_L^* = 1,100$ и $d_U^* = 1,537$ для $n = 20$, $m = 3$ при вероятности ошибки $\alpha = 0,05$.

3.4. Сравнение фактического и критических значений критерия $\bar{d} < d_L^*$ свидетельствует о том, что с вероятностью 95% можно отклонить нуль-гипотезу и сделать вывод о наличии положительной автокорреляции.

Наличие автокорреляции означает, что $e_t = r e_{t-1} + d_t$, т.е. не выполняются предположения классического регрессионного анализа, и, следовательно, можно найти более эффективную оценку, чем $\hat{\mathbf{b}}$. Кроме того, вычисленная в п.2.3 ковариационная матрица $\hat{\Sigma}$ является смещенной оценкой ковариационной матрицы $\hat{\mathbf{b}}$, и ее нельзя использовать для получения стандартных ошибок оценок коэффициентов регрессии.

4. Преобразование исходных данных с целью устранению автокорреляции в остатках.

4.1. Оценка параметра r авторегрессии первого порядка

$$e_t = r e_{t-1} + d_t .$$

4.1.1. Вычисление $\mathbf{e}_t \mathbf{e}_{t-1}$ и \mathbf{e}_t^2 . Оформление результатов расчетов в виде табл. 2.2.3

Таблица 2.2.3

t	e_t	$e_t e_{t-1}$	e_t^2	t	e_t	$e_t e_{t-1}$	e_t^2
1	1,31		1,72	11	-1,25	-0,44	1,56
2	-0,73	-0,95	0,53	12	-2,31	2,89	5,35
3	-5,99	4,35	35,83	13	-1,46	3,39	2,14
4	-4,31	25,81	18,60	14	-1,71	2,50	2,92
5	0,39	-1,68	0,15	15	0,55	-0,94	0,30
6	2,16	0,84	4,67	16	-0,21	-0,11	0,04
7	3,98	8,59	15,83	17	3,73	-0,77	13,89
8	7,88	31,34	62,05	18	0,24	0,88	0,06
9	1,61	12,71	2,60	19	-1,76	-0,41	3,09
10	0,35	0,57	0,12	20	-2,46	4,33	6,07
				Сумма:		92,91	177,37

4.1.2. Вычисление коэффициента автокорреляции

$$\hat{r} = \frac{\sum \hat{e}_t \hat{e}_{t-1}}{\sum \hat{e}_{t-1}^2} = 0,54.$$

4.2. Преобразование исходных данных по формулам

$$\mathbf{y}^* = \hat{\mathbf{P}}\mathbf{y}, \quad \mathbf{x}_0^* = \hat{\mathbf{P}}\mathbf{x}_0, \quad \mathbf{x}_1^* = \hat{\mathbf{P}}\mathbf{x}_1 \text{ и } \mathbf{x}_2^* = \hat{\mathbf{P}}\mathbf{x}_2.$$

$$\mathbf{y}: y_1^* = (1 - \hat{r}^2)^{1/2} y_1 = (1 - 0,29)^{1/2} 42,08 = 35,37,$$

$$y_2^* = y_2 - \hat{r} y_1 = 41,49 - 0,54 \cdot 42,08 = 18,68,$$

$$y_3^* = y_3 - \hat{r} y_2 = 39,06 - 0,54 \cdot 41,49 = 16,58.$$

.....

$$\mathbf{x}_0: x_{01}^* = (1 - \hat{r}^2)^{1/2} x_{01} = (1 - 0,29)^{1/2} 1,00 = 0,84,$$

$$x_{02}^* = x_{02} - \hat{r} x_{01} = 1,00 - 0,54 \cdot 1,00 = 0,46,$$

$$x_{03}^* = x_{03} - \hat{r} x_{02} = 1,00 - 0,54 \cdot 1,00 = 0,46.$$

.....

$$\mathbf{x}_1: x_{11}^* = (1 - \hat{r}^2)^{1/2} x_{11} = (1 - 0,29)^{1/2} 14,53 = 12,21,$$

$$x_{12}^* = x_{12} - \hat{r} x_{11} = 15,30 - 0,54 \cdot 14,53 = 7,47,$$

$$x_{13}^* = x_{13} - \hat{r} x_{12} = 15,92 - 0,54 \cdot 15,30 = 7,63.$$

.....

$$\mathbf{x}_2: x_{21}^* = (1 - \hat{r}^2)^{1/2} x_{21} = (1 - 0,29)^{1/2} 16,74 = 14,07,$$

$$x_{22}^* = x_{22} - \hat{r} x_{21} = 16,81 - 0,54 \cdot 16,74 = 7,74,$$

$$x_{23}^* = x_{23} - \hat{r} x_{22} = 19,50 - 0,54 \cdot 16,81 = 10,39.$$

Оформление результатов расчетов в виде табл. 2.2.4

Таблица 2.2.4

t	y	x_0	x_1	x_2	t	y	x_0	x_1	x_2
1	35,37	0,84	12,21	14,07	11	22,41	0,46	9,79	6,88
2	18,68	0,46	7,43	7,74	12	22,24	0,46	9,91	6,57
3	16,58	0,46	7,63	10,39	13	24,18	0,46	9,87	7,51
4	23,93	0,46	8,78	11,55	14	28,20	0,46	11,35	10,74
5	27,24	0,46	8,94	10,35	15	35,77	0,46	10,55	21,19
6	23,18	0,46	8,88	5,36	16	27,35	0,46	11,25	9,01
7	27,04	0,46	8,64	10,77	17	34,70	0,46	11,98	11,67
8	30,65	0,46	9,50	9,40	18	26,89	0,46	12,09	7,91
9	17,06	0,46	9,33	1,67	19	28,93	0,46	11,79	12,15
10	28,50	0,46	9,42	16,09	20	35,14	0,46	14,84	12,63

5. Нахождение параметров регрессии $\mathbf{y}^* = \mathbf{X}^* \mathbf{b} + \mathbf{e}^*$ с помощью обычного МНК, что эквивалентно применению обобщенного МНК к исходным данным с использованием матричных функций Excel (аналогично п. 2).

5.1. Оценка вектора коэффициентов регрессии $\hat{\mathbf{b}} = (\mathbf{X}^{*'} \mathbf{X}^*)^{-1} \mathbf{X}^{*'} \mathbf{y}$.

4,0352
1,6725
0,7593

5.2. Нахождение ковариационной матрицы $\hat{\Sigma} = \hat{\mathbf{S}}^2 (\mathbf{X}^{*'} \mathbf{X}^*)^{-1}$ для $\hat{\mathbf{b}}$

37,57	-1,51	-0,21
-1,51	0,09	-0,01
-0,21	-0,01	0,02

6. Получение прогнозного значения \mathbf{y} , задаваемого $\mathbf{x}_{T+1}' = (1, 20, 20)$.

6.1. Нахождение прогнозной оценки обобщенным МНК, игнорируя тот факт, что e_{T+1} коррелированно с предыдущим значением в выборочном периоде

$$\hat{y}_{T+1}^* = \mathbf{x}_{T+1}' \hat{\mathbf{b}} = 58,69.$$

6.2. Нахождение прогнозной оценки обобщенным МНК с учетом того, что e_{T+1} коррелированно с предыдущим значением в выборочном периоде

$$\hat{y}_{T+1} = \mathbf{x}_{T+1}' \hat{\mathbf{b}} + \hat{r} \left(y_T - \mathbf{x}_T' \hat{\mathbf{b}} \right) = 52,69 + 0,54(69,68 - 71,58) = 51,66.$$

2.3. Задания для самостоятельной работы

Задание 2.3.1. ОАО «Молоко Хавы» производит молочную продукцию, которую реализует не только в Воронежской области, но и за ее пределами, в частности в Липецкой области. Однако если рынок молочной продукции Воронежской области практически освоен и объем продаж на нем относительно постоянен, то ситуация в Липецкой области нуждается в более тщательном анализе, поскольку поставки в эту область начались совсем недавно. С целью проведения такого анализа были собраны данные о суточной выработке молока и ежедневных объемах продаж в Липецкой области за последний месяц. Эти данные представлены в табл. 2.3.1.

Таблица 2.3.1

День	Объем продаж молока, л.	Выработка молока, л.	День	Объем продаж молока, л.	Выработка молока, л.
1	414	4039	16	4313	7522
2	782	4358	17	4279	7401
3	1096	4660	18	4430	7425
4	1388	4818	19	4420	7437
5	1673	5161	20	4623	7834
6	1880	5392	21	5303	8272
7	2149	5497	22	5532	8670
8	2385	5677	23	5826	8991
9	2688	6080	24	5819	9223
10	3287	6407	25	6164	9558
11	3106	6257	26	6287	9682
12	3180	6366	27	6428	9706
13	3466	6776	28	6261	9544
14	3680	7080	29	6294	9435
15	4125	7399	30	6434	9567

Требуется: 1) построить уравнение регрессии, отражающее зависимость ежедневных объемов продаж в Липецкой области от суточной выработки молока; 2) для построенного уравнения вычислить остатки и проверить гипотезу о наличии автокорреляции в остатках с помощью критерия Дарбина – Уотсона; 3) в случае подтверждения гипотезы устранить автокоррелированность остатков путем соответствующего преобразования данных; 4) получить с помощью построенной модели прогнозную оценку объемов продаж молока в Липецкой области при планируемой суточной выработке молока 9650 литров.

Задание 2.3.2. В табл. 2.3.2 представлены данные об общих объемах выпускаемой продукции ОАО ХК «Мебель Черноземья» в период с 2000 по 2003 г., а также об объемах продаж в филиалах ХК.

Таблица 2.3.2

Месяц	Объемы продаж в филиалах (млн. руб.)	Общий объем выпускаемой продукции (млн. руб.)	Месяц	Объемы продаж в филиалах (млн. руб.)	Общий объем выпускаемой продукции (млн. руб.)
2000			2002		
1	5	17	1	30	61
2	6	14	2	31	62
3	12	27	3	30	62
4	7	21	4	29	54
5	9	24	5	28	56
6	9	23	6	29	61
7	12	35	7	33	72
8	13	37	8	39	83
9	13	36	9	41	92
10	13	42	10	42	90
11	18	44	11	43	98
12	33	70	12	49	102
2001			2003		
1	20	45	1	41	69
2	20	46	2	31	57
3	22	49	3	29	58
4	19	37	4	34	63
5	14	36	5	32	57
6	21	44	6	31	54
7	31	59	7	38	73
8	25	62	8	45	84
9	26	62	9	45	85
10	31	75	10	46	89
11	28	66	11	44	87
12	35	74	12	59	109

Предполагая существование линейной зависимости объемов продаж в филиалах от общего объема выпускаемой продукции, постройте обычным МНК линейное уравнение регрессии. Для построенного уравнения вычислите остатки и, используя критерий Дарбина – Уотсона, проверьте гипотезу о наличии автокорреляции в остатках. При наличии автокорреляции в остатках оцените параметры регрессии обобщенным МНК. Убедитесь в устранении автокоррелированности вновь полученных остатков, для чего снова примените критерий Дарбина–Уотсона. Рассчитайте прогнозное зна-

чение объемов продаж в филиалах при общем объеме выпуска продукции 110 млн. руб.

Задание 2.3.3.

Данные, представленные в табл. 2.3.3, отражают доходы населения (x) и расходы на покупку валюты (y) за период с 2001 по 2003г. Очевидно, что существует взаимосвязь между этими двумя показателями. Оцените степень этой взаимосвязи, построив соответствующее линейное уравнение регрессии. Для построенного уравнения вычислите остатки и, используя критерий Дарбина – Уотсона в матричной форме, проверьте гипотезу о наличии автокорреляции в остатках. В случае подтверждения этой гипотезы оцените параметры регрессии обобщенным МНК.

Таблица 2.3.3

t	x	y	t	x	y	t	x	y
2001			2002			2003		
1	166,2	14,28	1	215,0	15,48	1	290,2	21,25
2	186,0	15,66	2	261,3	19,92	2	337,5	23,42
3	197,9	16,49	3	286,5	21,15	3	376,1	25,26
4	220,5	17,72	4	291,5	21,16	4	395,4	26,13
5	212,5	17,46	5	284,5	21,03	5	372,1	25,09
6	226,5	18,18	6	315,1	22,45	6	428,2	27,67
7	226,6	18,23	7	308,1	22,12	7	424,9	27,57
8	239,1	18,84	8	322,7	23,00	8	437,2	28,36
9	239,8	18,98	9	331,5	23,16	9	436,1	28,29
10	250,8	19,56	10	325,5	22,95	10	438,6	28,32
11	257,0	19,46	11	348,5	23,96	11	448,3	28,9
12	354,9	23,25	12	452,3	28,85	12	580,6	35,45

3. ARCH МОДЕЛИ

3.1. Расчетные формулы

3.1.1. ARCH модель в простейшем случае

$$y_t = x_t b + e_t,$$

$$\text{где } e_t = u_t (a_0 + a_1 e_{t-1})^{\frac{1}{2}}, \quad u_t \sim N(0,1).$$

3.1.2. Четырехэтапная процедура построения **ARCH** модели:

1) оценка параметров уравнения регрессии обычным МНК

$$\hat{\mathbf{b}} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y} \text{ и вычисление остатков } \mathbf{e} = \mathbf{y} - \mathbf{X}\hat{\mathbf{b}};$$

2) построение с помощью МНК зависимости

$$s_t^2 = a_0 + a_1 e_{t-1}^2,$$

тестирование ARCH-эффектов по критерию TR^2 ;

3) вычисление расчетных значений дисперсии остатков

$$h_t = a_0 + a_1 e_{t-1}^2,$$

формирование

$$g_t = (e_t^2 / h_t - 1), \quad z_{t1} = 1/h_t, \quad z_{t2} = e_{t-1}^2 / h_t,$$

ввод обозначений

$$\mathbf{g} = [g_t]_2^T, \quad \mathbf{Z} = [z_{t1}, z_{t2}]_2^T,$$

вычисление поправочных коэффициентов

$$\mathbf{d}_a = (\mathbf{Z}'\mathbf{Z})^{-1} \mathbf{Z}'\mathbf{g},$$

корректировка коэффициентов

$$\hat{\mathbf{a}} = \mathbf{a} + \mathbf{d}_a, \quad \text{где } \mathbf{a} = (a_0, a_1)';$$

4) пересчет h_t по формуле $h_t = \hat{a}_0 + \hat{a}_1 e_{t-1}^2, \quad t = 2, \dots, T,$

вычисление

$$r_t = \sqrt{\frac{1}{h_t} + 2 \left(\frac{\hat{a}_1 e_t}{h_{t+1}} \right)^2}, \quad s_t = \frac{1}{h_t} - \left(\frac{\hat{a}_1}{h_{t+1}} \right) \left(\frac{e_{t+1}^2}{h_{t+1}} - 1 \right),$$

формирование $\mathbf{v} = [e_t s_t / r_t]_2^{T-1}, \quad \mathbf{W} = [r_t \mathbf{x}_t]_2^{t-1},$

вычисление поправочных коэффициентов

$$\mathbf{d}_b = (\mathbf{W}'\mathbf{W})^{-1} \mathbf{W}'\mathbf{v} \quad \text{и}$$

получение $\hat{\mathbf{b}} = \mathbf{b} + \mathbf{d}_b.$

3.1.3. Для тестирования ARCH-эффектов используется критерий TR^2 , где T - объем выборки и R^2 - коэффициент детерминации соответствующей модели для дисперсии. Если расчетное значение $TR^2 > \chi^2(p^*, V)$, где p^* - заданный уровень доверительной вероятности и V - число степеней свободы,

то гипотеза о присутствии ARCH-эффектов принимается, в противном случае – отвергается.

3.2. Решение типовой задачи

Задание 3.2.1. Господин А.В. Мудрин, владеющий 527 акциями компании «Юкос», собирается их продать в ближайшие дни. С целью определения примерной цены, по которой могут купить его акции, он собрал данные, представленные в табл. 3.2.1. По этим данным А.В. Мудрин решил построить авторегрессионную модель, отражающую динамику курса акции во времени, с помощью которой он сможет получить прогнозную оценку курса акций на следующие два дня.

Таблица 3.2.1

Дата	Курс акции, USD	Дата	Курс акции, USD	Дата	Курс акции, USD	Дата	Курс акции, USD
01.07.03	13,82	28.07.03	11,30	22.08.03	14,25	18.09.03	14,62
02.07.03	14,00	29.07.03	11,20	25.08.03	14,30	19.09.03	14,55
03.07.03	13,70	30.07.03	11,15	26.08.03	14,33	22.09.03	14,30
04.07.03	13,50	31.07.03	11,15	27.08.03	13,95	23.09.03	14,50
07.07.03	13,50	01.08.03	11,55	28.08.03	13,90	24.09.03	14,92
08.07.03	13,80	04.08.03	11,60	29.08.03	14,00	25.09.03	14,90
09.07.03	13,95	05.08.03	11,92	01.09.03	14,10	26.09.03	14,60
10.07.03	12,80	06.08.03	12,30	02.09.03	14,71	29.09.03	14,45
11.07.03	12,00	07.08.03	12,75	03.09.03	14,95	30.09.03	14,92
14.07.03	11,60	08.08.03	12,60	04.09.03	14,40	01.10.03	14,75
15.07.03	12,05	11.08.03	12,50	05.09.03	14,40	02.10.03	15,00
16.07.03	11,96	12.08.03	12,99	08.09.03	14,30	03.10.03	15,60
17.07.03	10,70	13.08.03	13,35	09.09.03	14,28	06.10.03	15,30
18.07.03	10,77	14.08.03	13,50	10.09.03	14,18	07.10.03	15,55
21.07.03	10,88	15.08.03	13,62	11.09.03	14,75	08.10.03	15,82
22.07.03	11,20	18.08.03	13,75	12.09.03	14,60	09.10.03	16,25
23.07.03	11,20	19.08.03	14,90	15.09.03	14,85	10.10.03	15,65
24.07.03	11,05	20.08.03	14,50	16.09.03	14,80	13.10.03	15,65
25.07.03	11,25	21.08.03	14,25	17.09.03	15,02	14.10.03	15,62
						15.10.03	15,75

Решение с помощью табличного процессора Excel

1. Ввод исходных данных и оформление в удобном для расчетов виде.
2. Определение порядка авторегрессионной модели

2.1. Формирование двух векторов: 1) вектора (размером 76×1), элементы которого y_{t-1} ; 2) вектора (размером 75×1), элементом которого y_{t-2} . Оформление результатов в виде табл. 3.2.2.

