

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
ДОНЕЦКОЙ НАРОДНОЙ РЕСПУБЛИКИ
ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
«ДОНЕЦКИЙ НАЦИОНАЛЬНЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

Л.Д. СЛЕПНЕВА, Е.А. РЫТИКОВА

ЭКОНОМЕТРИКА

Практикум

для обучающихся образовательных учреждений
высшего профессионального образования
по направлению подготовки 38.03.01 «Экономика»

Донецк

2019

УДК 330.43 (076.5)
ББК 65в6я73
С47

Рекомендовано Ученым советом
ГОУВПО «Донецкий национальный технический университет»
в качестве практикума для обучающихся образовательных учреждений
высшего профессионального образования
(Протокол №5 от 28.06.2018)

Рецензенты:

Корнев М.Н. – доктор экономических наук, доцент, заведующий кафедрой финансовых услуг и банковского дела ГОУ ВПО «Донецкая академия управления и государственной службы при Главе Донецкой Народной Республики;

Гавриленко В.А. – доктор экономических наук, профессор, профессор кафедры бухгалтерского учета и аудита ГОУ ВПО «Донецкий национальный технический университет»

Авторы:

Слепнева Л.Д. – кандидат экономических наук, доцент, доцент кафедры «Финансы и экономическая безопасность» ГОУ ВПО «ДОННТУ»;

Рытикова Е.А. – старший преподаватель кафедры «Финансы и экономическая безопасность» ГОУ ВПО «ДОННТУ»

Слепнева, Л. Д.

С47 Эконометрика : практикум / Л.Д. Слепнева, Е.А. Рытикова; – ГОУВПО «ДОННТУ». Донецк : ГОУВПО «ДОННТУ». 2019. – 187 с. : ил., табл.
ISBN

Практикум – дополнение к учебному пособию по эконометрике – обеспечивает методическую поддержку лабораторных занятий и самостоятельной работы при изучении дисциплины «Эконометрика». Содержит варианты заданий и примеры их выполнения, позволяющие студентам заочной формы обучения самостоятельно проводить необходимые расчеты, а также практические задания и упражнения, способствующие усвоению пройденного материала. Основное внимание уделено изложению особенностей применения эконометрических методов и современных программных средств для анализа экономических процессов.

Пособие предназначено для обучающихся дневной, очно-заочной и заочной форм обучения.

УДК 330.43 (076.5)
ББК 65в6я73

ISBN

© Слепнева Л.Д., Рытикова Е.А., 2019
© ГОУВПО «Донецкий национальный технический университет», 2019

ОГЛАВЛЕНИЕ

ОСНОВНЫЕ ПОЛОЖЕНИЯ	4
Тема 1. ПОСТРОЕНИЕ И АНАЛИЗ МОДЕЛИ ПАРНОЙ РЕГРЕССИИ	11
Тема2. ПОСТРОЕНИЕ И АНАЛИЗ МНОЖЕСТВЕННОЙ РЕГРЕССИОННОЙ МОДЕЛИ	37
Тема3. ТЕСТИРОВАНИЕ И УСТРАНЕНИЕМУЛЬТИКОЛЛИНЕАРНОСТИ	60
Тема4. ИСПОЛЬЗОВАНИЕ ФИКТИВНЫХ ПЕРЕМЕННЫХ В ЭКОНОМЕТРИЧЕСКОМ АНАЛИЗЕ	72
Тема5.ОБОБЩЕННАЯ МОДЕЛЬ ЛИНЕЙНОЙ РЕГРЕССИИ С ГЕТЕРОСКЕДАСТИЧНЫМИ ОСТАТКАМИ (ТЕСТИРОВАНИЕ И УСТРАНЕНИЕ ГЕТЕРОСКЕДАСТИЧНОСТИ)	90
Тема6. ОБОБЩЕННАЯ МОДЕЛЬ ЛИНЕЙНОЙ РЕГРЕССИИ С АВТОКОРРЕЛИРОВАННЫМИ ОСТАТКАМИ (ТЕСТИРОВАНИЕ И УСТРАНЕНИЕ ПОСЛЕДСТВИЙ АВТОКОРРЕЛЯЦИИ)	117
Тема7. СИСТЕМЫ ОДНОВРЕМЕННЫХ УРАВНЕНИЙ	142
ЛИТЕРАТУРА	182
ПРИЛОЖЕНИЕ А. МАТЕМАТИКО-СТАТИСТИЧЕСКИЕ ТАБЛИЦЫ	184

ОСНОВНЫЕ ПОЛОЖЕНИЯ

Настоящий практикум является дополнением к учебному пособию «Эконометрика» [18]. Оба издания следует рассматривать как две части – теоретическую и практическую.

Практикум охватывает основные темы курса. По каждой теме приводится

- решение типовых задач, теоретическое обоснование которого представлено в соответствующем разделе учебного пособия (теоретической части);
- варианты заданий для выполнения лабораторной работы;
- вопросы и практические задания для самостоятельной работы.

Цель преподавания дисциплины «Эконометрика» – приобретение студентами знаний в сфере эконометрического моделирования, умения использовать основные методы количественного анализа экономических явлений.