Таблица 3.2.2

t	y_t	y_{t-1}	y_{t-2}	t	y_t	y_{t-1}	y_{t-2}
1	14,00	13,82		39	14,30	14,25	14,25
2	13,70	14,00	13,82	40	14,33	14,30	14,25
3	13,50	13,70	14,00	41	13,95	14,33	14,30
4	13,50	13,50	13,70	42	13,90	13,95	14,33
5	13,80	13,50	13,50	43	14,00	13,90	13,95
6	13,95	13,80	13,50	44	14,10	14,00	13,90
7	12,80	13,95	13,80	45	14,71	14,10	14,00
8	12,00	12,80	13,95	46	14,95	14,71	14,10
9	11,60	12,00	12,80	47	14,40	14,95	14,71
10	12,05	11,60	12,00	48	14,40	14,40	14,95
11	11,96	12,05	11,60	49	14,30	14,40	14,40
12	10,70	11,96	12,05	50	14,28	14,30	14,40
13	10,77	10,70	11,96	51	14,18	14,28	14,30
14	10,88	10,77	10,70	52	14,75	14,18	14,28
15	11,20	10,88	10,77	53	14,60	14,75	14,18
16	11,20	11,20	10,88	54	14,85	14,60	14,75
17	11,05	11,20	11,20	55	14,80	14,85	14,60
18	11,25	11,05	11,20	56	15,02	14,80	14,85
19	11,30	11,25	11,05	57	14,62	15,02	14,80
20	11,20	11,30	11,25	58	14,55	14,62	15,02
21	11,15	11,20	11,30	59	14,30	14,55	14,62
22	11,15	11,15	11,20	60	14,50	14,30	14,55
23	11,55	11,15	11,15	62	14,92	14,50	14,30
24	11,60	11,55	11,15	62	14,90	14,92	14,50
25	11,92	11,60	11,55	63	14,60	14,90	14,92
26	12,30	11,92	11,60	64	14,45	14,60	14,90
27	12,75	12,30	11,92	65	14,92	14,45	14,60
28	12,60	12,75	12,30	66	14,75	14,92	14,45
29	12,50	12,60	12,75	67	15,00	14,75	14,92
30	12,99	12,50	12,60	68	15,60	15,00	14,75
31	13,35	12,99	12,50	69	15,30	15,60	15,00
32	13,50	13,35	12,99	70	15,55	15,30	15,60
33	13,62	13,50	13,35	71	15,82	15,55	15,30
34	13,75	13,62	13,50	72	16,25	15,82	15,55
35	14,90	13,75	13,62	73	15,65	16,25	15,82
36	14,50	14,90	13,75	74	15,65	15,65	16,25
37	14,25	14,50	14,90	75	15,62	15,65	15,65
38	14,25	14,25	14,50	76	15,75	15,62	15,65

2.2. Построение с помощью «Пакета анализа» авторегрессионной модели первого порядка $y_t = b_0 + b_1 y_{t-1} + e_t$ (см. Вывод итогов 2.1).

ВЫВОД ИТОГОВ 2.1

<i>Регрессионная статистика</i>					
Множественный R	0,969239				
R-квадрат	0,939424				
Нормированный R-квадрат	0,938606				
Стандартная ошибка	0,373152				
Наблюдения	76				

<i>Дисперсионный анализ</i>					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>
Регрессия	1	159,7957	159,7957	1147,609	8,4E-47
Остаток	74	10,30393	0,139242		
Итого	75	170,0997			

	<i>Коэффициенты</i>	<i>Стандартная ошибка</i>	<i>t-статистика</i>	<i>P-значение</i>	<i>Нижние 95%</i>	<i>Верхние 95%</i>
Y-пересечение	0,269471	0,397355	0,678161	0,499785	-0,52228	1,061219
Переменная X 1	0,982088	0,02899	33,87637	8,4E-47	0,924324	1,039853

Таким образом, построенная модель имеет вид

$$\hat{y}_t = 0,269 + 0,982 y_{t-1}.$$

Расчетное значение t-статистики свидетельствует о значимости коэффициента \hat{b}_1 .

2.3. Построение с помощью «Пакета анализа» авторегрессионной модели второго порядка $y_t = b_0 + b_1 y_{t-1} + b_2 y_{t-2} + e_t$ (см. Вывод итогов 2.2).

Таким образом, построенная модель имеет вид

$$\hat{y}_t = 0,318 + 1,064 y_{t-1} - 0,086 y_{t-2}.$$

Расчетные значения t-статистик свидетельствуют о значимости коэффициента \hat{b}_1 и незначимости коэффициента \hat{b}_2 , который, по сути, является частным коэффициентом автокорреляции второго порядка. Незначимость этого коэффициента позволяет сделать вывод, что исходные данные описываются авторегрессией первого порядка.

ВЫВОД ИТОГОВ 2.2

Регрессионная статистика	
Множественный R	0,969518
R-квадрат	0,939965
Нормированный R-квадрат	0,938298
Стандартная ошибка	0,37647
Наблюдения	75

Дисперсионный анализ					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	Значимость <i>F</i>
Регрессия	2	159,7725	79,88625	563,6533	1,05E-44
Остаток	72	10,20452	0,141729		
Итого	74	169,977			

	Коэффициенты	Стандартная ошибка	t-статистика	P-значение	Нижние 95%	Верхние 95%
Y-пересечение	0,3185	0,406484	0,78355	0,435874	-0,49181	1,12881
Переменная X 1	1,064107	0,117398	9,064117	1,59E-13	0,830079	1,298135
Переменная X 2	-0,08592	0,118837	-0,72303	0,472002	-0,32282	0,150974

3. Расчет остатков авторегрессионного уравнения первого порядка и оформление результатов расчетов в виде табл. 2.2.3.

Таблица 2.2.3

t	y_t	\hat{y}_t	e_t^2	t	y_t	\hat{y}_t	e_t^2
1	14,00	13,8419	0,0250	39	14,30	14,2642	0,0013
2	13,70	14,0187	0,1016	40	14,33	14,3133	0,0003
3	13,50	13,7241	0,0502	41	13,95	14,3428	0,1543
4	13,50	13,5277	0,0008	42	13,90	13,9696	0,0048
5	13,80	13,5277	0,0742	43	14,00	13,9205	0,0063
6	13,95	13,8223	0,0163	44	14,10	14,0187	0,0066
7	12,80	13,9696	1,3680	45	14,71	14,1169	0,3517
8	12,00	12,8402	0,7059	46	14,95	14,7160	0,0548
9	11,60	12,0545	0,2066	47	14,40	14,9517	0,3044
10	12,05	11,6617	0,1508	48	14,40	14,4115	0,0001
11	11,96	12,1036	0,0206	49	14,30	14,4115	0,0124
12	10,70	12,0152	1,7299	50	14,28	14,3133	0,0011
13	10,77	10,7778	0,0001	51	14,18	14,2937	0,0129
14	10,88	10,8466	0,0011	52	14,75	14,1955	0,3075
15	11,20	10,9546	0,0602	53	14,60	14,7553	0,0241
16	11,20	11,2689	0,0047	54	14,85	14,6080	0,0586
17	11,05	11,2689	0,0479	55	14,80	14,8535	0,0029
18	11,25	11,1215	0,0165	56	15,02	14,8044	0,0465
19	11,30	11,3180	0,0003	57	14,62	15,0204	0,1604
20	11,20	11,3671	0,0279	58	14,55	14,6276	0,0060
21	11,15	11,2689	0,0141	59	14,30	14,5589	0,0670

Продолжение табл. 2.2.3

t	y_t	\hat{y}_t	e_t^2	t	y_t	\hat{y}_t	e_t^2
22	11,15	11,2198	0,0049	60	14,50	14,3133	0,0348
23	11,55	11,2198	0,1091	62	14,92	14,5098	0,1683
24	11,60	11,6126	0,0002	62	14,90	14,9222	0,0005
25	11,92	11,6617	0,0667	63	14,60	14,9026	0,0916
26	12,30	11,9760	0,1050	64	14,45	14,6080	0,0250
27	12,75	12,3492	0,1607	65	14,92	14,4606	0,2110
28	12,60	12,7911	0,0365	66	14,75	14,9222	0,0297
29	12,50	12,6438	0,0207	67	15,00	14,7553	0,0599
30	12,99	12,5456	0,1975	68	15,60	15,0008	0,3590
31	13,35	13,0268	0,1045	69	15,30	15,5900	0,0841
32	13,50	13,3804	0,0143	70	15,55	15,2954	0,0648
33	13,62	13,5277	0,0085	71	15,82	15,5409	0,0779
34	13,75	13,645	0,0109	72	16,25	15,8061	0,1970
35	14,90	13,7732	1,2697	73	15,65	16,2284	0,3346
36	14,50	14,9026	0,1621	74	15,65	15,6392	0,0001
37	14,25	14,5098	0,0675	75	15,62	15,6392	0,0004
38	14,25	14,2642	0,0002	76	15,75	15,6097	0,0197

4. Вычисление дисперсий по формуле $s_t^2 = \frac{1}{t} \sum_{i=1}^t e_i^2$ и оформление результатов расчетов в виде табл. 2.2.4.

Таблица 2.2.4

y_t	s_t^2	y_t	s_t^2	y_t	s_t^2	y_t	s_t^2
13,70	0,0633	11,15	0,2201	14,33	0,1741	14,30	0,1448
13,50	0,0589	11,15	0,2103	13,95	0,1736	14,50	0,1430
13,50	0,0444	11,55	0,2059	13,90	0,1696	14,92	0,1434
13,80	0,0503	11,60	0,1974	14,00	0,1658	14,90	0,1411
13,95	0,0447	11,92	0,1921	14,10	0,1622	14,60	0,1403
12,80	0,2337	12,30	0,1888	14,71	0,1664	14,45	0,1385
12,00	0,2927	12,75	0,1877	14,95	0,1640	14,92	0,1396
11,60	0,2832	12,60	0,1823	14,40	0,1669	14,75	0,1380
12,05	0,2699	12,50	0,1768	14,40	0,1635	15,00	0,1368
11,96	0,2473	12,99	0,1775	14,30	0,1604	15,60	0,1401
10,70	0,3708	13,35	0,1751	14,28	0,1572	15,30	0,1393
10,77	0,3423	13,50	0,1701	14,18	0,1544	15,55	0,1382
10,88	0,3179	13,62	0,1652	14,75	0,1573	15,82	0,1374
11,20	0,3007	13,75	0,1607	14,60	0,1548	16,25	0,1382
11,20	0,2822	14,90	0,1923	14,85	0,1530	15,65	0,1409
11,05	0,2685	14,50	0,1915	14,80	0,1503	15,65	0,1390
11,25	0,2545	14,25	0,1881	15,02	0,1484	15,62	0,1371
11,30	0,2411	14,25	0,1832	14,62	0,1486	15,75	0,1356
11,20	0,2304	14,30	0,1785	14,55	0,1462		

5. Построение с помощью «Пакета анализа» регрессионного уравнения $\hat{S}_t^2 = a_0 + a_1 e_{t-1}^2$ (см. Вывод итогов 2.3).

ВЫВОД ИТОГОВ 2.3

<i>Регрессионная статистика</i>	
Множественный R	0,336649
R-квадрат	0,113333
Нормированный R-квадрат	0,101187
Стандартная ошибка	0,059354
Наблюдения	75

<i>Дисперсионный анализ</i>					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>
Регрессия	1	0,032872	0,032872	9,330783	0,003145
Остаток	73	0,257173	0,003523		
Итого	74	0,290045			

	<i>Коэффициенты</i>	<i>Стандартная ошибка</i>	<i>t-статистика</i>	<i>P-значение</i>	<i>Нижние 95%</i>	<i>Верхние 95%</i>
Y-пересечение	0,167468	0,007555	22,16555	3,79E-34	0,15241	0,182525
Переменная X 1	0,070833	0,023189	3,054633	0,003145	0,024618	0,117048

Таким образом, построенная модель имеет вид

$$\hat{S}_t^2 = 0,167 + 0,071 e_{t-1}^2.$$

Расчетные значения t-статистик свидетельствуют о значимости полученных коэффициентов регрессии.

6. Тестирование ARCH-модели на значимость с помощью TR^2 -критерия

$$TR^2 = 75 \cdot 0,113 = 8,50.$$

Сравнение полученного значения критерия с табличным значением распределения $\chi^2_{0,95}(1) = 3,84$ свидетельствует о присутствии ARCH-эффекта.

7. Вычисление расчетных значений \hat{S}_t^2 для $t = \overline{2, T}$. Оформление результатов в виде табл. 5.

8. Формирование новых переменных для $t = \overline{2, T}$:

$$1) g_t = (e_t^2 / \hat{S}_t^2 - 1); \quad 2) z_{t0} = 1 / \hat{S}_t^2; \quad 3) z_{t1} = e_{t-1}^2 / \hat{S}_t^2.$$

Оформление результатов расчетов в виде табл. 2.2.5.

Таблица 2.2.5

\hat{S}_t^2	g_t	z_{t0}	z_{t1}	\hat{S}_t^2	g_t	z_{t0}	z_{t1}
0,1692	-0,3998	5,9089	0,1476	0,1676	-0,9983	5,9681	0,0076
0,1747	-0,7125	5,7253	0,5816	0,1675	-0,0788	5,9706	0,0017
0,1710	-0,9955	5,8471	0,2936	0,1784	-0,9728	5,6055	0,8649
0,1675	-0,5573	5,9694	0,0046	0,1678	-0,9623	5,9591	0,0289
0,1727	-0,9056	5,7897	0,4294	0,1679	-0,9606	5,9554	0,0376
0,1686	7,1126	5,9304	0,0967	0,1679	1,0945	5,9547	0,0394
0,2644	1,6703	3,7826	5,1746	0,1924	-0,7154	5,1980	1,8284
0,2175	-0,0500	4,5983	3,2461	0,1713	0,7763	5,8361	0,3196
0,1821	-0,1720	5,4914	1,1345	0,1890	-0,9993	5,2903	1,6102
0,1781	-0,8842	5,6133	0,8464	0,1675	-0,9257	5,9710	0,0008
0,1689	9,2403	5,9196	0,1221	0,1683	-0,9934	5,9400	0,0739
0,2900	-0,9998	3,4483	5,9651	0,1675	-0,9229	5,9685	0,0066
0,1675	-0,9933	5,9711	0,0004	0,1684	0,8261	5,9388	0,0768
0,1675	-0,6406	5,9685	0,0067	0,1892	-0,8726	5,2841	1,6248
0,1717	-0,9724	5,8230	0,3507	0,1692	-0,6537	5,9110	0,1425
0,1678	-0,7145	5,9593	0,0283	0,1716	-0,9833	5,8269	0,3414
0,1709	-0,9034	5,8527	0,2803	0,1677	-0,7227	5,9641	0,0171
0,1686	-0,9981	5,9299	0,0978	0,1708	-0,0610	5,8561	0,2723
0,1675	-0,8333	5,9705	0,0019	0,1788	-0,9663	5,5920	0,8967
0,1694	-0,9166	5,9016	0,1647	0,1679	-0,6009	5,9561	0,0359
0,1685	-0,9711	5,9358	0,0839	0,1722	-0,7977	5,8067	0,3891
0,1678	-0,3501	5,9590	0,0290	0,1699	-0,0096	5,8846	0,2050
0,1752	-0,9991	5,7080	0,6225	0,1794	-0,9972	5,5745	0,9382
0,1675	-0,6016	5,9709	0,0009	0,1675	-0,4534	5,9701	0,0030
0,1722	-0,3902	5,8074	0,3875	0,1740	-0,8566	5,7487	0,5263
0,1749	-0,0814	5,7174	0,6003	0,1692	0,2468	5,9089	0,1474
0,1788	-0,7958	5,5913	0,8984	0,1824	-0,8374	5,4820	1,1567
0,1701	-0,8784	5,8805	0,2147	0,1696	-0,6468	5,8973	0,1749
0,1689	0,1692	5,9195	0,1224	0,1717	1,0910	5,8238	0,3488
0,1815	-0,4243	5,5109	1,0885	0,1929	-0,5639	5,1840	1,8613
0,1749	-0,9181	5,7186	0,5974	0,1734	-0,6263	5,7661	0,4851
0,1685	-0,9494	5,9354	0,0850	0,1721	-0,5474	5,8120	0,3767
0,1681	-0,9350	5,9498	0,0507	0,1730	0,1391	5,7809	0,4502
0,1682	6,5470	5,9439	0,0649	0,1814	0,8440	5,5119	1,0861
0,2574	-0,3703	3,8849	4,9327	0,1912	-0,9994	5,2311	1,7501
0,1789	-0,6230	5,5882	0,9057	0,1675	-0,9978	5,9710	0,0007
0,1722	-0,9988	5,8056	0,3917	0,1675	-0,8825	5,9704	0,0022
0,1675	-0,9924	5,9708	0,0012				

9. Построение регрессионного уравнения $g_t = d_0 z_{t0} + d_1 z_{t1}$ с нулевым свободным членом (см. Вывод итогов 2.4), коэффициенты которого являются поправочными коэффициентами уравнения дисперсии.

ВЫВОД ИТОГОВ 2.4

Регрессионная статистика	
Множественный R	0,028703
R-квадрат	0,000824
Нормированный R-квадрат	-0,02656
Стандартная ошибка	1,742834
Наблюдения	75

Дисперсионный анализ					Значимость
	df	SS	MS	F	F
Регрессия	2	0,18283	0,091415	0,030096	0,970365
Остаток	73	221,7354	3,037471		
Итого	75	221,9182			

	Коэффициенты	Стандартная ошибка	t-статистика	P-значение	Нижние 95%	Верхние 95%
Y-пересечение	0	#Н/Д	#Н/Д	#Н/Д	#Н/Д	#Н/Д
Переменная X 1	-0,03274	0,038801	-0,84389	0,401487	-0,11007	0,044587
Переменная X 2	-0,0577	0,171192	-0,33704	0,737051	-0,39889	0,283487

Таким образом, построенная модель имеет вид

$$\hat{g}_t = -0,033z_{t0} - 0,058z_{t1}.$$

10. Корректировка коэффициентов уравнения дисперсии

$$\hat{\hat{a}}_0 = \hat{a}_0 + \hat{d}_0 = 0,167 - 0,033 = 0,134,$$

$$\hat{\hat{a}}_1 = \hat{a}_1 + \hat{d}_1 = 0,071 - 0,058 = 0,013.$$

11. Расчет уточненных значений $\hat{\hat{S}}_t^2$ по уравнению

$$\hat{\hat{S}}_t^2 = 0,134 + 0,013e_{t-1}^2.$$

Оформление результатов расчетов в виде табл. 2.2.6.

12. Формирование новых переменных для $t = \overline{2, T-1}$:

$$1) r_t = \left(\frac{1}{\hat{\hat{S}}_t^2} + 2 \left(\frac{\hat{\hat{a}}_1 e_t}{\hat{\hat{S}}_{t+1}^2} \right)^2 \right)^{1/2}; \quad 2) s_t = \frac{1}{\hat{\hat{S}}_t^2} - \left(\frac{\hat{\hat{a}}_1}{\hat{\hat{S}}_{t+1}^2} \right) \left(\frac{e_{t+1}^2}{\hat{\hat{S}}_{t+1}^2} - 1 \right);$$

$$3) v_t = \frac{e_t s_t}{r_t}; \quad 4) w_{t0} = 1 \cdot r_t \text{ и } w_{t1} = r_t y_t.$$

Оформление результатов расчетов в виде табл. 2.2.6.

Таблица 2.2.6

r_t	s_t	v_t	w_{t0}	w_{t1}	r_t	s_t	v_t	w_{t0}	w_{t1}
2,4310	5,9624	-0,7817	2,4310	34,0347	2,4435	6,0490	0,0885	2,4435	34,8202
2,3929	5,8018	-0,5433	2,3929	32,7826	2,4430	5,9742	0,0408	2,4430	34,9344
2,4181	5,8908	-0,0674	2,4181	32,6441	2,4438	6,0422	-0,9712	2,4438	35,0200
2,4434	6,0382	0,6730	2,4434	32,9860	2,3676	5,6808	-0,1670	2,3676	33,0281
2,4062	5,2357	0,2779	2,4062	33,2058	2,4411	6,0342	0,1965	2,4411	33,9318
2,4366	5,8474	-2,8068	2,4366	33,9909	2,4404	5,8698	0,1955	2,4404	34,1653
1,9462	3,7857	-1,6343	1,9462	24,9117	2,4409	6,0035	1,4587	2,4409	34,4165
2,1449	4,6107	-0,9771	2,1449	25,7384	2,2800	5,1385	0,5274	2,2800	33,5395
2,3437	5,5566	0,9206	2,3437	27,1873	2,4164	5,9056	-1,3483	2,4164	36,1254
2,3693	4,8949	-0,2967	2,3693	28,5500	2,3001	5,3628	-0,0269	2,3001	33,1208
2,4345	5,9649	-3,2226	2,4345	29,1165	2,4436	6,0485	-0,2761	2,4436	35,1876
1,8570	3,5262	-0,0148	1,8570	19,8694	2,4372	6,0124	-0,0822	2,4372	34,8523
2,4436	6,0214	0,0824	2,4436	26,3175	2,4431	5,9041	-0,2748	2,4431	34,8872
2,4432	6,0428	0,6070	2,4432	26,5819	2,4376	5,9994	1,3648	2,4376	34,5649
2,4131	5,8789	-0,1678	2,4131	27,0267	2,2988	5,3348	-0,3604	2,2988	33,9069
2,4413	6,0288	-0,5405	2,4413	27,3425	2,4314	5,9863	0,5959	2,4314	35,4984
2,4193	5,9305	0,3149	2,4193	26,7331	2,4139	5,8835	-0,1304	2,4139	35,8465
2,4351	5,9953	-0,0442	2,4351	27,3954	2,4423	5,9688	0,5270	2,4423	36,1454
2,4435	6,0415	-0,4131	2,4435	27,6118	2,4203	5,9271	-0,9806	2,4203	36,3530
2,4294	5,9773	-0,2925	2,4294	27,2088	2,3648	5,6390	-0,1851	2,3648	34,5728
2,4364	5,9632	-0,1707	2,4364	27,1655	2,4407	6,0170	-0,6382	2,4407	35,5119
2,4414	6,0339	0,8162	2,4414	27,2212	2,4098	5,8075	0,4499	2,4098	34,4602
2,3891	5,7552	-0,0303	2,3891	27,5946	2,4262	5,9576	1,0074	2,4262	35,1797
2,4437	6,0007	0,6343	2,4437	28,3469	2,3610	5,6100	-0,0528	2,3610	35,2266
2,4101	5,8135	0,7816	2,4101	28,7284	2,4436	6,0347	-0,7473	2,4436	36,4094
2,3915	5,7758	0,9681	2,3915	29,4151	2,3977	5,7295	-0,3775	2,3977	35,0064
2,3647	5,6592	-0,4573	2,3647	30,1498	2,4313	5,9692	1,1278	2,4313	35,1320
2,4250	5,8673	-0,3479	2,4250	30,5552	2,3415	5,5321	-0,4069	2,3415	34,9345
2,4334	5,9502	1,0867	2,4334	30,4179	2,4286	5,8139	0,5859	2,4286	35,8216
2,3478	5,5799	0,7681	2,3478	30,4977	2,4139	5,8622	1,4551	2,4139	36,2091
2,3914	5,7926	0,2898	2,3914	31,9252	2,2771	5,2315	-0,6664	2,2771	35,5221
2,4363	6,0084	0,2277	2,4363	32,8898	2,4014	5,8079	0,6157	2,4014	36,7419
2,4393	5,4388	0,2330	2,4393	33,2227	2,4110	5,8014	0,6715	2,4110	37,4909
2,4394	5,9627	2,7544	2,4394	33,5412	2,4048	5,7198	1,0558	2,4048	38,0436
1,9715	3,9306	-0,8027	1,9715	29,3748	2,3484	5,5806	-1,3745	2,3484	38,1619
2,3641	5,6644	-0,6224	2,3641	34,2795	2,2872	5,3093	0,0252	2,2872	35,7940
2,4095	5,8834	-0,0347	2,4095	34,3352	2,4436	6,0402	-0,0473	2,4436	38,2418

13. Построение с помощью «Пакета анализа» регрессионного уравнения $v_t = c_0 w_{t0} + c_1 w_{t1}$ с нулевым свободным членом (см. Вывод итогов 2.5), параметры которого являются корректировочными коэффициентами исходной (авторегрессионной модели первого порядка) модели.

ВЫВОД ИТОГОВ 2.5

Регрессионная статистика	
Множественный	
R	0,015045
R-квадрат	0,000226
Нормированный	
R-квадрат	-0,02755
Стандартная	
ошибка	0,900989
Наблюдения	74

Дисперсионный анализ					Значимость
	df	SS	MS	F	F
Регрессия	2	0,013232	0,006616	0,00815	0,991884
Остаток	72	58,44825	0,811781		
Итого	74	58,46149			

	Коэффициенты	Стандартная ошибка	t- статистика	P- значение	Нижние 95%	Верхние 95%
Y-пересечение	0	#Н/Д	#Н/Д	#Н/Д	#Н/Д	#Н/Д
Переменная X 1	0,04081	0,408756	0,099839	0,92075	-0,77403	0,855649
Переменная X 2	-0,00271	0,029878	-0,0907	0,927984	-0,06227	0,056851

Таким образом, построенная модель с корректировочными коэффициентами имеет вид

$$v_t = 0,041w_{t0} - 0,003w_{t1}.$$

14.Корректировка коэффициентов построенного уравнения

$$\hat{b}_0 = \hat{b}_0 + \hat{c}_0 = 0,269 + 0,041 = 0,310 ,$$

$$\hat{b}_1 = \hat{b}_1 + \hat{c}_1 = 0,982 - 0,003 = 0,979 .$$

15. Выполнение п. 3-6. Тестирование скорректированной модели на наличие ARCH-эффекта с помощью TR^2 -критерия

$$TR^2 = 75 \cdot 0,011 = 0,828.$$

Сравнение полученного значения критерия с табличным значением распределения $c_{0,95}^2(1) = 3,84$ свидетельствует об отсутствии ARCH-эффекта. Следовательно, модель $y_t = 0,310 + 0,979 y_{t-1}$ пригодна для целей прогнозирования.

16. Прогноз курса акций на 16.10.03 и 17.10.03

$$\hat{y}_{t+1} = 0,310 + 0,979 \cdot 15,75 = 15,73;$$

$$\hat{y}_{t+2} = 0,310 + 0,979 \cdot 15,73 = 15,71.$$

3.3. Задания для самостоятельной работы

Задание 3.3.1. По данным табл. 3.3.1, отражающей динамику величины средней базовой ставки банковского процента в США, постройте авторегрессионную модель, предварительно установив ее порядок. Проведите расчеты, необходимые для тестирования ARCH-эффектов. Если наличие таких эффектов подтвердится, то скорректируйте коэффициенты авторегрессионной модели, применив четырехэтапную процедуру построения ARCH модели.

Таблица 3.3.1

Дата	Ставка, %	Дата	Ставка, %	Дата	Ставка, %	Дата	Ставка, %	Дата	Ставка, %
01.11.80	16,06	01.11.86	7,50	01.11.83	11,00	01.11.89	10,50	01.11.92	6,00
01.12.80	20,35	01.12.86	7,50	01.12.83	11,00	01.12.89	10,50	01.12.92	6,00
01.01.81	20,16	01.01.87	7,50	01.01.84	11,00	01.01.90	10,11	01.01.93	6,00
01.02.81	19,43	01.02.87	7,50	01.02.84	11,00	01.02.90	10,00	01.02.93	6,00
01.03.81	18,05	01.03.87	7,50	01.03.84	11,21	01.03.90	10,00	01.03.93	6,00
01.04.81	17,15	01.04.87	7,75	01.04.84	11,93	01.04.90	10,00	01.04.93	6,00
01.05.81	19,61	01.05.87	8,14	01.05.84	12,39	01.05.90	10,00	01.05.93	6,00
01.06.81	20,03	01.06.87	8,25	01.06.84	12,60	01.06.90	10,00	01.06.93	6,00
01.07.81	20,39	01.07.87	8,25	01.07.84	13,00	01.07.90	10,00	01.07.93	6,00
01.08.81	20,50	01.08.87	8,25	01.08.84	13,00	01.08.90	10,00	01.08.93	6,00
01.09.81	20,08	01.09.87	8,70	01.09.84	12,97	01.09.90	10,00	01.09.93	6,00
01.10.81	18,50	01.10.87	9,07	01.10.84	12,58	01.10.90	10,00	01.10.93	6,00
01.11.81	16,84	01.11.87	8,78	01.11.84	11,77	01.11.90	10,00	01.11.93	6,00
01.12.81	15,75	01.12.87	8,75	01.12.84	11,06	01.12.90	10,00	01.12.93	6,00
01.01.82	15,75	01.01.88	8,75	01.01.85	10,61	01.01.91	9,52	01.01.94	6,00
01.02.82	16,56	01.02.88	8,51	01.02.85	10,50	01.02.91	9,05	01.02.94	6,00
01.03.82	16,50	01.03.88	8,50	01.03.85	10,50	01.03.91	9,00	01.03.94	6,06
01.04.82	16,50	01.04.88	8,50	01.04.85	10,50	01.04.91	9,00	01.04.94	6,45
01.05.82	16,50	01.05.88	8,84	01.05.85	10,31	01.05.91	8,50	01.05.94	6,99
01.06.82	16,50	01.06.88	9,00	01.06.85	9,78	01.06.91	8,50	01.06.94	7,25
01.07.82	16,26	01.07.88	9,29	01.07.85	9,50	01.07.91	8,50	01.07.94	7,25
01.08.82	14,39	01.08.88	9,84	01.08.85	9,50	01.08.91	8,50	01.08.94	7,51
01.09.82	13,50	01.09.88	10,00	01.09.85	9,50	01.09.91	8,20	01.09.94	7,75
01.10.82	12,52	01.10.88	10,00	01.10.85	9,50	01.10.91	8,00	01.10.94	7,75
01.11.82	11,85	01.11.88	10,05	01.11.85	9,50	01.11.91	7,58	01.11.94	8,15
01.12.82	11,50	01.12.88	10,50	01.12.85	9,50	01.12.91	7,21	01.12.94	8,50
01.01.83	11,16	01.01.89	10,50	01.01.86	9,50	01.01.92	6,50	01.01.95	8,50
01.02.83	10,98	01.02.89	10,93	01.02.86	9,50	01.02.92	6,50	01.02.95	9,00
01.03.83	10,50	01.03.89	11,50	01.03.86	9,10	01.03.92	6,50	01.03.95	9,00
01.04.83	10,50	01.04.89	11,50	01.04.86	8,83	01.04.92	6,50	01.04.95	9,00
01.05.83	10,50	01.05.89	11,50	01.05.86	8,50	01.05.92	6,50	01.05.95	9,00
01.06.83	10,50	01.06.89	11,07	01.06.86	8,50	01.06.92	6,50	01.06.95	9,00
01.07.83	10,50	01.07.89	10,98	01.07.86	8,16	01.07.92	6,02	01.07.95	8,80
01.08.83	10,89	01.08.89	10,50	01.08.86	7,90	01.08.92	6,00	01.08.95	8,75
01.09.83	11,00	01.09.89	10,50	01.09.86	7,50	01.09.92	6,00	01.09.95	8,75
01.10.83	11,00	01.10.89	10,50	01.10.86	7,50	01.10.92	6,00	01.10.95	8,75

Задание 3.3.2. Один из акционеров ОАО «Клеопатра», осуществляющей производство натуральных эссенций и масел для разнообразной парфюмерной продукции, в связи с некоторым затруднением в материальном положении желает продать свои 2187 акций в следующем месяце. Для того чтобы определить примерную рыночную стоимость одной акции, он решил получить прогнозную оценку дивидендов на акцию в интересующем его периоде. С этой целью акционер сформировал табл. 3.3.2. Постройте модель, наилучшим образом отражающую динамику величины дивидендов, и осуществите требуемые прогнозные расчеты.

Таблица 3.3.2

Дата	Дивиденды на акцию, евро	Дата	Дивиденды на акцию, евро	Дата	Дивиденды на акцию, евро	Дата	Дивиденды на акцию, евро
01.01.98	7,30	01.07.99	9,79	01.01.01	11,23	01.07.02	13,44
01.02.98	7,63	01.08.99	9,77	01.02.01	12,29	01.08.02	13,51
01.03.98	7,65	01.09.99	9,89	01.03.01	11,53	01.09.02	13,63
01.04.98	7,97	01.10.99	10,00	01.04.01	12,58	01.10.02	13,70
01.05.98	8,00	01.11.99	10,24	01.05.01	10,82	01.11.02	13,61
01.06.98	8,23	01.12.99	10,24	01.06.01	11,41	01.12.02	13,72
01.07.98	8,19	01.01.00	10,35	01.07.01	10,91	01.01.03	13,98
01.08.98	8,35	01.02.00	10,43	01.08.01	12,02	01.02.03	14,03
01.09.98	8,53	01.03.00	10,75	01.09.01	12,41	01.03.03	13,99
01.10.98	7,90	01.04.00	10,77	01.10.01	12,56	01.04.03	13,88
01.11.98	8,25	01.05.00	11,02	01.11.01	12,33	01.05.03	14,41
01.12.98	7,69	01.06.00	10,85	01.12.01	13,14	01.06.03	15,02
01.01.99	7,66	01.07.00	11,18	01.01.02	12,49	01.07.03	14,59
01.02.99	8,38	01.08.00	9,98	01.02.02	12,20	01.08.03	13,33
01.03.99	8,63	01.09.00	11,62	01.03.02	12,50	01.09.03	15,08
01.04.99	7,72	01.10.00	12,00	01.04.02	13,39	01.10.03	15,31
01.05.99	9,00	01.11.00	11,73	01.05.02	11,90	01.11.03	14,85
01.06.99	9,24	01.12.00	11,48	01.06.02	13,24	01.12.03	15,97

4. МОДЕЛИ ДЛЯ ПАНЕЛЬНЫХ ДАННЫХ

4.1. Расчетные формулы

4.1.1. Модель с фиксированными эффектами

$$y_i = \mathbf{i}a_i + \mathbf{X}_i\mathbf{b} + e_i.$$

4.1.2. F -статистика, используемая для проверки значимости групповых эффектов

$$F(n-1, nT-n-K) = \frac{(R_u^2 - R_p^2)/(n-1)}{(1 - R_u^2)/(nT-n-K)},$$

где n - число панелей, T - число наблюдений в сбалансированной панели, K - число факторов в модели.

4.1.3. МНК-оценки \mathbf{b} для разделенной регрессии

$$\hat{\mathbf{b}} = [\mathbf{X}'\mathbf{M}_d\mathbf{X}]^{-1}[\mathbf{X}'\mathbf{M}_d\mathbf{y}],$$

где $\mathbf{M}_d = \mathbf{I} - \mathbf{D}(\mathbf{D}'\mathbf{D})^{-1}\mathbf{D}'$,

\mathbf{D} - матрица, элементы которой являются фиктивными переменными.

4.1.4. Коэффициенты фиктивных переменных

$$\mathbf{a} = [\mathbf{D}'\mathbf{D}]^{-1}\mathbf{D}'(\mathbf{y} - \mathbf{X}\hat{\mathbf{b}}).$$

4.1.5. Оценка ковариационной матрицы для $\hat{\mathbf{b}}$

$$Est.Var[\mathbf{b}] = s^2[\mathbf{X}'\mathbf{M}_d\mathbf{X}]^{-1},$$

$$\text{где } s^2 = \frac{\sum_{i=1}^n \sum_{t=1}^T (y_{it} - a_i - \mathbf{x}'_{it}\hat{\mathbf{b}})^2}{nT - n - K}.$$

4.1.6. Дисперсия для индивидуальных эффектов

$$Var[a_i] = \frac{s^2}{T} + \bar{\mathbf{x}}'_{i\bullet} Var[\hat{\mathbf{b}}] \bar{\mathbf{x}}_{i\bullet}.$$

4.1.7. Преобразование данных i -го блока \mathbf{y}_i и \mathbf{X}_i для применения обобщенного МНК

$$\Omega^{-1/2}\mathbf{y}_i = \frac{1}{s_e} \begin{bmatrix} y_{i1} - q \bar{y}_i \\ y_{i2} - q \bar{y}_i \\ \mathbf{M} \\ y_{iT} - q \bar{y}_i \end{bmatrix},$$

$$\text{где } q = 1 - \frac{s_e}{\sqrt{Ts_u^2 + s_e^2}}.$$

4.1.8. Остаточная дисперсия модели с фиксированными эффектами

$$\hat{s}_e^2 = \frac{\sum_{i=1}^n \sum_{t=1}^T (e_{it} - \bar{e}_{i\bullet})^2}{nT - n - K}.$$

4.1.9. Дисперсия, характеризующая вариацию наблюдений, относящихся к различным панелям

$$\hat{S}_u^2 = \hat{S}_{**}^2 - \frac{\hat{S}_e^2}{T},$$

$$\text{где } \hat{S}_{**}^2 = \frac{\mathbf{e}_{**}'\mathbf{e}_{**}}{n-K}, \quad e_{**i} = \bar{y}_{i\bullet} - \hat{a} - \hat{\mathbf{b}}'\bar{\mathbf{x}}_{i\bullet}.$$

4.1.10. Статистика хи-квадрат теста, основанного на критерии Вальда

$$W = c^2[K] = \left[\hat{\mathbf{b}} - \hat{\mathbf{b}} \right]' \hat{\Sigma}^{-1} \left[\hat{\mathbf{b}} - \hat{\mathbf{b}} \right],$$

где $\hat{\mathbf{b}}$ - оценка обычного МНК, $\hat{\mathbf{b}}$ - оценка обобщенного МНК, $\hat{\Sigma}^{-1}$ - ковариационная матрица оценок коэффициентов модели случайных эффектов, исключая константу.

4.2. Решение типовой задачи

Таблица 4.2.1

№ п.п.	у	x_1	x_2	№ п.п.	у	x_1	x_2
Орловская область				Воронежская область			
1	20,76	0,24	10,23	15	65,01	0,94	10,36
2	28,09	0,31	10,89	16	69,05	1,21	11,36
3	32,95	0,55	10,28	17	73,13	1,29	8,89
4	38,15	0,67	10,32	18	81,18	1,49	7,55
5	46,78	0,83	10,85	19	89,24	1,67	7,81
6	55,31	0,98	11,38	20	97,30	1,84	8,08
7	60,92	1,14	11,91	21	115,36	2,02	11,84
Белгородская область				Липецкая область			
8	41,08	0,45	1,45	22	91,26	1,12	10,72
9	56,29	0,78	2,02	23	99,84	1,29	11,27
10	68,51	0,98	3,77	24	108,55	1,49	13,02
11	82,72	1,24	5,52	25	117,17	1,67	13,41
12	96,43	1,49	7,51	26	125,81	1,85	13,62
13	110,15	1,74	9,04	27	134,46	2,04	14,34
14	123,86	1,99	12,01	28	143,10	2,22	14,85

Задание 4.2.1. Учредитель крупнейшей сети недорогих супермаркетов «Пятерочка», как и любой предприниматель, заинтересован в росте доходов от своего бизнеса. Очевидно, что доход напрямую зависит от величины товарооборота. С целью изыскания путей увеличения годового товарооборота (млн. руб., у), он поручил специалистам компании изучить фак

торы, влияющие на этот показатель, в четырех регионах России. В ходе исследования было выявлено, что такими факторами являются торговая площадь (тыс. кв. м., x_1) и среднее число посетителей в день (тыс. чел., x_2), и сформирована табл. 4.2.1. По представленным в этой таблице панельным данным было решено построить регрессионную модель, отражающую зависимость товарооборота от соответствующих факторов.

Решение с помощью табличного процессора Excel

1. Ввод исходных данных и оформление их в виде, удобном для проведения расчетов.
2. Построение регрессионной модели с фиксированными эффектами.
 - 2.1. Построение модели с использованием фиктивных переменных.
 - 2.1.1. Формирование фиктивных переменных и оформление результатов в виде табл. 4.2.2.