Дисциплина «Эконометрика» рассматривает вопросы применения современных методов эконометрического моделирования для описания взаимосвязи и прогнозирования экономических показателей.

Цель дисциплины – формирование системы знаний по методологии и инструментарию построения и использования разных типов эконометрических моделей.

Задачи – изучить основные принципы и инструментарий постановки задач, построения эконометрических моделей, методов их оценивания и анализа с целью использования в экономике.

В результате освоения дисциплины студент должен:

знать:

– перспективу и объективную необходимость использования эконометрических методов, возможность их применения как инструмента анализа и познания механизма, действий объективных экономических законов.

уметь:

– формировать и решать задачи повышения эффективности экономики на разных уровнях на основе методов эконометрического анализа.

Перечисленные результаты обучения являются основой для формирования следующих компетенций:

общекультурных:

- способностью к философскому подходу к изучению проблем науки и техники, к абстрактному мышлению, анализу, синтезу, обобщению, восприятию информации, постановке цели и выбору путей ее достижения (ОК-1);

- способностью использовать основы экономических, правовых знаний в различных сферах деятельности (ОК-4);

- готовностью действовать в нестандартных ситуациях, нести социальную и этическую ответственность за принятые решения (ОК-7);

общепрофессиональных:

- готовностью к коммуникации в устной и письменной формах на русском и иностранном языках для решения задач профессиональной деятельности (ОПК-1);

- готовностью работать в коллективе в сфере своей профессиональной деятельности, толерантно воспринимая социальные, этнические, конфессиональные и культурные различия, историческое наследие государства (ОПК-2);

готовности постигать проблемы общенаучного и профессионально-ориентированного характера на основе систематической проработки литературы по специальности (ОПК-3);

- готовностью самостоятельно принимать профессиональные решения на основе использования законодательных и нормативных актов, знания теории и практики (ОПК-4);

- владением и готовностью применять на практике методики по обработке и систематизации научной и практической информации, необходимой для решения профессиональных задач; пользоваться передовым опытом в сфере профессиональной деятельности (ОПК-5);

- способностью выбирать инструментальные средства для обработки экономических данных в соответствии с поставленной задачей, анализировать результаты расчетов и обосновывать полученные выводы (ОПК-6);

- способностью решать стандартные задачи профессиональной деятельности на основе информационной и библиографической культуры с применением информационно-коммуникационных технологий и с учетом умения работать с информацией в глобальных компьютерных сетях, соблюдением основных требований информационной безопасности, в том числе защиты государственной тайны (ОПК-7);

- способностью использовать необходимые формы повышения квалификации, образовательного и профессионального уровня, деловой квалификации (ОПК-8);

- способностью принимать организационно-управленческие решения в профессиональной деятельности и готовностью нести за них ответственность (ОПК-9).

профессиональных:

- способностью собирать и анализировать исходные данные, необходимые для расчета экономических и социально-экономических показателей, характеризующих деятельность хозяйствующих субъектов (ПК-1);

- способностью на основе типовых методик и действующей нормативно-правовой базы рассчитывать экономические, финансовые и социально-экономические показатели, характеризующие деятельность хозяйствующих субъектов (ПК-2);

- способностью выполнять необходимые для составления экономических и финансовых разделов планов расчеты, обосновывать их и представлять результаты работы в соответствии с принятыми в организации стандартами (ПК-3);

- способностью собирать и анализировать исходные данные, характеризующих финансовую деятельность учреждений, организаций,

предприятий различных организационно-правовых форм, включая органы государственной власти и местного самоуправления (ПК-4);

- способностью использовать статистические методы анализа при проведении анализа финансовых результатов предприятия, при оценке изменения стоимости, при определении эквивалентности процентных ставок, при проведении анализа финансовых потоков, при оценке долгосрочной задолженности и т.д. (ПК-5);

- способностью под контролем осуществлять стратегическое, тактическое и оперативное планирование и прогнозирование финансово-экономических показателей деятельности предприятий, учреждений, организаций различных организационно-правовых форм, включая органы государственной власти и местного самоуправления (ПК-11);

- способностью на основе описания экономических процессов и явлений строить стандартные теоретические и эконометрические модели, анализировать и содержательно интерпретировать полученные результаты с использованием современных технических средств и информационных технологий (ПК-15);

- способностью организовать научные исследования в области экономики: выбирать и обосновывать тему, составлять план исследований; уметь использовать информационное обеспечение; выбрать инструментальные средства для обработки экономических данных в соответствии с поставленной задачей, проанализировать результаты расчетов и обосновать полученные выводы; составлять и оформлять библиографию, применять компьютерные технологии в научных исследованиях и т.д. (ПК-18);

- способностью проводить анализ и давать оценку существующих финансово-экономических рисков, составлять и обосновывать прогноз динамики основных финансово-экономических показателей на микро-, макро- и мезоуровне (ПК-19);

- способностью принимать участие в разработке рабочих планов, программ и инструментов проведения научных исследований в области экономики, международной экономики, финансов, кредита, учета и аудита, анализ их

результатов, подготовка данных для составления финансовых обзоров, отчетов и научных публикаций (ПК-27);