Таблица 4.2.2

№ п.п.	y	i_1	i_2	i_3	i_4	x_1	x_2
1	20,76	1	0	0	0	0,24	10,23
2	28,09	1	0	0	0	0,31	10,89
3	32,95	1	0	0	0	0,55	10,28
4	38,15	1	0	0	0	0,67	10,32
5	46,78	1	0	0	0	0,83	10,85
6	55,31	1	0	0	0	0,98	11,38
7	60,92	1	0	0	0	1,14	11,91
8	41,08	0	1	0	0	0,45	1,45
9	56,29	0	1	0	0	0,78	2,02
10	68,51	0	1	0	0	0,98	3,77
11	82,72	0	1	0	0	1,24	5,52
12	96,43	0	1	0	0	1,49	7,51
13	110,15	0	1	0	0	1,74	9,04
14	123,86	0	1	0	0	1,99	12,01
15	65,01	0	0	1	0	0,94	10,36
16	69,05	0	0	1	0	1,21	11,36
17	73,13	0	0	1	0	1,29	8,89
18	81,18	0	0	1	0	1,49	7,55
19	89,24	0	0	1	0	1,67	7,81
20	97,30	0	0	1	0	1,84	8,08
21	115,36	0	0	1	0	2,02	11,84
22	91,26	0	0	0	1	1,12	10,72
23	99,84	0	0	0	1	1,29	11,27
24	108,55	0	0	0	1	1,49	13,02
25	117,17	0	0	0	1	1,67	13,41
26	125,81	0	0	0	1	1,85	13,62
27	134,46	0	0	0	1	2,04	14,34
28	143,10	0	0	0	1	2,22	14,85

2.1.2. Построение по данным табл. 4.2.2 регрессионного уравнения (без свободного члена) с использованием «Пакета анализа» (см. Вывод итогов 4.1).

ВЫВОД ИТОГОВ 4.1

<i>Регрессионная статистика</i>	
Множественный R	0,998463
R-квадрат	0,996929
Нормированный R-квадрат	0,950777
Стандартная ошибка	2,090854
Наблюдения	28

<i>Дисперсионный анализ</i>					<i>Значимость F</i>
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	
Регрессия	6	31222,94	5203,824	1190,352	1,82E-25
Остаток	22	96,17671	4,371669		
Итого	28	31319,12			

	<i>Коэффициенты</i>	<i>Стандартная ошибка</i>	<i>t-статистика</i>	<i>P-значение</i>	<i>Нижние 95%</i>	<i>Верхние 95%</i>
Y-пересечение	0	#Н/Д	#Н/Д	#Н/Д	#Н/Д	#Н/Д
Переменная X 1	-5,30202	2,446452	-2,16723	0,041316	-10,3757	-0,22839
Переменная X 2	20,10911	1,512595	13,29444	5,43E-12	16,97217	23,24604
Переменная X 3	5,288617	1,966483	2,689378	0,013394	1,210376	9,366858
Переменная X 4	25,06542	2,569854	9,753633	1,9E-09	19,73586	30,39497
Переменная X 5	43,27474	1,453606	29,77061	2,86E-19	40,26014	46,28933
Переменная X 6	1,526726	0,268973	5,676137	1,04E-05	0,96891	2,084541

2.1.3. Построение по данным табл. 4.2.1 регрессионного уравнения (со свободным членом) $y = a + b_1x_1 + b_2x_2 + e$ с использованием «Пакета анализа» (см. Вывод итогов 4.2).

2.1.4. Проверка гипотезы об отсутствии фиксированных групповых эффектов с помощью F -статистики

$$F(3, 22) = \frac{0,9969(4-1)}{0,9136(4 \cdot 7 - 4 - 2)} = 198,88.$$

Сравнение расчетного значения F -статистики с табличным $F_{0,95}(3, 22) = 3,05$ позволяет отвергнуть указанную гипотезу.

Следовательно, построенная модель

$$\hat{y} = -5,30i_1 + 20,11i_2 + 5,29i_3 + 25,05i_4 + 43,27x_1 + 1,53x_2, (*)$$

2.2.1.2. Вычисление $[\mathbf{X}'\mathbf{M}_d\mathbf{X}]^{-1}$ с учетом диагональной структуры матрицы \mathbf{M}_d

0,4833	-0,0641
-0,0641	0,0165

2.2.1.3. Вычисление $[\mathbf{X}'\mathbf{M}_d\mathbf{y}]$ с учетом диагональной матрицы \mathbf{M}_d

209,4182
903,6705

2.2.1.4. Вычисление $\hat{\mathbf{b}}$

43,2747
1,5267

2.2.2. Расчет оценок коэффициентов переменных по формуле $\mathbf{a} = [\mathbf{D}'\mathbf{D}]^{-1}\mathbf{D}'(\mathbf{y} - \mathbf{X}\hat{\mathbf{b}})$.

2.2.2.1. Вычисление $[\mathbf{D}'\mathbf{D}]^{-1}$

0,1428	0	0	0
0	0,1428	0	0
0	0	0,1428	0
0	0	0	0,1428

2.2.2.2. Вычисление $(\mathbf{X}\hat{\mathbf{b}})$ и $(\mathbf{y} - \mathbf{X}\hat{\mathbf{b}})$. Оформление результатов в виде табл. 4.2.3.

Таблица 4.2.3

№ п.п.	y	($\mathbf{X}\hat{\mathbf{b}}$)	($\mathbf{y} - \mathbf{X}\hat{\mathbf{b}}$)	№ п.п.	y	($\mathbf{X}\hat{\mathbf{b}}$)	($\mathbf{y} - \mathbf{X}\hat{\mathbf{b}}$)
1	20,76	26,00	-5,24	15	65,01	56,50	8,51
2	28,09	30,04	-1,95	16	69,05	69,71	-0,66
3	32,95	39,50	-6,55	17	73,13	69,40	3,73
4	38,15	44,75	-6,60	18	81,18	76,01	5,17
5	46,78	52,48	-5,70	19	89,24	84,19	5,05
6	55,31	59,78	-4,47	20	97,3	91,96	5,34
7	60,92	67,52	-6,60	21	115,36	105,49	9,87
8	41,08	21,69	19,39	22	91,26	64,83	26,43
9	56,29	36,84	19,45	23	99,84	73,03	26,81
10	68,51	48,16	20,35	24	108,55	84,36	24,19
11	82,72	62,09	20,63	25	117,17	92,74	24,43
12	96,43	75,95	20,48	26	125,81	100,85	24,96
13	110,15	89,10	21,05	27	134,46	110,17	24,29
14	123,86	104,45	19,41	28	143,1	118,74	24,36

2.2.2.3. Вычисление $\mathbf{D}'(\mathbf{y} - \mathbf{X}\hat{\mathbf{b}})$

-37,1142
140,7637
37,02032
175,4579

2.2.2.4. Вычисление \mathbf{a}

-5,3020
20,1091
5,2886
25,0654

2.2.3. Расчет стандартных ошибок коэффициентов модели по формуле $Est.Var[\mathbf{b}] = s^2[\mathbf{X}'\mathbf{M}_d\mathbf{X}]^{-1}$

2.2.3.1. Вычисление \hat{y} по построенному регрессионному уравнению (*), s^2 и оформление результатов расчетов в виде табл. 4.2.4.

Таблица 4.2.4

№ п.п.	y	\hat{y}	№ п.п.	y	\hat{y}
1	20,70	0,00	15	61,78	10,41
2	24,74	11,23	16	74,99	35,34
3	34,19	1,55	17	74,69	2,42
4	39,45	1,68	18	81,29	0,01
5	47,18	0,16	19	89,48	0,06
6	54,48	0,69	20	97,25	0,00
7	62,21	1,68	21	110,78	20,98
8	41,80	0,51	22	89,90	1,85
9	56,95	0,43	23	98,10	3,04
10	68,27	0,06	24	109,42	0,76
11	82,20	0,27	25	117,81	0,41
12	96,05	0,14	26	125,92	0,01
13	109,21	0,89	27	135,24	0,61
14	124,56	0,49	28	143,81	0,50
Сумма квадратов отклонений					96,17671
Остаточная дисперсия					4,371669

2.2.3.2. Вычисление $Est.Var[\mathbf{b}]$

1,4536
0,2689

2.2.4. Расчет стандартных ошибок оценок коэффициентов фиктивных

переменных по формуле $Var[a_i] = \frac{S^2}{T} + \bar{x}_{i\bullet}' Var[\hat{b}] \bar{x}_{i\bullet}$.

2.2.4.1. Вычисление $\bar{x}_{i\bullet}$

	Орловская область	Белгородская область	Воронежская область	Липецкая область
$\bar{x}_{1\bullet}$	0,6743	1,2386	1,4943	1,6686
$\bar{x}_{2\bullet}$	10,8371	5,9029	9,4129	13,0329

2.2.4.2. Вычисление \hat{S}^2

$$\hat{S}^2 = \sqrt{4,3716} = 2,0908.$$

2.2.4.3. Вычисление $Var[a_i]$

2,4465
1,5126
1,9665
2,5699

3. Построение регрессионной модели со случайными эффектами.

3.1. Вычисление средних значений $\bar{y}_{i\bullet}$ для каждой панели данных.

3.2. Расчет квадратов отклонений $(\bar{y}_{i\bullet} - \hat{a} - \hat{b}'\bar{x}_{i\bullet})^2$ и оформление результатов расчетов в виде табл. 4.2.5

Таблица 4.2.5

$\bar{y}_{i\bullet}$	$\bar{x}_{1\bullet}$	$\bar{x}_{2\bullet}$	$a + b'\bar{x}_{i\bullet}$	$(\bar{y}_{i\bullet} - a - b'\bar{x}_{i\bullet})^2$
40,4229	0,6743	10,8371	47,1698	45,5213
82,7200	1,2386	5,9029	77,7750	24,4533
84,3243	1,4943	9,4129	94,0451	94,4945
117,1700	1,6686	13,0329	105,6473	132,7736
Сумма квадратов отклонений				297,2428

3.3. Вычисление \hat{S}_u^2

$$\hat{S}_u^2 = 297,2428 - \frac{4,3716}{7} = 296,6183.$$

3.4. Расчет \hat{q}

$$\hat{q} = 1 - \left(\frac{4,3716}{7 \cdot 296,6183 + 4,3716} \right)^{0,5} = 0,9541.$$

3.4. Преобразование данных i -го блока \mathbf{y}_i исходных данных по формуле

$$\Omega^{-1/2} \mathbf{y}_i = \frac{1}{s_e} \begin{bmatrix} y_{i1} - q \bar{y}_i \\ y_{i2} - q \bar{y}_i \\ \mathbf{M} \\ y_{iT} - q \bar{y}_i \end{bmatrix}$$

и аналогично для \mathbf{X}_i . Оформление результатов расчетов в виде табл. 4.2.6.

Таблица 4.2.6.

№ п.п.	y^*	x_1^*	x_2^*	№ п.п.	y^*	x_1^*	x_2^*
1	-4,0740	-0,0923	-0,0253	15	-3,5339	-0,1111	0,3153
2	-2,3972	-0,0763	0,1257	16	-2,6098	-0,0494	0,5441
3	-1,2855	-0,0214	-0,0138	17	-1,6765	-0,0311	-0,0209
4	-0,0961	0,0061	-0,0047	18	0,1649	0,0147	-0,3274
5	1,8780	0,0427	0,1166	19	2,0086	0,0559	-0,2680
6	3,8292	0,0770	0,2378	20	3,8523	0,0947	-0,2062
7	5,1125	0,1136	0,3590	21	7,9834	0,1359	0,6539
8	-8,6576	-0,1674	-0,9567	22	-4,6983	-0,1080	-0,3924
9	-5,1784	-0,0919	-0,8263	23	-2,7356	-0,0691	-0,2666
10	-2,3831	-0,0462	-0,4260	24	-0,7433	-0,0234	0,1337
11	0,8673	0,0133	-0,0257	25	1,2285	0,0178	0,2229
12	4,0034	0,0705	0,4295	26	3,2049	0,0590	0,2710
13	7,1418	0,1277	0,7795	27	5,1835	0,1025	0,4357
14	10,2779	0,1849	1,4589	28	7,1599	0,1436	0,5523

3.5. Построение по данным табл. 4.2.6 регрессионного уравнения с помощью «Пакета анализа» (см. Вывод итогов 4.3).

4. Проверка наличия случайных эффектов с помощью теста Хаусмена.

4.1. Формирование векторов $(\mathbf{x}_1^* - \bar{\mathbf{x}}_1^*)$, $(\mathbf{x}_2^* - \bar{\mathbf{x}}_2^*)$ и оформление результатов расчетов в виде табл. 4.2.7.

4.2. Нахождение матрицы $\hat{\Sigma}^{-1} = \left(\mathbf{X}^{**'} \mathbf{X}^{**} \right)^{-1}$ по данным табл. 4.2.7.

9,1818	-1,2150
-1,2150	0,3140

4.3. Вычисление матрицы $s^{2*} \left(\mathbf{X}^{**'} \mathbf{X}^{**} \right)^{-1}$, где s^{2*} - остаточная дисперсия по преобразованным данным ($s^{2*} = 0,2191$)

2,0114	-0,2662
-0,2662	0,0688

ВЫВОД ИТОГОВ 4.3

Регрессионная статистика	
Множественный R	0,995042
R-квадрат	0,990109
Нормированный R-квадрат	0,989318
Стандартная ошибка	0,468049
Наблюдения	28

Дисперсионный анализ

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>
Регрессия	2	548,2222	274,1111	1251,252	8,72E-26
Остаток	25	5,476736	0,219069		
Итого	27	553,699			

	<i>Коэффициенты</i>	<i>Стандартная ошибка</i>	<i>t-статистика</i>	<i>P-значение</i>	<i>Нижние 95%</i>	<i>Верхние 95%</i>
У-пересечение	0,118466	0,090433	1,309978	0,202113	-0,06779	0,304717
Переменная X 1	43,37632	1,418254	30,58432	2,5E-21	40,45537	46,29726
Переменная X 2	1,512727	0,262271	5,767808	5,19E-06	0,972571	2,052884

Таблица 4.2.7

№ п.п.	$(\mathbf{x}_1^* - \bar{\mathbf{x}}_1^*)$	$(\mathbf{x}_2^* - \bar{\mathbf{x}}_2^*)$	№ п.п.	$(\mathbf{x}_1^* - \bar{\mathbf{x}}_1^*)$	$(\mathbf{x}_2^* - \bar{\mathbf{x}}_2^*)$
1	-0,1056	-0,1280	15	-0,1244	0,2126
2	-0,0896	0,0230	16	-0,0627	0,4414
3	-0,0347	-0,1165	17	-0,0444	-0,1236
4	-0,0072	-0,1074	18	0,0014	-0,4301
5	0,0294	0,0139	19	0,0426	-0,3707
6	0,0637	0,1351	20	0,0814	-0,3089
7	0,1003	0,2563	21	0,1226	0,5512
8	-0,1807	-1,0594	22	-0,1213	-0,4951
9	-0,1052	-0,9290	23	-0,0824	-0,3693
10	-0,0595	-0,5287	24	-0,0367	0,0310
11	0,0000	-0,1284	25	0,0045	0,1202
12	0,0572	0,3268	26	0,0457	0,1682
13	0,1144	0,6768	27	0,0892	0,3329
14	0,1716	1,3562	28	0,1303	0,4496

4.4. Определение вектора $\begin{bmatrix} \hat{\mathbf{b}} - \hat{\mathbf{b}} \end{bmatrix}$, где $\hat{\mathbf{b}}$ - вектор оценок параметров уравнения регрессии, построенного по исходным данным

(без свободного члена), $\hat{\mathbf{b}}$ - вектор оценок параметров уравнения регрессии, построенного по преобразованным данным (без свободного члена). Оформление результатов в виде табл. 4.2.8.

Таблица 4.2.8

$\hat{\mathbf{b}}$	$\hat{\mathbf{b}}$	$[\hat{\mathbf{b}} - \hat{\mathbf{b}}]$
57,8908	43,3763	14,5145
0,4178	1,5127	-1,0949

4.5. Расчет критерия Вальда

$$W = c^2[2] = [\hat{\mathbf{b}} - \hat{\mathbf{b}}]' \hat{\Sigma}^{-1} [\hat{\mathbf{b}} - \hat{\mathbf{b}}] = 432,29.$$

Сравнение расчетного значения $c^2[2] = 432,29$ с табличным $c_{0,95}^2[2] = 5,99$ позволяет отвергнуть нулевую гипотезу об отсутствии случайных эффектов, порождаемых структурой панельных данных, и принять альтернативную гипотезу, в соответствии с которой в модели необходимо учитывать случайные эффекты, а для нахождения параметров этой модели применять обобщенный МНК.

5. Выбор наилучшей модели (с фиксированными или случайными эффектами) по остаточной дисперсии.

5.1. Получение расчетных значений \hat{y}_{it}^*

$$\hat{y}_{it}^* = \hat{b}_0 + \hat{b}_1 x_{1it}^* + \hat{b}_2 x_{2it}^* = 0,1185 + 43,3763 x_{1it}^* + 1,5127 x_{2it}^*,$$

$$\hat{y}_{it} = \hat{y}_{it}^* \hat{S}_e + \hat{q} \bar{y}_i = 4,3717 \hat{y}_{it}^* + 0,9542 \bar{y}_i.$$

5.2. Расчет остаточной дисперсии. (Заметим, что число степеней свободы для модели случайных эффектов равно 26, а для модели с фиксированными эффектами – 22).

Оформление результатов вычислений в виде табл. 4.2.9.

Со статистической точки зрения модель со случайными эффектами лучше, поскольку обладает меньшей остаточной дисперсией.

Таблица 4.2.9

№ п.п.	y	\hat{y}	$(y - \hat{y})^2$	\hat{y}	$(y - \hat{y})^2$
1	20,76	21,4238	0,4406	20,7023	0,0033
2	28,09	25,4586	6,9245	24,7392	11,2279
3	32,95	34,9461	3,9844	34,1938	1,5471
4	38,15	40,2118	4,2509	39,4479	1,6844
5	46,78	47,9537	1,3776	47,1810	0,1608
6	55,31	55,2619	0,0023	54,4814	0,6867
7	60,92	63,0039	4,3425	62,2145	1,6757
8	41,08	41,3768	0,0881	41,7965	0,5134
9	56,29	56,5533	0,0693	56,9474	0,4322
10	68,51	67,8758	0,4022	68,2741	0,0556
11	82,72	81,8009	0,8447	82,1973	0,2732
12	96,43	95,6553	0,6001	96,0542	0,1412
13	110,15	108,8139	1,7852	109,2087	0,8860
14	123,86	124,1508	0,0845	124,5618	0,4925
15	65,01	61,9906	9,1169	61,7837	10,4087
16	69,05	75,2149	38,0061	74,9946	35,3389
17	73,13	74,9486	3,3072	74,6856	2,4199
18	81,18	81,5968	0,1737	81,2948	0,0132
19	89,24	89,7978	0,3112	89,4812	0,0582
20	97,3	97,5802	0,0785	97,2501	0,0025
21	115,36	111,0758	18,3541	110,7800	20,9763
22	91,26	89,2446	4,0617	89,8996	1,8506
23	99,84	97,4506	5,7092	98,0960	3,0415
24	108,55	108,7731	0,0498	109,4227	0,7617
25	117,17	117,1708	0,0000	117,8076	0,4066
26	125,81	125,2963	0,2639	125,9177	0,0116
27	134,46	134,6269	0,0279	135,2391	0,6070
28	143,1	143,2062	0,0113	143,8072	0,5001
Сумма квадратов отклонений расчетных от фактических значений					
104,6686				96,17671	
Остаточная дисперсия					
4,0257				4,3717	

4.3. Задание для самостоятельной работы

Задание 4.3.1. Отделу труда и заработной платы завода ОАО «Тяжмехпресс» было поручено провести исследование факторов, существенно влияющих на среднемесячный размер оплаты труда, выплачиваемой рабочим этого предприятия (y , руб.). В результате исследования удалось выяснить, что такими факторами являются: 1) процент перевыполнения месячного плана (x_1); 2) разряд рабочего (x_2). На основании этого результата

по данным четырех цехов была сформирована табл. 4.3.1. Выполните следующие задания: 1) постройте модель с фиксированными эффектами: а) с помощью фиктивных переменных; б) с использованием процедуры раздельного оценивания коэффициентов; 2) постройте модель со случайными эффектами; 3) выберите из построенных моделей наиболее подходящую для аналитических целей.