- способностью осуществлять сбор, обработку, анализ и систематизацию информации по теме исследования, выбор методов и средств решения задач исследования (ПК-28);

- способностью принимать участие в разработке теоретических и новых эконометрических моделей исследуемых процессов, явлений и объектов, относящихся к сфере профессиональной деятельности по направлению подготовки, давать оценку и интерпретировать полученные в ходе исследования результаты (ПК-29);

- способностью выбирать цели, методы, задачи товарной, ценовой, коммуникационной политики предприятия на рынке товаров и услуг (ПК-37);

- способностью преподавать экономические дисциплины в образовательных организациях среднего общего образования, используя существующие программы и учебно-методические материалы (ПК-38);

- способностью принимать участие в разработке образовательных программ и учебно-методических материалов по экономическим дисциплинам (ПК-39).

Поскольку при изучении данной дисциплины в центре внимания находятся практические аспекты использования инструментария эконометрики для описания и моделирования социально-экономических явлений и процессов, то для закрепления полученных знаний и приобретения практических навыков, необходимых для решения задач эконометрического моделирования с применением ПК при обработке исходных данных, выполняются лабораторные работы.

В результате выполнения лабораторных работ студент:

должен знать:

- основные методические подходы и приемы изучения экономических процессов и явлений;
- методы статистического анализа;

- способы определения степени и характера влияния отдельных факторов на экономические показатели;
- методы контроля правильности решения;
- методы количественного обоснования экономических и управленческих решений;

должен уметь:

- владеть методикой сбора, обработки экономической информации и прогнозировать состояние и развитие экономических процессов;
- применять общие и специальные методы экономических и статистических расчетов,
- выбирать наиболее подходящий метод исследования из спектра имеющихся;
- доводить решение задач до практически приемлемых результатов;
- работать с современными программными средствами, предназначенными для решения экономико-математических задач.

Целью выполнения лабораторных работ является более глубокое усвоение материала, а также развитие практических навыков, направленных на реализацию вычислительных процедур.

Перед выполнением заданий рекомендуется внимательно изучить теоретический материал соответствующей темы, изложенный в учебном пособии по эконометрике [18], ответить на содержащиеся в ней вопросы для самопроверки, провести самостоятельный анализ всех разобранных примеров.

Требования к оформлению отчета по лабораторной работе

Оформление отчета по лабораторной работе осуществляется студентом самостоятельно во время или после занятий в установленные преподавателем сроки.

Полный отчет по лабораторной работе должен содержать:

- 1) титульный лист (с указанием всех необходимых реквизитов);
- 2) постановку экономической задачи (задач);

3) содержательный отчет по каждому пункту задания. Обязательно соблюдение нумерации пунктов задания. Каждый пункт исследования необходимо завершать содержательной экономической интерпретацией полученных результатов;

4) при построении графиков и гистограмм подписываются оси (в том числе указываются единицы измерения) и числовые метки. Графики, помимо этого, должны содержать титульные подписи, чтобы было понятно, что на них изображено.

Не следует приводить лишние результаты! Некоторые второстепенные таблицы и графики можно опускать (главное, чтобы они содержались в рабочем файле!).

К экзамену допускаются студенты, выполнившие все пункты задания, оформившие результаты работы в соответствии с установленными требованиями и получившие зачет по всем лабораторным работам.

Для получения зачета студент должен:

- знать теоретические основы тематики лабораторной;
- уметь ответить на вопросы преподавателя по содержанию выполненной лабораторной работы.

ТЕМА 1. ПОСТРОЕНИЕ И АНАЛИЗ МОДЕЛИ ПАРНОЙ РЕГРЕССИИ

Данная тема охватывает большой круг вопросов, связанных с эконометрическим оцениванием взаимосвязи двух экономических показателей. Это оценка тесноты и выбор формы связи рассматриваемых показателей, применение статистических методов оценивания (в частности, метода наименьших квадратов), проверка качественных характеристик модели, осуществление прогноза изменения результативного признака.

Цель выполнения лабораторной работы: научиться самостоятельно применять формулы МНК-оценщика для расчета оценок параметров линейной и нелинейной парной регрессии, а также связанных с этим величин; интерпретировать и критически оценивать результаты применения данных формул.

Для выполнения лабораторной работы № 1 студенты должны

знать:

- формулы оценок параметров линейной регрессии;
- формулы оценок стандартных ошибок параметров регрессии;
- формулы оценок коэффициентов корреляции (парного и множественного);
- формулы оценки коэффициента детерминации;
- t – критерий оценки статистической надежности оценок параметров модели;
- F – критерий оценки адекватности модели статистическим данным;
- формулу оценки коэффициента эластичности;

уметь:

- пользоваться таблицами для проверки статистических гипотез;
- пользоваться пакетом EXCEL:
 - строить таблицы; составлять и копировать формулы;
 - пользоваться МАСТЕРОМ ФУНКЦИЙ: статистическими и математическими функциями.

Задание для выполнения лабораторной работы №1

В табл. 1.1 приведены показатели, характеризующие деятельность российских банков по состоянию на 1 марта 2005 г., и их количественные характеристики. По имеющимся данным (табл. 1.1) для соответствующего варианта (табл. 1.2):

1. Сформулировать гипотезу о взаимосвязи рассматриваемых показателей, определив, какой из них будет рассматриваться как зависимая переменная, и какой – в качестве факторного признака.
2. Рассчитать показатели вариации каждого признака, отграничить однородную совокупность единиц, устранив аномальные объекты наблюдения, а также единицы с неполной информацией.
3. Вычислить выборочный коэффициент корреляции и оценить его значимость на 5 % уровне.
4. Построить диаграмму рассеяния заданных переменных и сформулировать гипотезу о виде связи. Оценить параметры уравнения парной регрессионной модели методом наименьших квадратов.
5. Оценить качество построенного уравнения регрессии с помощью коэффициента детерминации R^2 , а также статистическую значимость построенного уравнения регрессии с помощью t и F -критерия.
6. Построить 95 % -е интервалы для оценок параметров уравнения регрессии.
7. Оценить с помощью коэффициента эластичности (среднего) силу связи независимой переменной с зависимой.
8. Построить точечный и интервальный прогноз величины рассматриваемого экономического показателя, характеризующего деятельность системы банков и отдельного банка, при условии прироста влияющего фактора на 10% от его среднего значения.
9. Оценить параметры уравнений регрессий других форм: степенной, логарифмической, экспоненциальной и др.
10. Определить лучшее уравнение регрессии по величине стандартной ошибки регрессии, на основе средней ошибки аппроксимации и с помощью скорректированного коэффициента детерминации.

Таблица 1.1 – Показатели российских банков на 1 марта 2005 г. [12]

Банк	Работающие активы, млн. руб.	Собственный капитал, млн. руб.	Средства частных лиц, млн. руб.	Кредиты частным лицам, млн. руб.	Средства предприятий и организаций, млн. руб.	Кредиты предприятиям и организациям, млн. руб.	Акции, млн. руб.	Облигации, млн. руб.	Векселя, млн. руб.
1	2	3	4	5	6	7	8	9	10
1.Сбербанк	1917403	209933	1235105	308437	389016	1073255	13571	359499	2
2.Внешторгбанк	426484	72057	58557	5205	111103	189842	23152	50012	24859
3.Газпромбанк	362532	30853	33574	5084	141437	207118	18660	35676	6107
4.Альфа-банк	186700	25581	30001	1361	58489	138518	4505	8471	3153
5.Банк Москвы	157286	18579	49300	5768	44636	90757	3026	24838	873
6.Росбанк	151849	12879	31989	4466	93007	62388	4474	5667	4538
7.Ханты-Мансийский банк	127440	3345	6310	1392	11655	4142	406	15601	6231
8.МДМ-банк	111285	13887	9903	7266	28779	51731	2656	13186	3987
9.ММБ	104372	8380	10871	4119	66525	48400	721	14213	59
10.Райффайзенбанк	96809	7572	21602	10828	42129	46393	284	5273	0
11.Промстройбанк	85365	9528	22829	2719	27362	45580	2781	18727	2480
12.Ситибанк	81296	8953	10401	3576	38895	33339	13	23442	0
13.Уралсиб	76617	13979	18665	8170	15953	43073	6705	4026	784
14.Межпромбанк	67649	28770	1072	511	5457	60154	63	2577	1185
15.Промсвязьбанк	54848	5222	6404	822	26302	32761	68	5250	6009
16.Петрокоммерц	53701	8373	14844	1693	21544	23053	3561	9417	863
17.Номос-банк	52473	6053	3556	476	9488	28511	2126	9416	2100
18.Зенит	50666	7373	5075	421	18923	25412	2743	8264	1357
19.Русский стандарт	46086	9078	3246	38799	599	3599	0	377	370
20.Транскредитбанк	41332	3768	3495	993	19923	18506	827	7350	322
21.Ак Барс	40521	10000	8442	2811	16607	23841	1586	4353	1158
22.Глобэкс	40057	11163	6739	46	8685	29420	330	872	0
23.Еврофинанс- Моснарбанк	38245	6077	2145	171	9127	18114	400	7949	751
24.Никойл	36946	8790	4355	245	9040	13117	9160	5231	349
25.Автобанк-Никойл	34762	7094	12494	1773	8435	19135	5174	3238	612
26.Импэксбанк	34032	4806	14025	3993	7759	15047	4098	2584	1867
27.Союз	33062	4579	3028	3254	12213	15507	3172	5187	3641
28.БИН-банк	32948	4218	6736	764	12013	24980	17	2172	192
29.Возрождение	30713	3024	16244	1218	9826	20665	459	294	0
30.Гута-банк	30596	3350	4083	906	5178	11556	531	1601	7378
31.Менатеп СПб	29970	3422	4953	1226	20277	7593	157	2541	807
32.Коммерцбанк	26724	3695	0	4	5623	18158	0	1809	0
33.ХКФБ	26388	2995	1598	22267	7177	28	0	67	17
34.Дойче банк	26015	3289	1637	0	5910	2014	1	8546	550
35.АБН Амро банк	25691	2855	4488	31	17569	11044	1	2828	0
36.ПЧРБ	25206	3987	1092	23	15893	5415	1134	361	1468
37.МБРР	24639	3146	2062	311	13103	14216	6	1105	2072
38.НРБ	24027	7436	922	610	4053	10408	5254	3156	909
39.Россельхозбанк	23863	5088	3360	1178	5733	13953	102	1628	858
40.Сургутнефтегазбанк	22894	2359	11974	2600	9341	3254	307	4239	151