Таблица 4.3.1

№ п.п.	У	x_1	x_2	№ п.п.	У	x_1	x_2
<i>Первый цех</i>				<i>Третий цех</i>			
1	2170	26,4	6	21	2520	13,4	10
2	1000	17,3	3	22	4190	29,7	10
3	3350	23,8	8	23	3130	21,6	6
4	2200	17,6	5	24	2780	25,1	7
5	1760	26,2	5	25	1320	14,1	2
6	1610	21,1	1	26	2790	24,1	6
7	1900	17,5	3	27	3330	10,5	9
8	1810	22,9	5	28	2050	22,1	2
9	1490	22,9	4	29	1670	17,0	2
10	2320	14,9	7	30	2070	20,5	2
<i>Второй цех</i>				<i>Четвертый цех</i>			
11	2180	19,6	3	31	1530	14,2	4
12	2100	22,8	2	32	2590	18,0	10
13	3380	27,8	6	33	2190	29,9	2
14	1800	14,0	4	34	1550	14,1	5
15	1220	11,4	2	35	1670	18,4	6
16	3000	16,0	9	36	2610	20,1	8
17	2170	28,8	1	37	3480	27,6	9
18	2490	16,8	5	38	2260	27,4	5
19	2190	11,8	6	39	2900	28,5	8
20	2360	18,6	5	40	3710	28,6	9

5. МОДЕЛИ РАСПРЕДЕЛЕННЫХ ЛАГОВ

5.1. Расчетные формулы

5.1.1. Модель с конечным числом лагов

$$y_t = a_0 + b_0 x_t + b_1 x_{t-1} + \mathbf{L} + b_l x_{t-l} + e_t$$

и с бесконечным числом лагов

$$y_t = a_0 + b_0 x_t + b_1 x_{t-1} + b_2 x_{t-2} + \mathbf{L} + e_t,$$

где b_0 - краткосрочный мультипликатор;

$\sum_j b_j$ - долгосрочный мультипликатор.

5.1.2. Замена лаговых переменных одной интегрированной

$$z_t = x_t + l x_{t-1} + l^2 x_{t-2} + l^3 x_{t-3} + \mathbf{L} + l^p x_{t-p}$$

и преобразование исходной модели в уравнение вида

$$y_t = a_0 + b_0 z_t + e_t.$$

5.1.3. Уравнение, полученное в результате преобразования Койка,

$$y_t = (1 - l) a_0 + b_0 x_t + l y_{t-1} + v_t,$$

где $v_t = e_t - l e_{t-1}$ - скользящая средняя.

5.1.4. Модель с распределенными лагами, согласно методу Алмон, может быть представлена в виде регрессионной модели

$$y_t = a_0 + c_0 z_0 + c_1 z_1 + c_2 z_2 + \mathbf{L} + c_k z_k + e_t,$$

где $z_0 = \sum_{j=0}^l x_{t-j}$, $z_1 = \sum_{j=1}^l j x_{t-j}$, $z_2 = \sum_{j=1}^l j^2 x_{t-j}$, ..., $z_k = \sum_{j=1}^l j^k x_{t-j}$.

5.1.5. Каждый из коэффициентов исходной лаговой модели вычисляется следующим образом:

$$b_0 = c_0$$

$$b_1 = c_0 + c_1 + c_2 + \mathbf{L} + c_k$$

$$b_2 = c_0 + 2c_1 + 4c_2 + \mathbf{L} + 2^k c_k$$

$$b_3 = c_0 + 3c_1 + 9c_2 + \mathbf{L} + 3^k c_k$$

$$\mathbf{L} \mathbf{L}$$

$$b_l = c_0 + l c_1 + l^2 c_2 + \mathbf{L} + l^k c_k.$$

5.2. Решение типовых задач

Задание 5.2.1. Компания «Автоматика», как и любая другая компания, желающая добиться успеха в мире современного бизнеса, с целью увеличения своей прибыли периодически проводит маркетинговые исследования, ориентированные на выявление изменений в предпочтениях потребителей, а также анализ динамики рыночной конъюнктуры. Данные, отражающие зависимость прибыли компании от расходов на маркетинговые исследования, представлены в табл. 5.2.1. Руководство этой компании

заинтересовано в получении ответа вопрос: какой эффект дает дополнительное вложение в маркетинговые исследования одной тысячи рублей и каков средний лаг, существующий между вложением средств в маркетинговые исследования и получением прибыли от этих вложений.

Таблица 5.2.1

Период	Прибыль компании, тыс. руб.	Расходы на маркетинговые исследования, тыс. руб.	Период	Прибыль компании, тыс. руб.	Расходы на маркетинговые исследования, тыс. руб.
1	988	60	11	1281	78
2	1035	66	12	1253	67
3	1089	73	13	1302	74
4	1082	67	14	1382	94
5	1073	54	15	1426	93
6	1126	65	16	1468	91
7	1177	75	17	1513	93
8	1234	83	18	1593	96
9	1265	83	19	1612	98
10	1258	74	20	1628	93

Решение с помощью табличного процессора Excel

1. Ввод исходных данных.
2. Преобразование исходных данных x_t в новые переменные z_{0t} , z_{1t} и z_{2t} по следующим формулам:

$$z_{0t} = x_t + x_{t-1} + x_{t-3} + x_{t-4};$$

$$z_{1t} = x_{t-1} + 2x_{t-2} + 3x_{t-3} + 4x_{t-4};$$

$$z_{2t} = x_{t-1} + 4x_{t-2} + 9x_{t-3} + 16x_{t-4};$$

и оформление результатов в виде табл. 5.2.2.

3. Расчет параметров линейной регрессии \hat{a} , \hat{c}_0 , \hat{c}_1 , \hat{c}_2 по преобразованным данным с помощью пакета «Анализ данных» (см. Вывод итогов 5.1).

Таким образом, построенная по преобразованным данным модель записывается в следующем виде:

$$\hat{y}_t = 31,6184 + 5,9333z_{0t} - 3,1978z_{1t} + 0,6317z_{2t}.$$

4. Расчет оценок коэффициентов регрессии \hat{a} , \hat{b}_0 , \hat{b}_1 , \hat{b}_2 исходной модели.

Таблица 5.2.2

y	z_0	z_1	z_2
988			
1035			
1089			
1082			
1073	320	651	1913
1126	325	671	2035
1177	334	666	2052
1234	344	635	1893
1265	360	644	1832
1258	380	734	2130
1281	393	789	2353
1253	385	807	2449
1302	376	777	2373
1382	387	738	2228
1426	406	755	2241
1468	419	771	2207
1513	445	855	2493
1593	467	930	2798
1612	471	927	2775
1628	471	933	2775

ВЫВОД ИТОГОВ 5.1

Регрессионная статистика

Множественный R	0,992168
R-квадрат	0,984398
Нормированный R-квадрат	0,980498
Стандартная ошибка	24,31472
Наблюдения	16

Дисперсионный анализ

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>
Регрессия	3	447623,5	149207,8	252,3789	4,2E-11
Остаток	12	7094,468	591,2057		
Итого	15	454717,9			

	<i>Коэффициенты</i>	<i>Стандартная ошибка</i>	<i>t-статистика</i>	<i>P-значение</i>	<i>Нижние 95%</i>	<i>Верхние 95%</i>
Y-пересечение	31,6184	49,35158	0,640676	0,533777	-75,9095	139,1463
Переменная X 1	5,933285	0,532509	11,14213	1,1E-07	4,773047	7,093522
Переменная X 2	-3,19784	0,757169	-4,22341	0,001182	-4,84757	-1,54811
Переменная X 3	0,63166	0,184014	3,43267	0,004961	0,230728	1,032593

$$\hat{a} = 31,6184;$$

$$\hat{b}_0 = \hat{c}_0 = 5,9333;$$

$$\hat{b}_1 = \hat{c}_0 + \hat{c}_1 + \hat{c}_2 = 5,9333 - 3,1978 + 0,6317 = 3,3671;$$

$$\hat{b}_2 = \hat{c}_0 + 2\hat{c}_1 + 4\hat{c}_2 = 5,9333 - 2 \cdot 3,1978 + 4 \cdot 0,6317 = 2,0642;$$

$$\hat{b}_3 = \hat{c}_0 + 3\hat{c}_1 + 9\hat{c}_2 = 5,9333 - 3 \cdot 3,1978 + 9 \cdot 0,6317 = 2,0247;$$

$$\hat{b}_4 = \hat{c}_0 + 4\hat{c}_1 + 16\hat{c}_2 = 5,9333 - 4 \cdot 3,1978 + 16 \cdot 0,6317 = 3,2485.$$

Следовательно, модель с распределенным лагом имеет вид

$$\hat{y}_t = 31,6184 + 5,9333x_t + 3,3671x_{t-1} + 2,0642x_{t-2} + 2,0247x_{t-3} + 3,2485x_{t-4}.$$

5. Расчет долгосрочного мультипликатора

$$b = 5,9333 + 3,3671 + 2,0642 + 2,0247 + 3,2485 = 16,6379.$$

Мультипликатор показывает, что увеличение средств на проведение маркетинговых исследований на 1 тыс. руб. в настоящий момент времени через 4 периода приведет к увеличению прибыли на 16637 руб.

6. Расчет относительных коэффициентов регрессии

$$b_0 = \hat{b}_0 / b = 5,9333 / 16,6379 = 0,3566;$$

$$b_1 = \hat{b}_1 / b = 3,3671 / 16,6379 = 0,2024;$$

$$b_2 = \hat{b}_2 / b = 2,0642 / 16,6379 = 0,1241;$$

$$b_3 = \hat{b}_3 / b = 2,0247 / 16,6379 = 0,1217;$$

$$b_4 = \hat{b}_4 / b = 3,2485 / 16,6379 = 0,1952.$$

7. Расчет среднего лага

$$T = 0 \cdot 0,3566 + 1 \cdot 0,2024 + 2 \cdot 0,1241 + 3 \cdot 0,1217 + 4 \cdot 0,1952 = 1,5966.$$

Таким образом, в среднем увеличении затрат на маркетинговые исследования приведет к увеличению прибыли компании через 1,6 периода.

Задание 5.2.2. Администрация области проводит комплексное исследование социально-экономической ситуации в регионе с целью выработки обоснованной политики его развития. Одной из поставленных задач исследования является определение среднего срока воздействия инфляции на реальные доходы населения. Для проведения необходимых для решения этой задачи расчетов была сформирована табл. 5.2.3. Специалисты выдвигают

нули гипотезу о том, что годовой уровень инфляции оказывает воздействие на реальные доходы населения с бесконечным временным лагом, который имеет геометрическую структуру.

Таблица 5.2.3

Год	Уровень инфляции, %	Реальные доходы населения, млн. руб.	Год	Уровень инфляции, %	Реальные доходы населения, млн. руб.
1	13,9	1704	11	10,6	2198
2	12,6	1749	12	9,6	2297
3	11,8	1821	13	8,5	2407
4	13,3	1870	14	9,2	2468
5	13,2	1869	15	8,8	2455
6	12,0	1927	16	7,4	2498
7	10,4	2020	17	6,8	2444
8	8,7	2125	18	6,5	2472
9	10,8	2111	19	5,9	2692
10	11,8	2094			

Решение с помощью табличного процессора Excel

1. Ввод исходных данных.
2. Формирование вектора y_{t-1} и оформление результатов в виде табл. 5.2.4.

Таблица 5.2.4

t	x_t	y_{t-1}	y_t	t	x_t	y_{t-1}	y_t
2	12,6	1704	1749	11	10,6	2094	2198
3	11,8	1749	1821	12	9,6	2198	2297
4	13,3	1821	1870	13	8,5	2297	2407
5	13,2	1870	1869	14	9,2	2407	2468
6	12,0	1869	1927	15	8,8	2468	2455
7	10,4	1927	2020	16	7,4	2455	2498
8	8,7	2020	2125	17	6,8	2498	2444
9	10,8	2125	2111	18	6,5	2444	2472
10	11,8	2111	2094	19	5,9	2472	2692

3. Расчет параметров модели двухфакторной линейной авторегрессии

$$y_t = (1 - I)a + b_0 x_t + I y_{t-1} + v_t$$

по данным табл. 5.2.4 с помощью пакета «Анализ данных» (см. Вывод итогов 5.2).

Таким образом, уравнение, полученное в результате преобразования Койка, имеет вид

$$y_t = 885,76 - 31,94x_t + 0,76y_{t-1} + v_t.$$

ВЫВОД ИТОГОВ 5.2

Регрессионная статистика	
Множественный R	0,981887
R-квадрат	0,964102
Нормированный R-квадрат	0,959315
Стандартная ошибка	56,84734
Наблюдения	18

Дисперсионный анализ					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	Значимость <i>F</i>
Регрессия	2	1301856	650928	201,4247	1,46E-11
Остаток	15	48474,3	3231,62		
Итого	17	1350330			

	Коэффициенты	Стандартная ошибка	t-статистика	P-значение	Нижние 95%	Верхние 95%
Y-пересечение	885,7605	340,4847	2,601469	0,02004	160,0341	1611,487
Переменная X 1	-31,9375	12,53875	-2,5471	0,022326	-58,6632	-5,21172
Переменная X 2	0,759298	0,105577	7,191857	3,11E-06	0,534265	0,984331

4. Вычисление параметров исходной модели

$$y_t = a + b_0 x_t + b_1 x_{t-1} + b_2 x_{t-2} + \mathbf{K} + e_t.$$

$$\hat{a} = \frac{885,76}{1 - I} = \frac{885,76}{1 - 0,76} = 3679,90; \quad \hat{b}_0 = -31,94;$$

$$\hat{b}_1 = \hat{b}_0 I = -31,94 \cdot 0,76 = -24,25;$$

$$\hat{b}_2 = \hat{b}_0 I^2 - 31,94 \cdot 0,76^2 = -18,41 \text{ и т.д.}$$

Таким образом, модель с бесконечным числом лаговых переменных в рассматриваемом случае записывается следующим образом:

$$y_t = 3679,90 - 31,94x_t - 24,25x_{t-1} - 18,41x_{t-2} - \mathbf{K} - e_t$$

5. Расчет среднего лага модели

$$T = \frac{I}{1 - I} = \frac{0,76}{1 - 0,76} = 3,15.$$

Следовательно, очередной рост цен в среднем оказывает инфляционное воздействие на реальные доходы населения в течение отрезка времени, равного 3,15 года.

5.3. Задания для самостоятельной работы

Задание 5.3.1. Совет директоров крупной компании «Эксклюзив», имеющей возможности для увеличения степени компьютеризации управ

ления производством, для реализации своих стратегических планов желал бы иметь представление о том, на сколько и когда могут снизиться производственные затраты (y , тыс. руб.) при росте степени компьютеризации (x , %) на 1% в текущем периоде. Очевидно, что для ответа на этот вопрос целесообразно воспользоваться регрессионной моделью с распределенными лагами. Постройте такого рода модель с лагом, равным четырем, в предположении, что структура лага описывается полиномом третьей степени. Данные для построения модели представлены в табл. 5.3.1.

Таблица 5.3.1

Год	y	x	Год	y	x
1	542	53,5	11	419	70,1
2	537	59,1	12	421	60,0
3	531	65,7	13	411	66,6
4	504	60,3	14	392	84,2
5	489	48,6	15	375	82,8
6	475	57,8	16	357	81,7
7	460	66,7	17	360	83,2
8	434	73,8	18	363	85,9
9	417	74,4	19	345	87,7
10	427	66,0	20	329	82,7

Задание 5.3.2.

Таблица 5.3.2

Год	y	x	Год	y	x
1	2160,57	0,54	11	1794,86	1,17
2	2141,29	0,59	12	1809,43	1,31
3	2118,86	0,65	13	1821,43	1,20
4	2094,86	0,72	14	1731,43	0,96
5	2141,14	0,75	15	1651,71	1,15
6	2104,29	0,82	16	1602,00	1,33
7	2115,43	0,97	17	1602,86	1,47
8	2063,14	1,02	18	1560,86	1,47
9	1968,86	0,94	19	1499,14	1,31
10	1882,43	1,06	20	1460,57	1,40

Департамент экономического развития городской администрации проводит мониторинг социально-экономического развития региона. В частности, исследуется проблема взаимосвязи среднемесячной реальной заработной платы (y , руб.) и уровня регистрируемой безработицы (x , %). Рассчитайте по данным табл. 5.3.2 средний срок воздействия первого из

указанного фактора на другой, предварительно построив для этого модель распределенных лагов.

6. РЕКУРРЕНТНЫЙ МНК

6.1. Расчетные формулы

6.1.1. Рекуррентная формула пересчета коэффициентов регрессии

$$\hat{\mathbf{b}}_n = \hat{\mathbf{b}}_{n-1} + \frac{\mathbf{C}_{n-1}^{-1} \mathbf{x}'_n}{\mathbf{x}_n \mathbf{C}_{n-1}^{-1} \mathbf{x}'_n + 1} [\mathbf{y}_n - \mathbf{x}_n \hat{\mathbf{b}}_{n-1}],$$

где $\mathbf{C}_n^{-1} = (\mathbf{X}'_n \mathbf{X}_n)^{-1}$.

6.1.2. Формула Шермана-Моррисона для рекуррентного обращения матриц

$$(\mathbf{C}_{n-1} + \mathbf{x}'_n \mathbf{x}_n)^{-1} = \mathbf{C}_{n-1}^{-1} - \frac{\mathbf{C}_{n-1}^{-1} \mathbf{x}'_n \mathbf{x}_n \mathbf{C}_{n-1}^{-1}}{\mathbf{x}_n \mathbf{C}_{n-1}^{-1} \mathbf{x}'_n + 1}.$$

6.2. Решение типовой задачи

Задание 6.2.1. Агентство «Земной дар», занимающееся куплей-продажей земельных участков, постоянно осуществляет мониторинг цен на эти участки. Для выработки стратегии, которой необходимо придерживаться, руководство агентства решило выяснить механизм формирования цен на земельные участки. В ходе исследования было выявлено, что цена (USD, y), главным образом, зависит от таких факторов как площадь участка (сотки, x_1), доля площади участка, занимаемая лесом, (x_2) а также расстояния от участка до железнодорожного вокзала (км., x_3). Для того чтобы получить количественные оценки влияния этих факторов на уровень цен, было решено использовать регрессионный анализ. Данные о цене и факторах, влияющих на ее величину, представлены в табл. 6.2.1. Таблица продолжает пополняться данными по мере совершения очередной сделки. Поэтому модель, отражающую зависимость цены от соответствующих факторов, приходилось все время пересчитывать. Для того чтобы не тратить на это время, руководству агентства посоветовали воспользоваться рекуррентным МНК. Убедитесь в том, что рекуррентный метод приводит к тем же самым результатам, которые получаются при использовании обычного МНК.

Таблица 6.2.1

№ п.п.	y	x_1	x_2	x_3	№ п.п.	y	x_1	x_2	x_3
1.	5556	36,0	1	12,1	16.	4537	38,0	1	18,0
2.	5236	38,2	1	12,1	17.	3700	5,0	0	17,2
3.	5952	21,0	1	12,0	18.	2020	5,0	1	34,2
4.	7000	40,0	0	16,0	19.	5000	3,5	0	11,1
5.	3750	40,0	0	15,5	20.	4764	23,7	0	14,2
6.	7000	20,0	0	13,7	21.	8071	23,7	1	14,2
7.	5952	21,0	0	14,5	22.	3500	20,0	1	11,1
8.	2009	65,0	0	16,1	23.	8156	5,0	1	14,7
9.	2583	60,0	1	15,2	24.	4764	30,0	0	12,1
10.	2449	56,0	0	15,5	25.	9568	3,8	1	14,8
11.	2500	40,0	0,5	15,2	26.	9873	7,9	1	14,8
12.	3000	13,0	0	15,5	27.	5175	40,0	0,25	14,2
13.	3704	27,0	0	13,5	28.	3977	8,8	0	11,4
14.	3500	10,0	0	15,5	29.	5500	10,0	0,2	18,5
15.	3500	20,0	0	17,5	30.	7500	8,0	0	16,5

Решение с помощью табличного процессора Excel

1. Ввод исходных данных.
2. Формирование дополнительной переменной x_0 , принимающей единственное значение, равное 1.
3. Построение модели с помощью матричного МНК для $n = \overline{1, 29}$, используя для этого следующие функции Excel: **ТРАНСП**, **МОБР**, **МУМНОЖ**.

- 3.1. Вычисление матрицы, обратной к матрице системы нормальных уравнений $(X'X)^{-1}$

0,65373	-0,00407	-0,03724	-0,03300
-0,00407	0,00012	0,00013	0,00007
-0,03724	0,00013	0,15789	-0,00205
-0,03300	0,00007	-0,00205	0,00212

- 3.2. Расчет правой части системы нормальных уравнений $(X'y)$

142296,00
3174480,40
68695,75
2090740,40

- 3.3. Получение вектора оценок коэффициентов $\hat{b} = (X'X)^{-1} X'y$

8540,298
-52,865
1677,008
-196,950

4. Пересчет модели в связи с появлением нового наблюдения

$$y_{30} = 7500 \text{ и } \mathbf{x}_{30} = (1; 8; 0; 16,5).$$

4.1. Пересчет модели с помощью обычного МНК

4.1.1. Вычисление матрицы, обратной к матрице системы нормальных уравнений

0,64839	-0,00393	-0,03236	-0,03317
-0,00393	0,00012	0,00001	0,00007
-0,03236	0,00001	0,15343	-0,00190
-0,03317	0,00007	-0,00190	0,00211

4.1.2. Расчет правой части системы нормальных уравнений

149796,00
3234480,40
68695,75
2214490,40

4.1.3. Получение вектора оценок коэффициентов модели

8723,473
-57,701
1509,564
-191,009

4.2. Пересчет модели с помощью рекуррентного МНК

4.2.1. Вычисление вектора корректирующих коэффициентов

$$(\mathbf{C}_n^{-1} \mathbf{x}'_{n+1})$$

0,07665
-0,00202
-0,07007
0,00249

4.2.2. Расчет величины нормирующего коэффициента

$$(\mathbf{x}_{n+1} \mathbf{C}_n^{-1} \mathbf{x}'_{n+1} + 1)$$

1,10148

4.2.3. Нормировка корректирующих коэффициентов

$$\frac{(\mathbf{C}_n^{-1} \mathbf{x}'_{n+1})}{(\mathbf{x}_{n+1} \mathbf{C}_n^{-1} \mathbf{x}'_{n+1} + 1)}$$

0,06959
-0,00184
-0,06361
0,00226

4.2.4. Вычисление прогнозной ошибки, возникающей вследствие использования нескорректированной модели

$$(y_{n+1} - \hat{\mathbf{b}}_n \mathbf{x}_{n+1})$$

2632,297

4.2.5. Корректировка коэффициентов модели

$$\hat{\mathbf{b}}_{n+1} = \hat{\mathbf{b}}_n + \frac{(\mathbf{C}_n^{-1} \mathbf{x}'_{n+1})}{(\mathbf{x}_{n+1} \mathbf{C}_n^{-1} \mathbf{x}'_{n+1} + 1)} (y_{n+1} - \hat{\mathbf{b}}_n \mathbf{x}_{n+1})$$

8723,473

-57,701

1509,564

-191,009

5. Сравнение полученных результатов

Вывод: коэффициенты регрессионной модели, полученные с помощью обычного и рекуррентного МНК, идентичны.

6.3. Задания для самостоятельной работы

Задание 6.3.1. Группа бизнесменов планирует создать сеть баз для зимнего отдыха в горах Северного Кавказа. Одной из важных проблем, которую им необходимо решить для реализации этого проекта, является определение оптимальной стоимости одного дня пребывания на каждой базе отдыха. Очевидно, что стоимость должна устанавливаться в зависимости от целого ряда факторов. Для того чтобы принять наиболее обоснованное решение, было проведено исследование, в ходе которого удалось собрать информацию о 25 уже существующих горнолыжных лагерях. Кроме того, было выявлено, что стоимость (руб., у) обычно складывается под влиянием следующих факторов: общей площади территории лагеря (x_1), количества жилых помещений (x_2), наличия сауны (x_3), наличия плавательного бассейна (x_4), наличия канатных подъемников (x_5) и наличия дополнительных мест проведения досуга (бара, кафе, дискотеки, бильярда и др.) (x_6). Данные о стоимости одного дня пребывания в лагере и соответствующих факторах представлены в табл. 6.3.1. Исследование еще не закончилось, поэтому таблица продолжает пополняться новыми данными. Не

давно стали известны данные еще о двух лагерях: 1) $y_{26} = 985$, $\mathbf{x}_{26} = (27; 125; 1; 0; 1; 3)$; 2) $y_{27} = 875$, $\mathbf{x}_{27} = (30; 120; 0; 0; 1; 2)$. Постройте модель, отражающую зависимость стоимости от соответствующих факторов, 1) обычным МНК для $n = \overline{1, 27}$; 2) обычным МНК для $n = \overline{1, 25}$ с пересчетом коэффициентов модели по формулам рекуррентного МНК в связи с появлением новых наблюдений. Сравните полученные результаты.

Таблица 6.3.1

Лагерь	y	x_1	x_2	x_3	x_4	x_5	x_6
1	700	40	32	0	0	0	2
2	850	20	47	1	0	1	2
3	900	45	18	1	1	1	1
4	800	110	32	1	0	1	3
5	800	30	54	1	0	1	2
6	700	50	30	1	0	1	3
7	775	35	30	1	0	1	2
8	800	18	40	1	0	1	1
9	850	23	60	1	1	1	1
10	850	9	60	1	0	1	3
11	900	52	50	1	1	1	2
12	700	25	21	0	0	1	1
13	900	250	30	1	0	1	2
14	850	140	70	1	1	1	2
15	900	120	80	1	1	1	1
16	750	60	50	1	1	1	2
17	850	120	35	1	0	1	2
18	900	173	25	1	1	1	2
19	800	100	75	1	0	1	2
20	950	134	35	1	1	1	1
21	750	114	120	0	1	1	2
22	750	2	17	0	0	1	2
23	750	32	15	0	1	0	3
24	900	25	30	1	1	1	2
25	750	66	100	1	0	1	2

Задания 6.3.2. Руководство фирмы «Доступное чтение», осуществляющей продажу художественной и научной литературы через ряд магазинов Воронежской области, решило начать освоение рынка электронной коммерции. У фирмы уже есть свой сайт, который пока имеет только рекламный характер; торговля через него не осуществляется. С целью выработки стратегии успешного осуществления интернет-бизнеса было решено провести маркетинговое исследование, ориентированное на изучение фак

торов, наиболее существенным образом влияющих на объем продаж книг Интернет-магазинами. В результате исследования было выявлено, что объем продаж (руб.) зависит от затрат на поддержание сайта (руб.), средней цены продаваемой книги (руб.), категории литературы (художественная – 0, научная – 1), наличия обложки и аннотации книги на сайте (нет – 0, есть – 1). Для построения модели, отражающей указанную зависимость, было использовано 1500 наблюдений. Это маркетинговое исследование еще не закончилось, и для того чтобы постоянно не пересчитывать модель по такому большому числу наблюдений, аналитики решили пользоваться рекуррентным МНК. Текущая обратная матрица имеет вид

0,0000005	-0,0000024	-0,0007783	-0,0003367
-0,0000024	0,0000195	0,0029822	0,0004840
-0,0007783	0,0029822	1,5809786	0,5299996
-0,0003367	0,0004840	0,5299996	0,5860256

а текущие коэффициенты модели равны

11900,0022
0,9804
0,8068
0,6550
0,8421

Пересчитайте эти коэффициенты в связи с появлением следующих новых наблюдений

Интернет-магазин	y	x_1	x_2	x_3	x_4
11	14543	2543	185	1	0
12	13884	1967	70	0	0

7. МНОГОФАКТОРНЫЕ АДАПТИВНЫЕ МОДЕЛИ

7.1. Расчетные формулы

7.1.1. Многофакторная регрессионная модель с адаптивным механизмом в виде рекуррентных формул

$$\hat{y}_t = \mathbf{x}_t \hat{\mathbf{B}}(t-1, \mathbf{a});$$

$$\hat{\mathbf{B}}(t, \mathbf{a}) = \hat{\mathbf{B}}(t-1, \mathbf{a}) + \frac{\mathbf{C}_{t-1}^{-1} \mathbf{x}_t'}{\mathbf{x}_t \mathbf{C}_{t-1}^{-1} \mathbf{x}_t' + \mathbf{a}} [y_t - \hat{y}_t];$$

$$\mathbf{C}_t^{-1} = \frac{1}{a} \left[\mathbf{C}_{t-1}^{-1} - \frac{\mathbf{C}_{t-1}^{-1} \mathbf{x}_t' \mathbf{x}_t \mathbf{C}_{t-1}^{-1}}{\mathbf{x}_t \mathbf{C}_{t-1}^{-1} \mathbf{x}_t' + a} \right]$$

где $\hat{\mathbf{B}}(0, a)$, \mathbf{C}_0^{-1} - начальные значения, определяемые по методу наименьших квадратов.

7.1.2. Критерии настройки параметра адаптации

$$S_t^1(a) = \sum_{j=1}^{t-t} \sum_{k=1}^t |y_{j+k} - \hat{y}_{j+k}|,$$

$$S_t^2(a) = \sum_{j=1}^{t-t} \max_{1 \leq k \leq t} |y_{j+k} - \hat{y}_{j+k}|,$$

$$S_t^3(a) = \sum_{j=1}^{t-t} \max_{1 \leq k \leq t} \frac{|y_{j+k} - \hat{y}_{j+k}|}{|y_{j+k}|},$$

где $\hat{y}_{j+k} = \mathbf{x}_{j+k} \hat{\mathbf{B}}(j, a)$; $j = 1, 2, \mathbf{K}, t-t$; $k = 0, 1, \mathbf{K}, t$.

7.1.3. Дисперсионное отношение Фишера для адаптивных регрессионных моделей

$$F_p = \frac{N-m-1}{m} \cdot \frac{\sum (\hat{y}_i - \tilde{y}_i)^2}{\sum (y_i - \hat{y}_i)^2},$$

где \tilde{y}_i – экспоненциально взвешенное среднее значение;

\hat{y}_i – расчетные значения адаптивной модели.

7.2. Решение типовых задач

Таблица 7.2.1

Год	Месяц	Средняя цена 1 кг. говядины, у.е.		Среднемесячная заработная плата, у.е.
T	t	y_t	y_{t-1}	x_t
2001	сентябрь	1,2	1,1	6,1
	октябрь	1,5	1,2	7,2
	ноябрь	1,5	1,5	11,2
	декабрь	2,3	1,5	11,6
2002	январь	3,4	2,3	13,9
	февраль	4,3	3,4	17,7
	март	4,5	4,3	22,1
	апрель	5,8	4,5	25,7
	май	7,2	5,8	35,5
	июнь	8,9	7,2	45,4

Задание 7.2.1. По данным табл. 7.2.1 необходимо построить адаптивную регрессионную модель для прогнозирования цены говядины y_t на следующий период в зависимости от уровня средней заработной платы x_t и цены говядины в предшествующий момент времени y_{t-1} . Результаты прогнозирования по адаптивной модели следует сравнить с результатами прогнозирования по статической модели.

Решение с помощью табличного процессора Excel

1. Ввод исходных данных табл. 7.2.1.
2. Выбор начальных значений для построения адаптивной регрессионной модели \mathbf{C}_0^{-1} и \mathbf{B}_0 по первым восьми наблюдениям с помощью матричных функций ТРАНСП, МУМНОЖ, МОБР

2.1. Вычисление \mathbf{C}_0

$$\mathbf{C}_0 = \mathbf{X}'\mathbf{X} = \begin{bmatrix} 8,00 & 115,50 & 19,80 \\ 115,50 & 2004,45 & 352,38 \\ 19,80 & 352,38 & 62,74 \end{bmatrix},$$

и \mathbf{C}_0^{-1}

$$\mathbf{C}_0^{-1} = \begin{bmatrix} 0,9775 & -0,1658 & 0,6230 \\ -0,1658 & 0,0677 & -0,3277 \\ 0,6230 & -0,3277 & 1,6597 \end{bmatrix}.$$

2.2. Нахождение вектора

$$\mathbf{X}'\mathbf{y} = \begin{bmatrix} 24,50 \\ 433,48 \\ 76,71 \end{bmatrix}.$$

2.3. Расчет начальных значений вектора оценок коэффициентов

$$\mathbf{B}_0 = \mathbf{C}_0^{-1}\mathbf{X}'\mathbf{Y} = \begin{bmatrix} -0,1527 \\ 0,1301 \\ 0,5399 \end{bmatrix}.$$

Таким образом, для начальных значений регрессионная модель записывается в следующем виде:

$$y_t = -0,1527 + 0,1301x_t + 0,5399y_{t-1}.$$

3. Адаптивная корректировка коэффициентов регрессии

3.1. Определение с помощью построенного уравнения регрессии расчетного значения \hat{y}_8

$$\hat{y}_8 = 5,6216$$

3.2. Вычисление прогнозной оценки \hat{y}_9

$$\hat{y}_9 = 7,5025.$$

3.3. Расчет

$$\mathbf{C}_0^{-1} \mathbf{x}'_9 = \begin{bmatrix} -1,4079 \\ 0,3940 \\ -1,6792 \end{bmatrix}$$

3.4. Вычисление $\mathbf{x}_9 \mathbf{C}_0^{-1} \mathbf{x}'_9 + a$, выбрав в качестве сглаживающего параметра величину $a = 0,35$

$$\mathbf{x}_9 \mathbf{C}_0^{-1} \mathbf{x}'_9 + a = 2,9443.$$

3.5. Получение корректирующего вектора

$$\frac{\mathbf{C}_0^{-1} \mathbf{x}'_9}{\mathbf{x}_9 \mathbf{C}_0^{-1} \mathbf{x}'_9 + a} = \begin{bmatrix} -0,4404 \\ 0,1140 \\ -0,4698 \end{bmatrix}.$$

3.6. Расчет прогнозной ошибки для вновь поступившего наблюдения

$$y_9 - \mathbf{x}_9 \mathbf{B}_0 = 7,2 - 7,5988 = -0,3988$$

и умножение на эту ошибку корректирующего вектора

$$\frac{\mathbf{C}_0^{-1} \mathbf{x}'_9}{\mathbf{x}_9 \mathbf{C}_0^{-1} \mathbf{x}'_9 + a} [y_9 - \mathbf{x}_9 \mathbf{B}_0] = \begin{bmatrix} 0,1756 \\ -0,0455 \\ 0,1873 \end{bmatrix}.$$

3.7. Получение скорректированного по вновь поступившему наблюдению вектора коэффициентов регрессионной модели

$$\mathbf{B}_1 = \mathbf{B}_0 + \frac{\mathbf{C}_0^{-1} \mathbf{x}'_9}{\mathbf{x}_9 \mathbf{C}_0^{-1} \mathbf{x}'_9 + a} [y_9 - \mathbf{x}_9 \mathbf{B}_0] = \begin{bmatrix} 0,0229 \\ 0,0847 \\ 0,7273 \end{bmatrix}.$$

Таким образом, регрессионная модель с обновленными коэффициентами имеет вид

$$y_t = 0,0229 + 0,0847x_t + 0,7273y_{t-1}.$$

4. Сравнение результатов прогнозирования по адаптивной и статической моделям.

Прогнозные расчеты по статической модели \hat{y}_{10} и по адаптивной \hat{y}_{10}^c дают следующие результаты:

$$\hat{y}_{10} = 9,8584 ; \quad \hat{y}_{10}^c = 9,1384 .$$

Соответственно ошибки прогнозирования равны

$$y_{10} - \hat{y}_{10} = -0,9584 ; \quad y_{10} - \hat{y}_{10}^c = -0,2384 .$$

Следовательно, в рассматриваемом примере краткосрочный прогноз с помощью адаптивной модели более точен.

Задание 7.2.2. Проверить адекватность построенной в предыдущем задании многофакторной адаптивной модели, используя для этого критерий (7.1.3).

Решение с помощью табличного процессора Excel

1. Вычисление дисперсионного отношения Фишера по первым восьми наблюдениям.

$$F_p = \frac{N - m - 1}{m} \frac{\sum (\hat{y}_i - \bar{y})^2}{\sum (y_i - \hat{y}_i)^2} = \frac{5}{2} \cdot \frac{19,0581}{1,0806} \approx 44,0903.$$

Сравнение расчетного значения F -критерия с табличным $F_{2;5;0,95} = 5,79$ позволяет сделать вывод об адекватности построенной модели.

2. Расчет дисперсионного отношения Фишера после добавления девятого наблюдения. Оформление расчетов в виде табл. 7.2.2.

Для упрощения расчетов модифицированного критерия имеет смысл, учитывая логику построения регрессионных уравнений и последовательность проверки их адекватности, к имеющимся суммам квадратов отклонений числителя и знаменателя дисперсионного отношения Фишера первой модели добавить квадраты отклонений, полученные для скорректированной модели. К сумме, стоящей в знаменателе, добавляется квадрат отклонения расчетного значения от фактического

$$(y_9 - \hat{y}_9)^2 = (7,2 - 7,2474)^2 = (-0,0474)^2.$$

Расчетное значение здесь определено по скорректированной модели.

К сумме, стоящей в числителе, добавляется квадрат отклонения среднего значения от расчетного

$$(\hat{y}_9 - \tilde{y})^2 = (7,2474 - 4,5106)^2 = (2,7368)^2,$$

при вычислении которого использовано экспоненциальное среднее

$$\tilde{y} = \bar{y} + a(y_9 - \bar{y}) = 3,0625 + 0,35(7,2 - 3,0625) = 4,5106.$$

После внесения этих изменений, окончательно получаем

$$F_p = \frac{N - m - 1}{m} \frac{\sum (\hat{y}_i - \tilde{y})^2}{\sum (y_i - \hat{y}_i)^2} = \frac{6}{2} \cdot \frac{26,5481}{1,0828} \approx 73,5488.$$

Сравнение полученного значения с табличным $F_{2;6;0,95} = 5,14$ позволяет сделать вывод о том, что после корректировки модель остается адекватной.

Таблица 7.2.2

№ п.п.	y	\hat{y}	$(y - \hat{y})$	$(y - \hat{y})^2$	$(y - \bar{y})$	$(y - \bar{y})^2$
1	1,2	1,1654	-0,0346	0,0012	-1,8274	3,3395
2	1,5	1,5678	0,0678	0,0046	-1,6303	2,6579
3	1,5	0,8853	-0,6147	0,3779	-0,9478	0,8982
4	2,3	2,433	0,133	0,0177	-0,8957	0,8023
5	3,4	3,9019	0,5019	0,2519	-0,1644	0,027
6	4,3	4,6135	0,3135	0,0983	0,924	0,8538
7	4,5	3,9549	0,5451	0,2971	1,9826	3,9306
8	5,8	5,9783	0,1783	0,0318	2,559	6,5488
Суммы				1,0806		19,0581
9	7,2	7,1531	-0,0469	0,0022	2,7368	7,49
Суммы				1,0828		26,5481

7.3. Задания для самостоятельной работы

Задание 7.3.1. По данным табл. 7.3.1 построить многофакторную адаптивную модель для прогнозирования дохода на акцию компании «Пинпод» (y , руб.) в зависимости от объема продаж (x_1 , руб.) и чистой прибыли (x_2 , руб.). Начальные значения для ее построения получить с помощью МНК по первым четырем наблюдениям. Последние два наблюдения использовать для настройки параметра a . Рассчитать прогнозные значения дохода на акцию на 2000-2001гг. Сравнить прогнозные расчеты, по

лученные с помощью адаптивной модели и с помощью обычной регрессии.