Банк	Работающие активы, млн. руб.	Собственный капитал, млн. руб.	Средства частных лиц, млн. руб.	Кредиты частным лицам, млн. руб.	Средства предприятий и организаций, млн. руб.	Кредиты предприятиям и организациям, млн. руб.	Акции, млн. руб.	Облигации, млн. руб.	Векселя, млн. руб.
1	2	3	4	5	6	7	8	9	10
41.Кредит Свисс	22570	3703	3	1	14367	8	1	3035	0
42.Собинбанк	20905	5332	3292	1680	5158	15405	18	322	70
43.Траст	19513	5391	746	209	5591	2584	921	9397	465
44.Запсибкомбанк	19346	1966	4273	3989	3859	8586	226	2009	1164
45.Судостроительный банк	18991	3105	320	179	4486	6811	20	950	3315
46.Банк Санкт-Петербург	18389	1966	5375	240	7428	11911	140	2862	352
47.ИНГ банк	18090	2946	1086	20	10375	11672	1	1921	0
48.Балтийский банк	17674	2480	10000	759	4882	11422	17	1057	251
49.МИНБ	16965	2221	7265	746	6655	11788	15	1886	295
50.ВестЛБ Восток	16910	1116	6	3	5773	1169	1	4125	194
51.Авангард	16070	3169	2044	727	3174	11839	718	227	0
52.Российский кредит	15474	4845	36	11	988	1406	7343	762	953
53.Кредитагропромбанк	15332	1399	2357	3153	7603	5334	40	622	1936
54.Инвестсбербанк	15326	2650	4192	3085	5436	6249	814	1192	719
55.Сосьете Женераль Восток	14957	1417	2644	1337	7862	9128	5	0	0
56.Русь-банк	14555	2579	428	843	8145	9710	137	1136	364
57.Пробизнесбанк	13026	1729	1232	548	4089	6913	794	680	6
58.Национальный стандарт	12545	1693	0	28	8390	2233	1	2975	2822
59.ВБРР	12070	1265	565	106	7553	5881	292	2683	655
60.Татфондбанк	11949	2817	2589	544	3549	9897	69	561	67
61.МБСП	11889	1573	1583	295	5704	5404	12	4676	984
62.Абсолют банк	11831	1524	1308	639	2626	7872	40	413	275
63.Визави	11768	4719	127	183	1264	7208	257	483	0
64.Центрокредит	11674	3878	668	364	3859	5097	864	3438	348
65.Связь-банк	11250	2307	987	532	8002	3692	236	1534	2304
66.Россия	11048	1404	1752	260	6698	5893	61	2687	300
67.БНП Париба	10911	2330	0	0	1770	6638	0	409	0
68.МКБ	10864	2042	1465	1596	2002	6739	0	603	479
69.Газэнергопромбанк	10848	961	2286	897	6477	6295	24	246	941
70.Росевробанк	10830	1680	1473	161	5981	5372	489	507	59
71.Северная казна	10792	1120	5544	1694	2883	4915	249	1552	1074
72.МДМ-банк СПб	10645	976	5075	677	2204	1222	1	650	0
73.РБР	10533	1423	4723	314	3071	5031	101	784	3286
74.Вэб-инвестбанк	10503	2687	59	343	6749	3784	184	3889	87
75.Транскапиталбанк	10497	1307	1888	466	3662	5660	283	857	823
76.Кредит Урал банк	10400	2170	3498	966	4305	4621	0	0	6

Таблица 1.2 - Варианты моделей к лабораторной работе № 1

Вариант	Столбцы из таблицы 1.1	Виды моделей				
		Линейная	Степенная	Показательная	Линейно-логарифмическая	Гиперболическая
1	2	3	4	5	6	7
1	2, 3	*	*		*	
2	2, 4	*	*		*	
3	2, 6	*	*		*	
4	2, 8	*	*		*	
5	2, 9	*	*		*	
6	2, 10	*		*		*
7	3, 4	*		*		*
8	3, 5	*		*		*
9	3, 6	*		*		*
10	3, 7	*		*		*
11	3, 8	*	*		*	
12	3, 9	*	*		*	
13	3, 10	*	*		*	
14	5, 3	*	*		*	
15	5, 4	*	*		*	
16	5, 6	*		*		*
17	5, 8	*		*		*
18	5, 9	*		*		*
19	5, 10	*		*		*
20	7, 3	*		*		*
21	7, 4	*	*		*	
22	7, 6	*	*		*	
23	7, 8	*	*		*	
24	7, 9	*	*		*	
25	7, 10	*	*		*	

Для выполнения работы следует изучить теоретический материал, представленный в теме 2 Учебного пособия по эконометрике [18].

РЕШЕНИЕ ТИПОВЫХ ЗАДАЧ

Пример 1.1. Выделение аномальных наблюдений.