Таблица 7.3.1

Год	y	x_1	x_2
1994	8005	4437	0,75
1995	15209	6592	0,79
1996	27642	9566	0,82
1997	56943	12979	0,87
1998	69265	21565	1,27
1999	73134	2845	1,62

Задание 7.3.2. По данным табл. 7.3.2 построить многофакторную адаптивную модель для прогнозирования валовой выручки от продаж (y , руб.) в зависимости от расходов на рекламу (x_1 , руб.) и персонал (x_2 , руб.). Определить начальные значения с помощью МНК по первым пяти наблюдениям. Осуществить настройку параметра адаптации, используя для этого последние два наблюдения. Оценить адекватность построенной адаптивной модели по модифицированному критерию Фишера. В случае, если модель адекватна, осуществить прогнозные расчеты на следующие два периода.

Таблица 7.2.3

год	y	x_1	x_2
1994	29191925	1110577	7634465
1995	34209290	1144408	10482865
1996	43775216	1732440	12429585
1997	52150998	1832064	15544310
1998	62816819	1916779	18332890
1999	75439535	1974583	21197454
2000	90387247	2015704	24144711

8. МОДЕЛИ БИНАРНОГО ВЫБОРА

8.1. Решение типовой задачи

Задание 8.1.1. Фирма «Бэст-Перспектива» заключила договор с ОАО «Центртелеком», предметом которого является оказание услуг связи, в том числе и обеспечение доступа в Интернет. С целью увеличения интеллектуального потенциала фирмы руководство решило расширить круг своих со

трудников, обладающих правом бесплатного доступа. Выбор таких сотрудников было решено осуществлять с учетом ряда факторов, тем или иным образом характеризующих претендентов с точки зрения наиболее эффективного использования Интернет-ресурсов. В связи с этим возник вопрос: «Кому из претендентов на бесплатный доступ предоставить такую возможность в первую очередь?» Для того, чтобы получить обоснованный ответ на этот вопрос, руководство «Бэст-Перспективы» поручило экономико-аналитическому отделу разработать модель, позволяющую по каждому претенденту рассчитать прогнозную оценку целесообразности предоставления ему права бесплатного доступа к ресурсам Интернета. В основу построения такой модели была положена идея применения бинарной переменной

$$y = \begin{cases} 1, & \text{если сотрудник, обладающий правом бесплатного доступа к} \\ & \text{интернет – ресурсам, по оценке экспертной группы эффективно} \\ & \text{пользуется этим правом} \\ 0, & \text{в противном случае} \end{cases},$$

зависящей от некоторых анкетных данных, а также результатов тестирования, характеризующих навыки использования Интернет.

Построение модели руководство фирмы предложило провести по пяти факторам: возраст, стаж профессиональной деятельности, заработная плата, число случаев поступления полезной для фирмы информации от сотрудника, результат тестирования (в баллах) на предмет оценки навыков работы в Интернет. Значения этих показателей, а также значения бинарной переменной представлены в табл. 8.8.1.

Имеются следующие претенденты на право бесплатного доступа:

- 1) возраст – 27 лет, стаж – 3 года, заработная плата – 3200 руб., количество случаев нахождения полезной для фирмы информации – 9 раз, тест – 15 баллов;
- 2) возраст – 44 года, стаж – 12 лет, заработная плата – 5600 руб., количество случаев нахождения полезной для фирмы информации – 2 раза, тест – 5 баллов;

- 3) возраст – 35 лет, стаж – 10 лет, заработная плата – 4100 руб., количество случаев нахождения полезной для фирмы информации – 4 раза, тест – 7 баллов;
- 4) возраст – 39 лет, стаж – 13 лет, заработная плата – 7500 руб., количество случаев нахождения полезной для фирмы информации – 11 раз, тест – 15 баллов.

Используя построенную прогнозную модель, определить среди имеющихся претендентов тех, кому в первую очередь следует предоставить право бесплатного доступа к ресурсам Интернета.

Решение с помощью пакета STATISTICA.

1. Подготовка данных для проведения расчетов с учетом группировки по интервалам, приведенным в табл. 8.1.2. В результате группировки исходный массив данных будет иметь вид табл. 3 (где q - частота появления соответствующего наблюдения в выборочной совокупности).

2. Ввод данных табл. 8.1.3. Для этого необходимо загрузить пакет STATISTICA и создать чистую таблицу, установив ее размеры (27 строк и 7 столбцов) с помощью кнопок «Переменные» (столбцы) и «Случаи» (строки).

3. Открыть меню «Статистика», в нем выбрать «Дополнительные линейные и нелинейные модели» — «Нелинейная оценка» — «Логистическая регрессия».

4. В открывшемся окне установить с помощью стрелочки опцию «Коды и числа» и выбрать в качестве независимых переменных «Переменные 2-6», в качестве зависимой переменной – «Переменную 1» и в качестве переменной, содержащей частоту появления каждого из наблюдаемых событий – «Переменную 7» и щелкнуть на кнопке «ОК».

5. На появившейся панели «Оценивание модели» в окне «Метод оценивания», используя стрелку, выбрать квазинытоновский метод и, не изменяя параметров итерационного процесса, щелкнуть на кнопке «ОК».

6. Для просмотра параметров модели в открывшемся окне «Результаты» необходимо выбрать пункт «Обзор: оценки параметров». Полученные результаты представлены в табл. 8.1.4.

Таблица 8.8.1

№ п.п.	Бинарная переменная	Возраст претендента, лет	Стаж работы, лет	Заработная плата, тыс. руб.	Количество случаев нахождения полезной информации	Результаты тестирования, балл	№ п.п.	Бинарная переменная	Возраст претендента, лет	Стаж работы, лет	Заработная плата, тыс. руб.	Количество случаев нахождения полезной информации	Результаты тестирования, баллы
1	2	3	4	5	6	7	8	9	10	11	12	13	14
1.	0	22	1	2,5	2	6	26.	0	52	17	9,7	2	1
2.	0	24	1	3	3	8	27.	0	54	12	8,9	3	4
3.	1	25	1	2,1	1	7	28.	1	18	1	3,4	4	11
4.	1	27	4	4,6	8	11	29.	0	19	1	3,7	5	12
5.	1	28	3	5,9	9	13	30.	1	20	1	2,9	6	13
6.	0	21	1	3,3	5	14	31.	0	19	1	3,8	5	15
7.	1	22	1	3,2	9	15	32.	0	20	1	4,3	1	8
8.	0	29	5	6	1	13	33.	1	28	2	6,9	9	14
9.	1	26	4	2,7	8	11	34.	0	27	4	6,5	5	12
10.	0	27	6	3,8	0	3	35.	0	43	8	9,2	2	9
11.	1	28	4	4,2	9	11	36.	1	44	13	10,3	11	13
12.	0	29	3	6,7	5	13	37.	0	56	12	10,8	2	4
13.	1	29	9	6,5	10	14	38.	1	54	14	9,4	12	13
14.	1	30	4	3,4	9	13	39.	0	44	11	9,8	2	6
15.	1	41	7	7,2	7	15	40.	1	28	4	6,0	8	12
16.	0	32	8	7,9	4	7	41.	0	19	1	3,4	2	7
17.	1	46	9	6,4	9	12	42.	0	20	1	4,3	5	11
18.	1	34	3	6,3	8	13	43.	0	19	1	4,9	5	12
19.	0	33	3	6,1	1	9	44.	1	21	1	3,2	6	13
20.	1	47	9	9,7	5	13	45.	0	45	12	12,4	2	4
21.	0	37	2	6,4	3	7	46.	0	19	1	3,1	4	13
22.	0	38	5	10,5	2	8	47.	0	48	13	12,7	2	5
23.	0	49	13	12,3	3	3	48.	1	20	1	3,5	4	13
24.	1	46	7	10,1	5	14	49.	0	58	18	9,9	2	2
25.	1	48	8	7,9	8	15	50.	0	49	12	11,8	3	3

Продолжение таблицы 8.8.1

1	2	3	4	5	6	7	8	9	10	11	12	13	14
51.	0	57	18	10,2	2	2	76.	1	26	8	5,4	11	15
52.	0	28	4	6,7	1	11	77.	0	39	5	7,6	2	7
53.	0	52	14	9,9	2	2	78.	0	32	9	5,5	5	8
54.	0	45	15	13,1	3	4	79.	1	33	4	5,4	7	11
55.	0	53	16	9,1	1	4	80.	1	23	1	2,8	8	13
56.	0	54	20	9,7	1	5	81.	0	36	3	6,5	2	9
57.	0	30	4	6,8	3	12	82.	0	35	9	6,9	5	8
58.	0	46	12	12,2	2	1	83.	1	23	1	3,2	9	11
59.	0	28	4	3,9	0	2	84.	1	34	3	6,1	8	12
60.	0	44	14	12,5	1	2	85.	0	35	8	5,9	5	9
61.	0	22	1	2,8	3	7	86.	0	38	2	7,8	3	10
62.	0	29	4	3,4	0	2	87.	0	32	9	7,3	5	8
63.	0	23	1	4,9	4	15	88.	0	33	2	6,4	3	10
64.	1	26	8	6,7	12	15	89.	1	34	3	6,5	8	11
65.	1	27	9	6,6	11	13	90.	0	35	10	7,9	5	10
66.	0	39	4	8,9	2	8	91.	0	33	2	7,8	1	9
67.	1	27	7	5,5	11	11	92.	1	24	1	3,0	8	11
68.	0	26	6	5,7	2	11	93.	1	34	4	6,0	9	13
69.	1	28	9	6,1	12	12	94.	0	33	3	5,4	2	10
70.	1	28	8	7,3	10	13	95.	1	22	1	2,9	7	13
71.	0	40	4	9,2	2	7	96.	1	32	5	6,4	8	15
72.	1	28	10	7,9	10	11	97.	0	33	6	5,7	2	7
73.	1	22	1	2,5	8	12	98.	1	23	1	3,6	9	15
74.	1	26	9	6,8	11	15	99.	1	31	4	6,0	8	13
75.	0	40	5	6,0	2	9	100.	0	33	4	6,1	3	10

Таблица 8.1.2

ВОЗРАСТ		СТАЖ ПРОФЕССИО- НАЛЬНОЙ ДЕЯ- ТЕЛЬНОСТИ		ЗАРАБОТНАЯ ПЛАТА		КОЛИЧЕСТВО СЛУЧАЕВ ПОСТУ- ПЛЕНИЯ ПОЛЕЗ- НОЙ ИНФОРМА- ЦИИ ОТ СОТРУД- НИКА		РЕЗУЛЬТАТЫ ТЕСТИРОВА- НИЯ (МАКСИМАЛЬ- НЫЙ БАЛЛ - 15)	
Границы интерва- лов	Ме- диана	Границы интерва- лов	Ме- диана	Грани- цы ин- терва- лов	Медиа- на	Границы интервалов	Ме- диана	Граница интер- валов	Ме- диана
До 21 го- да	19	До 1года	1	2-5тыс	3	Не прино- сил	0	Ни од- ного	0
21-25	23	2-6	4	5-8тыс	6	1-3	2	1-5	3
26-30	28	7-11	9	8-11	9	4-6	5	6-10	8
31-35	33	12-16	14	От 11 и более	12	7-9	8	11-15	13
36-40	38	17-21	19			10-12	11		
41-50	45	Больше 21	24			Больше12	14		
Старше 50	55								

Таблица 8.1.3

№ п.п.	y	x_1	x_2	x_3	x_4	x_5	q
1.	0	23	1	3	2	8	3
2.	1	23	1	3	2	8	1
3.	0	23	1	3	5	13	2
4.	1	23	1	3	8	13	7
5.	0	28	4	6	2	13	4
6.	0	28	4	3	0	3	3
7.	1	28	9	6	11	13	9
8.	1	28	4	3	8	13	4
9.	0	33	9	6	5	8	6
10.	1	33	4	6	8	13	7
11.	0	33	4	6	2	8	6
12.	0	38	4	6	2	8	5
13.	0	38	4	9	2	8	3
14.	0	45	14	12	2	3	7
15.	1	45	9	9	5	13	2
16.	1	45	9	6	8	13	3
17.	0	55	19	9	2	3	4
18.	0	55	14	9	2	3	3
19.	1	19	1	3	5	13	4
20.	0	19	1	3	5	13	5
21.	0	19	1	3	2	8	2
22.	1	28	4	6	8	13	3
23.	0	28	4	6	5	13	2
24.	0	45	9	9	2	8	2
25.	1	45	14	9	11	13	1
26.	0	55	14	9	2	3	1
27.	1	55	14	9	11	13	1

Таблица 8.1.4

N=27	Model: Logistic regression (logit) N of 0's:58 1's:42 Dep. var: Var1 Loss: Max likelihood Final loss: 15,847791601 Chi?(5)=104,36 p=0,0000					
	Const.B0	Var2	Var3	Var4	Var5	Var6
Estimate	-13,2000	0,27	-0,521844	-0,315524	1	0,150857
Odds ratio (unit ch)	0,0000	1,32	0,593425	0,729407	4	1,162831
Odds ratio (range)		19876,47	0,000083	0,058442	3704998	4,520271

Аналитическое представление построенной логит-модели можно записать следующим образом:

$$P(y_i = 1 / \tilde{x}_i) = (1 + e^{13,2 - 0,27\tilde{x}_{1i} + 0,521844\tilde{x}_{2i} + 0,315524\tilde{x}_{3i} - \tilde{x}_{4i} - 0,150857\tilde{x}_{5i}})^{-1}. (*)$$

7. Для оценки прогнозных возможностей построенной модели выполнить пункт «Наблюдаемые, предсказанные значения, остатки». Полученные результаты приведены в табл. 8.1.5.

Таблица 8.1.5

№ п.п.	Наблю-денные значения	Предска-занные значения	Остатки	№ п.п.	Наблю-денные значения	Предска-занные значения	Остатки
1.	0	0,01227	-0,01227	14.	0	0,000164	-0,00016
2.	1	0,01227	0,98773	15.	1	0,615739	0,384261
3.	0	0,620354	-0,62035	16.	1	0,996101	0,003899
4.	1	0,990206	0,009794	17.	0	0,000485	-0,00049
5.	0	0,008397	-0,0084	18.	0	0,006548	-0,00655
6.	0	0,000308	-0,00031	19.	1	0,352368	0,647632
7.	1	0,99327	0,00673	20.	0	0,352368	-0,35237
8.	1	0,98817	0,01183	21.	0	0,004119	-0,00412
9.	0	0,0669	-0,0669	22.	1	0,970073	0,029927
10.	1	0,992257	0,007743	23.	0	0,343795	-0,3438
11.	0	0,015502	-0,0155	24.	0	0,012035	-0,01204
12.	0	0,058606	-0,05861	25.	1	0,99779	0,00221
13.	0	0,023589	-0,02359	26.	0	0,006548	-0,00655
				27.	1	0,999858	0,000142

При сравнении первого и второго столбцов этой таблицы можно сделать следующий вывод: с достаточным уровнем надежности не удалось предсказать поведение для случаев, описываемых 2, 3 и 19 наблюдениями. Но в 93 случаев из 100 удалось точно предсказать стратегию поведения фирмы в отношении сотрудников (предоставлять или нет право бесплатно-

го доступа к ресурсам Интернет). Надежность модели также подтверждается расчетным значением хи-квадрат (104,36), которое значительно боль

ше теоретического значения (32,67), и почти нулевой вероятностью не отвергнуть нулевую гипотезу.

8. Получение значений коэффициентов абсолютного роста по каждой переменной. Для этого необходимо выражение (*), описывающее полученную логит-зависимость, продифференцировать по \tilde{x}_k и вычислить значения производной в каждом наблюдении $\tilde{\mathbf{x}}_i$

$$\frac{\partial P(y_i = 1 / \tilde{\mathbf{x}}_i)}{\partial \tilde{x}_k} = (1 + e^{-\tilde{\mathbf{x}}_i b})^{-1} (1 + e^{\tilde{\mathbf{x}}_i b})^{-1} b_k.$$

Результаты расчетов представлены в табл.8.1.6.

Таблица 8.1.6

№ п.п.	Коэффициенты абсолютного роста				
	x_1	x_2	x_3	x_4	x_5
1.	0,00140	-0,00271	-0,00164	0,00518	0,00078
2.	0,00140	-0,00271	-0,00164	0,00518	0,00078
3.	0,04034	-0,07797	-0,04714	0,14940	0,02254
4.	0,04020	-0,07770	-0,04698	0,14889	0,02246
5.	0,00093	-0,00181	-0,00109	0,00346	0,00052
6.	0,00007	-0,00014	-0,00008	0,00027	0,00004
7.	0,05922	-0,11446	-0,06921	0,21934	0,03309
8.	0,04577	-0,08847	-0,05349	0,16953	0,02557
9.	0,00248	-0,00479	-0,00289	0,00917	0,00138
10.	0,03549	-0,06860	-0,04148	0,13145	0,01983
11.	0,00169	-0,00326	-0,00197	0,00624	0,00094
12.	0,00627	-0,01213	-0,00733	0,02324	0,00351
13.	0,00251	-0,00485	-0,00293	0,00929	0,00140
14.	0,00002	-0,00003	-0,00002	0,00006	0,00001
15.	0,03711	-0,07172	-0,04336	0,13744	0,02073
16.	0,02197	-0,04247	-0,02568	0,08138	0,01228
17.	0,00005	-0,00009	-0,00006	0,00017	0,00003
18.	0,00064	-0,00123	-0,00075	0,00236	0,00036
19.	0,01772	-0,03425	-0,02071	0,06563	0,00990
20.	0,01772	-0,03425	-0,02071	0,06563	0,00990
21.	0,00048	-0,00093	-0,00056	0,00177	0,00027
22.	0,06558	-0,12675	-0,07664	0,24289	0,03664
23.	0,01651	-0,03190	-0,01929	0,06113	0,00922
24.	0,00123	-0,00238	-0,00144	0,00457	0,00069
25.	0,03368	-0,06510	-0,03936	0,12475	0,01882
26.	0,00064	-0,00123	-0,00075	0,00236	0,00036
27.	0,00303	-0,00586	-0,00355	0,01124	0,00170

Анализ коэффициентов абсолютного роста показывает, что второй и третий факторы (стаж профессиональной деятельности и заработная плата)

имеют отрицательные коэффициенты абсолютного роста, а остальные – положительные. Это свидетельствует о том, что при увеличении значений таких факторов, как стаж и зарплата, вероятность предоставления бесплатного доступа в Интернет снижается, при увеличении таких факторов как возраст, результаты тестирования, количество случаев нахождения полезной для фирмы информации – вероятность возрастает.

9. Использование построенной модели для выбора среди имеющихся претендентов тех, кому в первую очередь следует предоставить право бесплатного доступа к ресурсам Интернета:

$$1) (1 + e^{13,2-0,27 \cdot 27+0,521844 \cdot 3+0,315524 \cdot 3,2-9-0,150857 \cdot 15})^{-1} = 0,941;$$

$$2) (1 + e^{13,2-0,27 \cdot 44+0,521844 \cdot 12+0,315524 \cdot 5,6-2-0,150857 \cdot 5})^{-1} = 0,001;$$

$$3) (1 + e^{13,2-0,27 \cdot 35+0,521844 \cdot 10+0,315524 \cdot 4,1-4-0,150857 \cdot 7})^{-1} = 0,0054;$$

$$4) (1 + e^{13,2-0,27 \cdot 39+0,521844 \cdot 13+0,315524 \cdot 7,5-11-0,150857 \cdot 15})^{-1} = 0,808.$$

Следовательно, первому и четвертому претендентам целесообразно предоставить право бесплатного доступа, а второму и третьему – нет.

8.2. Задание для самостоятельной работы

Задание 8.2. Ежегодно НИИ «Интеллектуальный бизнес-успех», занимающийся, в частности, проблемами макро- и микромоделирования, обеспечивает своим сотрудникам, специализирующимся в этой области, стажировки в англоязычные страны. При отборе претендентов на поездку учитываются такие факторы, как: 1) знание английского языка (претенденты сдают тест, максимальное число баллов по которому можно получить - 50); 2) стаж профессиональной деятельности; 3) ученая степень (для ее обозначения используется следующая кодировка: 00 – без степени, 01 – кандидат наук, 10 – доктор наук); 4) количество публикаций по указанной проблематике. В настоящее время руководству НИИ снова предстоит решить, кого из сотрудников отправить на стажировку. Для того чтобы выбор претендента осуществлялся исключительно на научной основе, а также для того, чтобы иметь возможность предвидеть, что даст организации ста-

жировка того или иного претендента, аналитику НИИ было поручено построить модель, позволяющую реализовать эти цели. Сложность поставленной задачи потребовало использования не совсем традиционного под

хода. В качестве основы построения модели было решено использовать бинарную переменную

$$y = \begin{cases} 1, & \text{стажер вернулся с дипломом} \\ 0, & \text{стажер вернулся с сертификатом} \end{cases},$$

поскольку сертификат свидетельствует о том, что сотрудник всего лишь прослушал курс по макро- и микромоделированию, а диплом – сотрудник не только прослушал курс, но и сдал квалификационный экзамен. Очевидно, что диплом, полученный стажером, положительным образом сказывается не только на его собственной карьере, но на имидже НИИ, а также увеличивает «интеллект» организации. Становится понятным желание руководства отбор провести таким образом, чтобы максимально возможное число претендентов вернулось с дипломами. Данные для построения модели представлены в табл. 8.2.1. В текущий момент руководству необходимо выяснить, кого из претендентов следует отправить на стажировку в первую очередь. Потенциальные стажеры обладают следующими характеристиками:

1) результаты тестирования – 27 баллов, стаж профессиональной деятельности – 3 года, кандидат наук, количество публикаций – 21;

2) результаты тестирования – 48 баллов, стаж профессиональной деятельности – 12 лет, без степени, количество публикаций – 10;

3) результаты тестирования – 32 балла, стаж профессиональной деятельности – 25 лет, доктор наук, количество публикаций – 98;

Таблица 8.2.1

№ п.п.	Бинарная переменная	Результаты тестирования, балл	Стаж работы, лет	Ученая степень	Количество публикаций	№ п.п.	Бинарная переменная	Результаты тестирования, балл	Стаж работы, лет	Ученая степень	Количество публикаций
1	2	3	4	5	6	7	8	9	10	11	12
1.	1	39	7	кандидат	15	26.	0	12	4	кандидат	9
2.	0	7	2	кандидат	7	27.	1	21	14	кандидат	18
3.	0	27	4	без степени	2	28.	0	31	3	без степени	3
4.	0	29	5	без степени	1	29.	0	4	20	кандидат	13
5.	0	34	4	без степени	3	30.	0	34	3	без степени	3
6.	1	18	11	кандидат	18	31.	0	12	8	кандидат	6
7.	1	21	13	кандидат	17	32.	0	8	4	кандидат	7
8.	1	40	9	кандидат	13	33.	0	3	17	кандидат	13
9.	1	28	11	без степени	6	34.	0	10	7	кандидат	5
10.	1	24	11	кандидат	16	35.	1	26	11	кандидат	16
11.	0	12	5	без степени	9	36.	1	30	21	кандидат	13
12.	1	25	13	кандидат	20	37.	0	36	2	без степени	1
13.	1	20	18	доктор	43	38.	0	19	7	без степени	7
14.	1	30	15	без степени	5	39.	0	36	5	без степени	3
15.	0	18	7	без степени	3	40.	1	46	19	кандидат	37
16.	1	18	11	кандидат	16	41.	0	13	1	кандидат	7
17.	1	39	18	кандидат	30	42.	1	36	14	кандидат	26
18.	1	24	15	кандидат	19	43.	0	21	10	без степени	9
19.	0	4	16	без степени	5	44.	1	20	11	кандидат	19
20.	1	46	15	доктор	45	45.	0	36	4	без степени	3
21.	0	9	5	кандидат	8	46.	0	14	8	без степени	3
22.	1	16	11	кандидат	17	47.	0	31	4	без степени	2
23.	1	24	7	кандидат	14	48.	0	7	5	без степени	6
24.	0	20	9	без степени	7	49.	0	33	3	без степени	1
25.	0	3	18	без степени	4	50.	0	14	1	кандидат	8

Продолжение таблицы 8.2.1

№ п.п.	Бинарная переменная	Результаты тестирования, балл	Стаж работы, лет	Ученая степень	Количество публикаций	№ п.п.	Бинарная переменная	Результаты тестирования, балл	Стаж работы, лет	Ученая степень	Количество публикаций
1	2	3	4	5	6	7	8	9	10	11	12
51.	1	27	11	без степени	4	76.	0	3	16	кандидат	13
52.	0	16	7	без степени	7	77.	1	38	16	кандидат	30
53.	0	21	7	без степени	2	78.	0	2	20	кандидат	15
54.	1	36	14	без степени	6	79.	0	3	16	без степени	4
55.	0	21	6	без степени	9	80.	0	4	19	кандидат	14
56.	1	41	17	кандидат	31	81.	0	9	7	кандидат	4
57.	0	25	7	без степени	8	82.	0	9	8	кандидат	4
58.	1	33	12	без степени	5	83.	0	11	6	кандидат	6
59.	0	19	9	без степени	8	84.	0	5	18	кандидат	14
60.	0	13	1	без степени	8	85.	0	4	17	без степени	5
61.	0	14	4	без степени	7	86.	0	4	16	без степени	6
62.	1	28	13	кандидат	21	87.	0	2	17	кандидат	13
63.	1	29	13	без степени	6	88.	0	2	16	без степени	4
64.	1	27	21	кандидат	13	89.	1	16	9	кандидат	13
65.	1	29	23	кандидат	14	90.	0	3	19	кандидат	15
66.	1	31	21	кандидат	15	91.	0	12	7	без степени	3
67.	0	27	5	без степени	3	92.	0	13	7	без степени	3
68.	0	29	4	без степени	2	93.	0	3	16	кандидат	14
69.	0	27	3	без степени	2	94.	1	19	8	кандидат	14
70.	0	33	5	без степени	3	95.	1	18	8	кандидат	15
71.	1	31	15	кандидат	28	96.	0	5	16	кандидат	13
72.	1	36	11	кандидат	21	97.	1	16	6	кандидат	13
73.	0	36	2	без степени	1	98.	0	4	20	кандидат	15
74.	0	31	4	без степени	2	99.	1	17	7	кандидат	13
75.	0	34	3	без степени	3	100.	0	3	18	кандидат	15

ПРИЛОЖЕНИЕ

Таблица 1

Двусторонние квантили распределения Стьюдента $t_a(n)$

(n – число степеней свободы, a – доверительный интервал)

	0,20	0,40	0,50	0,60	0,80	0,90	0,95	0,98	0,99
1	0,325	0,727	1,000	1,376	3,078	6,314	12,706	31,821	63,657
2	0,289	0,617	0,816	1,061	1,886	2,920	4,303	6,965	
3	0,277	0,584	0,765	0,978	1,638	2,353	3,182	4,541	5,841
4	0,271	0,569	0,741	0,941	1,533	2,132	2,776	3,747	4,604
5	0,267	0,559	0,727	0,920	1,476	2,015	2,571	3,365	4,032
6	0,265	0,553	0,718	0,906	1,440	1,943	2,447	3,143	3,707
7	0,263	0,549	0,711	0,896	1,415	1,895	2,365	2,998	3,499
8	0,262	0,546	0,706	0,889	1,397	1,860	2,306	2,896	3,355
9	0,261	0,543	0,703	0,883	1,383	1,833	2,262	2,821	3,250
10	0,260	0,542	0,700	0,879	1,372	1,812	2,228	2,764	3,169
11	0,260	0,540	0,697	0,876	1,363	1,796	2,201	2,718	3,106
12	0,259	0,539	0,695	0,873	1,356	1,782	2,179	2,681	3,055
13	0,259	0,538	0,694	0,870	1,350	1,771	2,160	2,650	3,012
14	0,258	0,537	0,692	0,868	1,345	1,761	2,145	2,624	2,977
15	0,258	0,536	0,691	0,866	1,341	1,753	2,131	2,602	2,947
16	0,258	0,535	0,690	0,865	1,337	1,746	2,120	2,583	2,291
17	0,257	0,534	0,689	0,863	1,333	1,740	2,110	2,567	2,898
18	0,257	0,534	0,688	0,862	1,330	1,734	2,101	2,552	2,878
19	0,257	0,533	0,688	0,861	1,328	1,729	2,093	2,539	2,861
20	0,257	0,533	0,687	0,860	1,325	1,725	2,086	2,528	2,845
21	0,257	0,532	0,686	0,859	1,323	1,721	2,080	2,518	2,831
22	0,256	0,532	0,686	0,858	1,321	1,717	2,074	2,508	2,819
23	0,256	0,532	0,685	0,858	1,319	1,714	2,069	2,500	2,807
24	0,256	0,531	0,685	0,857	1,318	1,711	2,064	2,492	2,797
25	0,256	0,531	0,684	0,856	1,316	1,708	2,060	2,485	2,787
26	0,256	0,531	0,684	0,856	1,315	1,706	2,056	2,479	2,779
27	0,256	0,531	0,684	0,855	1,314	1,703	2,052	2,473	2,771
28	0,256	0,530	0,683	0,855	1,313	1,701	2,048	3,467	2,763
29	0,256	0,530	0,683	0,854	1,311	1,699	2,045	2,462	2,756
30	0,256	0,530	0,683	0,854	1,310	1,697	2,042	2,457	2,750
40	0,255	0,529	0,681	0,851	1,303	1,684	2,021	2,423	2,704
60	0,254	0,527	0,679	0,848	1,296	1,671	2,000	2,390	2,660
100	0,254	0,526	0,677	0,845	1,290	1,660	1,984	2,364	2,626
200	0,254	0,525	0,676	0,843	1,286	1,652	1,972	2,345	2,601
∞	0,253	0,524	0,675	0,842	1,282	1,645	1,96	2,326	2,576

Пример. Пусть t - случайная величина, распределенная по закону Стьюдента с пятью степенями свободы.

$t_{0,95}(5) = 2,571$, т.е. $P(-2,571 < t < 2,571) = 0,95$ (см. пятая строка, третий столбец).

Таблица 2

Квантили распределения $\chi^2(n)$
 (n – число степеней свободы, α - доверительный интервал)

	0,005	0,010	0,025	0,050	0,100	0,900	0,950	0,975	0,990	0,995
1	0,000039	0,00016	0,00098	0,0039	0,0158	2,71	3,84	5,02	6,63	7,88
2	0,0100	0,0201	0,0506	0,1026	0,2107	4,61	5,99	7,38	9,21	10,60
3	0,0717	0,115	0,216	0,352	0,584	6,25	7,81	9,35	11,34	12,84
4	0,207	0,297	0,484	0,711	1,064	7,78	9,49	11,14	13,28	14,86
5	0,412	0,554	0,831	1,15	1,61	9,24	11,07	12,83	15,09	16,75
6	0,676	0,872	1,24	1,64	2,20	10,64	12,59	14,45	16,81	18,55
7	0,989	1,24	1,69	2,17	2,83	12,02	14,07	16,01	18,48	20,28
8	1,34	1,65	2,18	2,73	3,49	13,36	15,51	17,53	20,09	21,96
9	1,73	2,09	2,70	3,33	4,17	14,68	16,92	19,02	21,67	23,59
10	2,16	2,56	3,25	3,94	4,87	15,99	18,31	20,48	23,21	25,19
11	2,60	3,05	3,82	4,57	5,58	17,28	19,68	21,92	24,73	26,76
12	3,07	3,57	4,40	5,23	6,30	18,55	21,03	23,34	26,22	28,30
13	3,57	4,11	5,01	5,89	7,04	19,81	22,36	24,74	27,69	29,82
14	4,07	4,66	5,63	6,57	7,79	21,06	23,68	26,12	29,14	31,32
15	4,60	5,23	6,26	7,26	8,55	22,31	25,00	27,49	30,58	32,80
16	5,14	5,81	6,91	7,96	9,31	23,54	26,30	28,85	32,00	34,27
18	6,26	7,01	8,23	9,39	10,86	25,99	28,87	31,53	34,81	37,16
20	7,43	8,26	9,59	10,85	12,44	28,41	31,41	34,17	37,57	40,00
24	9,89	10,86	12,40	13,85	15,66	33,20	36,42	39,36	42,98	45,56
30	13,79	14,95	16,79	18,49	20,60	40,26	43,77	46,98	50,89	53,67
40	20,71	22,16	24,43	26,51	29,05	51,81	55,76	59,34	63,69	66,77
60	35,53	37,48	40,48	43,19	46,46	74,40	79,08	83,30	88,38	91,95
80	51,17	53,54	57,15	60,39	64,28	96,58	101,88	106,63	112,33	116,32
100	67,33	70,06	74,22	77,93	82,36	118,5	124,34	129,56	135,81	140,17
120	83,85	86,92	91,58	95,7	100,62	140,23	146,57	152,21	158,95	163,64

Пример. Пусть χ^2 - случайная величина, распределенная по закону χ^2 с пятью степенями свободы. $\chi^2_{0,95}(5) = 11,07$, т.е. $P(\chi^2 < 11,07) = 0,95$ (см. пятая строка, седьмой столбец).

Таблица 4

95%-ные квантили распределения Фишера

 $(k_1$ - число степеней свободы числителя, k_2 - число степеней свободы знаменателя)

	1	2	3	4	5	6	7	8	9	10
1	161	200	216	225	230	234	237	239	241	242
2	18,5	19,0	19,2	19,2	19,3	19,3	19,4	19,4	19,4	19,4
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,3
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99
120	3,92	3,07	2,68	2,45	2,29	2,18	2,09	2,02	1,96	1,91
∞	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83

Таблица 4(окончание)

95% -ные квантили распределения Фишера

 $(k_1$ - число степеней свободы числителя, k_2 - число степеней свободы знаменателя)

	12	15	20	24	30	40	60	120	∞
1	244	246	248	249	250	251	252	253	254
2	19,4	19,4	19,4	19,4	19,5	19,5	19,5	19,5	19,5
3	8,74	8,70	8,66	8,64	8,62	8,59	8,57	8,55	8,53
4	5,91	5,86	5,8	5,77	5,75	5,72	5,69	5,66	5,63
5	4,68	4,62	4,56	4,53	4,50	4,46	4,43	4,40	4,37
6	4,00	3,94	3,87	3,84	3,81	3,77	3,74	3,70	3,67
7	3,57	3,51	3,44	3,41	3,38	3,34	3,30	3,27	3,23
8	3,28	3,22	3,15	3,12	3,08	3,04	3,01	2,97	2,93
9	3,07	3,01	2,94	2,90	2,86	2,83	2,79	2,75	2,71
10	2,91	2,85	2,77	2,74	2,70	2,66	2,62	2,58	2,54
11	2,79	2,72	2,65	2,61	2,57	2,53	2,49	2,45	2,40
12	2,69	2,62	2,54	2,51	2,47	2,43	2,38	2,34	2,30
13	2,60	2,53	2,46	2,42	2,38	2,34	2,3	2,25	2,21
14	2,53	2,46	2,39	2,35	2,31	2,27	2,22	2,18	2,13
15	2,48	2,40	2,33	2,29	2,25	2,20	2,16	2,11	2,07
16	2,42	2,35	2,28	2,24	2,19	2,15	2,11	2,06	2,01
17	2,38	2,31	2,23	2,19	2,15	2,10	2,06	2,01	1,96
18	2,34	2,27	2,19	2,15	2,11	2,06	2,02	1,97	1,92
19	2,31	2,23	2,16	2,11	2,07	2,03	1,98	1,93	1,88
20	2,28	2,20	2,12	2,08	2,04	1,99	1,95	1,90	1,84
21	2,25	2,18	2,10	2,05	2,01	1,96	1,92	1,87	1,81
22	2,23	2,15	2,07	2,03	1,98	1,94	1,89	1,84	1,78
23	2,20	2,13	2,05	2,01	1,96	1,91	1,86	1,81	1,76
24	2,18	2,11	2,03	1,98	1,94	1,89	1,84	1,79	1,73
25	2,16	2,09	2,01	1,96	1,92	1,87	1,82	1,77	1,71
30	2,09	2,01	1,93	1,89	1,84	1,79	1,74	1,68	1,62
40	2,00	1,92	1,84	1,79	1,74	1,69	1,64	1,58	1,51
60	1,92	1,84	1,75	1,70	1,65	1,59	1,53	1,47	1,39
120	1,83	1,75	1,66	1,61	1,55	1,50	1,43	1,35	1,25
∞	1,75	1,67	1,57	1,52	1,46	1,39	1,32	1,22	1,00

Пример. Пусть F - случайная величина, распределенная по закону Фишера $F(3, 5)$. $F_{0,95}(3, 5) = 5,41$, т.е. $P(F < 5,41) = 0,95$ (см. пятая строка, третий столбец).

СПИСОК ЛИТЕРАТУРЫ

Основная литература:

1. Магнус Я.Р. Эконометрика. Начальный курс / Я.Р. Магнус, П.К. Каттышев, А.А. Пересецкий. – М.: Дело, 2000. – 400 с.
2. Прикладная статистика. Основы эконометрика: Учебник для вузов: В 2 т. – 2-е изд., испр. - Т.2: Айвазян С.А. Основы эконометрики / С.А. Айвазян. – М.: ЮНИТИ-ДАНА, 2001. – 432 с.

Дополнительная литература:

1. Green W.H. Econometric analysis / W.H. Greene. – 4th ed. Prentice-Hall, Inc., 2000. – 1004 p.
2. Maddala G.S. Introduction to econometrics / G.S. Maddala. – 3rd ed. John Wiley&Sons, Ltd., 2001. – 636 p.

Электронные ресурсы:

1. Электронный каталог научной библиотеки Воронежского государственного университета [http: //www.lib.vsu.ru/](http://www.lib.vsu.ru/)
2. Социальные и гуманитарные науки. Экономика: Библиографическая база данных. 1986-2002гг. / ИНИОН РАН. – М.: 2003. – (CD-ROM).

ОГЛАВЛЕНИЕ

Предисловие	3
1. Мультиколлинеарность	4
2. Автокоррелированность остатков	10
3. АРНС модели	22
4. Модели для панельных данных	34
5. Модели распределенных лагов	47
6. Рекуррентный МНК	55
7. Многофакторные адаптивные модели	60
8. Модели бинарного выбора	66
Приложение	78
Список литературы	82

Рецензент: Доктор физико-математических наук, профессор, заведующий кафедрой математического моделирования Воронежского государственного университета В.А. Костин

Составители: Давнис Валерий Владимирович
Тинякова Виктория Ивановна
Мокшина Светлана Ивановна
Воищева Ольга Станиславовна
Щекунских Светлана Станиславовна

Редактор Бунина Т.Д.