Аномальными наблюдениями (выбросами) называются элементы совокупности (выборки), значительно отличающиеся от остальных по значениям изучаемого признака. Для их выявления можно использовать z-оценки и модифицированные z-оценки.

Метод z-оценок подходит для больших выборок, соответствующих нормальному закону распределения. Из свойств нормального распределения известно, что 99,7% значений случайной величины, находящихся в рамках распределения, отклоняются от среднего значения не более, чем на 3σ . Таким образом, для i -го наблюдения рассчитывается значение z_i :

$$z_i = \frac{x_i - \bar{x}}{\sigma},$$

где z_i – стандартизованное значение переменной, рассчитанное для подозрительного наблюдения; x_i – значение подозрительного наблюдения, которое тестируется на аномальность; \bar{x} – среднее значение по совокупности (характеристика центра распределения); σ – СКО по совокупности (стандартное отклонение – характеристика разброса).

Если для i -го наблюдения полученное значение z_i превышает 3, то его можно считать аномальным.

Для вычисления *z-оценок* можно воспользоваться функцией НОРМАЛИЗАЦИЯMicrosoftExcel, которая находится в категории **Статистические**:

НОРМАЛИЗАЦИЯ(х;среднее;стандартное_откл),

предварительно рассчитав аргументы функции:

- **х** Нормализуемое значение – выделяется массив данных, для которых надо найти нормализованные значения;
- **Среднее**. Среднее арифметическое распределения – функция СРЗНАЧ (см. рис. 1.1);
- **Стандартное_откл.** Стандартное отклонение распределения – функция СТАНДОТКЛОН (см. рис. 1.1).

Модифицированная z-оценка основана на медиане:

$$z_{Mi} = \frac{0,6745(x_i - Me)}{MED},$$

где z_{Mi} – фактическая величина критерия, рассчитанная для подозрительного наблюдения; x_i – значение подозрительного наблюдения, которое тестируется на аномальность; Me – медианное значение по совокупности данных, характеристика центра распределения; MED – медиана модуля отклонения значений признака от медианы Me .

Для этого показателя значение, превышающее 3,5, считается аномальным.

Для вычисления медианы Me можно использовать функцию МЕДИАНА MicrosoftExcel из категории *Статистические*, указав в качестве аргумента массив значений, для которого вычисляется медиана.

Расчет z_{Mi} представлен на рис. 1.1. В столбцах С, Е, G рис. 1.1 показаны формулы для расчета величин столбцов D, F, H соответственно.

A	B	C	D	E	F	G	H
Модифицированный z-метод	Y						Z_{Mi}
	17674	=B2-\$B\$13	12127	=ABS(D2)	12127	=0,6745*D2/\$F\$13	3,679
	16965	=B3-\$B\$13	11418	=ABS(D3)	11418	=0,6745*D3/\$F\$13	3,464
	16910	=B4-\$B\$13	11363	=ABS(D4)	11363	=0,6745*D4/\$F\$13	3,447
	16070	=B5-\$B\$13	10523	=ABS(D5)	10523	=0,6745*D5/\$F\$13	3,193
	15474	=B6-\$B\$13	9927	=ABS(D6)	9927	=0,6745*D6/\$F\$13	3,012

	2820	=B8-\$B\$13	-2727	=ABS(D8)	2727	=0,6745*D8/\$F\$13	-0,827
	2809	=B9-\$B\$13	-2738	=ABS(D9)	2738	=0,6745*D9/\$F\$13	-0,831
	2776	=B10-\$B\$13	-2771	=ABS(D10)	2771	=0,6745*D10/\$F\$13	-0,841
=СРЗНАЧ(B2:B160)	6617,22						
=СТАНДОТКЛОН(B2:B160)	3620,32						
=МЕДИАНА(B2:B160)	5547			=МЕДИАНА(F2:F160)	2223	MED	

Рис. 1.1. Вычисление модифицированной z-оценки

Таким образом, аномальным значением следует признать величину 17674, поскольку ее $z_M = 3.679 > 3,5$.

Модифицированная z-оценка часто выделяет гораздо больше аномальных значений (если медиана имеющегося распределения меньше средней арифметической, то есть, расстояние от каждого значения в правой части распределения до медианы будет больше, чем до средней).

Пример 1.2. На основе информации, характеризующей деятельность 17 предприятий одной отрасли (табл. 1.3), построить регрессионные модели зависимости чистой прибыли от стоимости основных производственных фондов (ОПФ), выбрать «лучшую» модель и сделать прогноз.

Таблица 1.3 – Исходные данные [9]

№ п/п (i)	Чистая прибыль, млн. руб. (Y)	Основные фонды, млн. руб. (X)	№ п/п (i)	Чистая прибыль, млн. руб. (Y)	Основные фонды, млн. руб. (X)
1	13,2	21,4	10	30,57	26,87
2	8,16	19,27	11	33,1	25,98
3	15,14	19,21	12	38,65	29,5
4	7,45	15,34	13	40,54	28,57
5	18,79	24,45	14	39,62	29,41
6	21,52	23,21	15	41,3	30,23
7	25,33	21,12	16	41,95	30,5
8	26,78	24,4	17	42,69	31,1
9	28,72	25,44			

Изучая взаимосвязь между двумя показателями – величиной чистой прибыли и стоимостью основных производственных фондов предприятий, – в качестве зависимой переменной выберем чистую прибыль (обозначим ее через Y), а независимой – стоимость основных производственных фондов – (обозначим через X). Представим зависимость графически: построим диаграмму рассеяния признаков (рис. 1.2).

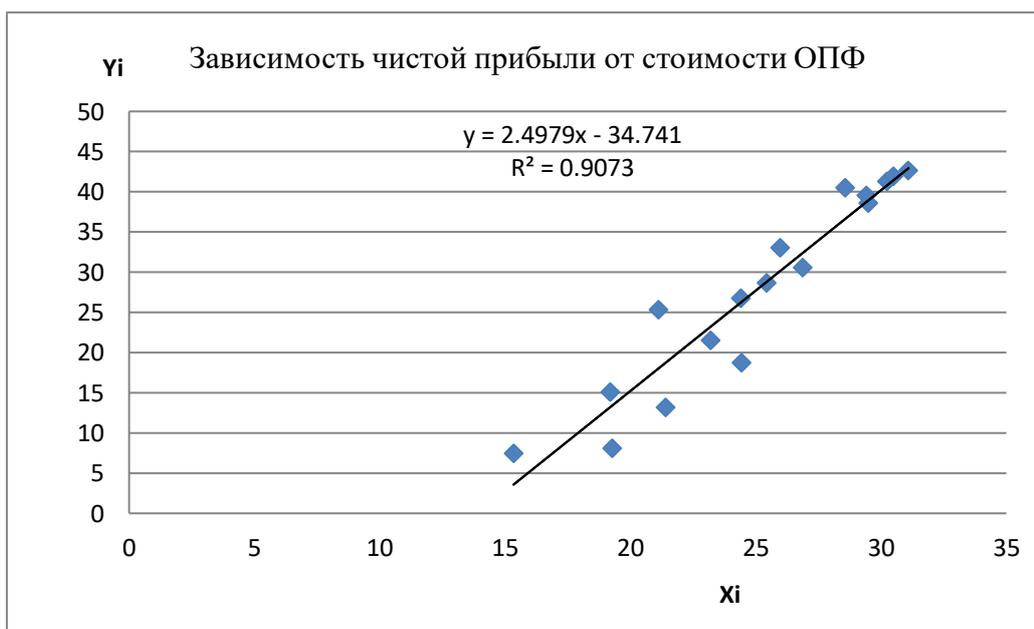


Рис. 1.2 – Диаграмма рассеяния признаков

Анализ диаграммы рассеяния позволяет сделать вывод о наличии тесной прямой связи между изменением чистой прибыли и стоимости ОПФ (то есть, величина чистой прибыли и стоимость основных производственных фондов предприятий рассматриваемой отрасли изменяются в одном направлении).

Для описания этой связи можно использовать как линейную, так и нелинейные функции.

Для построения различных моделей, характеризующих влияние изменения стоимости ОПФ на изменчивость чистой прибыли, выполним промежуточные расчеты (см. табл. 1.4), результаты которых используем при вычислениях следующих показателей.

1. ЛИНЕЙНАЯ МОДЕЛЬ

1) *Найдем коэффициент корреляции:*

$$r_{YX} = \frac{n \sum_{i=1}^n X_i Y_i - \sum_{i=1}^n X_i \sum_{i=1}^n Y_i}{\sqrt{n \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2} \sqrt{n \sum_{i=1}^n Y_i^2 - \left(\sum_{i=1}^n Y_i \right)^2}} = \frac{17 \cdot 12723,39 - 426 \cdot 473,51}{\sqrt{17 \cdot 11018,46 - 426^2} \sqrt{17 \cdot 15550,46 - 473,51^2}} = 0,953$$

Величина коэффициента корреляции подтверждает вывод графического анализа (рис. 1.2) о наличии тесной прямой связи между чистой прибылью и стоимостью ОПФ.

Проверим статистическую надежность коэффициента корреляции:

$$t = \frac{r \sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0,953 \sqrt{17-2}}{\sqrt{1-0,953^2}} = 12,1826 > t_{0,05;15} = 2,13$$

Гипотеза о статистической незначимости коэффициента корреляции с вероятностью 0,95 отклоняется в пользу альтернативной.

2) **Определим оценки параметров модели**, используя данные табл. 1.4.

$$b_1 = \frac{n \sum_{i=1}^n X_i Y_i - \sum_{i=1}^n X_i \sum_{i=1}^n Y_i}{n \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2} = \frac{17 \cdot 12723,39 - 426 \cdot 473,51}{17 \cdot 11018,46 - 426^2} = 2,498,$$

$$b_0 = \bar{Y} - b_1 \bar{X} = \frac{473,51}{17} - 2,498 \frac{426}{17} = -34,741.$$

Таким образом, уравнение парной линейной регрессии с оцененными параметрами имеет вид

$$\hat{Y} = -34,741 + 2,498X.$$

Коэффициент регрессии $b_1 = 2,498$ является по своей сути **средним абсолютным приростом**. Его значение показывает, что увеличение стоимости ОПФ на 1 млн. грн. способствует увеличению чистой прибыли в среднем на 2,498 млн. руб.

3) В табл. 1.5 приведены результаты оценивания модели с применением функции ЛИНЕЙН. Используя соответствующую информацию, **сделаем выводы о статистической значимости модели**.

Имеем $b_0 = -34,741$, $b_1 = 2,498$,

$$s_{b_0} = 5,2476, \quad s_{b_1} = 0,2061, \quad t_0 = \frac{|-34,741,51|}{5,2476} = 6,62, \quad t_1 = \frac{2,498}{0,2061} = 12,12$$

Табличное значение статистики Стьюдента при уровне значимости 5% и числе степеней свободы, равном $17 - 2 = 15$, составляет: $t_{5\%, \nu=15} = 2,13$. Сравнивая вычисленные t-статистики с табличным значением, делаем вывод о статистической надежности коэффициентов.

4) **Величина коэффициента детерминации** $R^2 = 0,9073$ говорит о том, что на 90,73% модель характеризует изменчивость исследуемого показателя, и только 9,27% дисперсии чистой прибыли объясняется влиянием неучтенных факторов.

F-статистика = 146,86 больше критического значения $F_{5\%;1;15} = 4,54$, следовательно, коэффициент детерминации и модель в целом с вероятностью 0,95 можно считать статистически значимыми.

5) **Построим доверительные интервалы для параметров модели:**

$$b_j - \Delta_j \leq \beta_j \leq b_j + \Delta_j,$$

$$\Delta_0 = t_{\alpha, \nu} s_0 = 2,13 \cdot 5,2476 = 11,185 \quad \Delta_1 = t_{\alpha, \nu} s_1 = 2,13 \cdot 0,2061 = 0,439$$

$$-34,741 - 11,185 \leq \beta_0 \leq -34,741 + 11,185 \quad 2,498 - 0,439 \leq \beta_1 \leq 2,498 + 0,439$$

$$-45,926 \leq \beta_0 \leq -23,556 \quad 2,059 \leq \beta_1 \leq 2,937$$

Поскольку в границы доверительных интервалов ноль не попал, то оцениваемые параметры считаются статистически надежными. Таким образом, при изменении стоимости ОПФ на 1 млн. руб. величина чистой прибыли с вероятностью 0,95 изменится не менее, чем на 2,059 млн. руб. и не более, чем на 2,937 млн. руб.

6) **Вычислим коэффициент эластичности:**

$$E_{Y/x} = b_1 \frac{\bar{X}}{\bar{Y}} = 2,498 \frac{426/17}{473,51/17} = 2,247$$

То есть, рост стоимости ОПФ на 1% вызывает в среднем рост чистой прибыли на 2,25%.

7) **Определим среднюю относительную ошибку аппроксимации:**

$$A = \frac{1}{n} \cdot \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \cdot 100 \% \approx 0,8 \cdot \frac{S_{\text{пер}}}{\bar{y}} \cdot 100 \% = 0,8 \cdot \frac{1,19}{23,7} \cdot 100 \% \approx 4,0 \%,$$

где $\bar{y} = 23,7$ млн. руб. — средний объем выпускаемой продукции, определенный с помощью встроенной функции «СРЗНАЧ» (см. «Исходные данные» в *прил. 1*).

2. НЕЛИНЕЙНЫЕ МОДЕЛИ

Выполним *оценивание параметров нелинейных моделей* регрессии методом наименьших квадратов. При вычислении оценок будем использовать результаты предварительных расчетов, собранные в табл. 1.4. Кроме того, в табл. 1.5 помещены характеристики моделей, полученные с помощью Microsoft Excel – функции ЛИНЕЙН. Используя данные табл. 1.5, сделаем выводы относительно качественных характеристик моделей.

1. *Экспоненциальная модель* $Y = e^{\beta_0 + \beta_1 X} \varepsilon$,

где $e=2,718\dots$ – основание натуральных логарифмов; $\exp(a) = e^a$ – функция экспоненты (в Excel встроенная функция «EXP»).

Данная модель относится к нелинейным относительно оцениваемых параметров, так как она включает параметры β_0 и β_1 не аддитивно. Но путем логарифмирования уравнение приводится к линейному виду:

$$\ln Y = (\beta_0 + \beta_1 X) \ln e + \ln \varepsilon, \text{ или (с учетом } \ln(e) = 1) \ln Y = \beta_0 + \beta_1 X + \ln \varepsilon$$

Расчет оценок параметров выполним по формулам:

$$b_1 = \frac{n \sum X_i \ln Y_i - \sum X_i \cdot \sum \ln Y_i}{n \sum X_i^2 - (\sum X_i)^2} = \frac{17 \cdot 1403,321 - 426 \cdot 54,451}{17 \cdot 11018,46 - 426^2} = 0,113,$$

$$b_0 = \frac{1}{n} \sum \ln Y_i - b_1 \bar{X} = \frac{1}{17} \cdot 54,451 - 0,113 \cdot \frac{426}{17} = 0,370.$$

Тогда модель, характеризующая изменение чистой прибыли под влиянием изменения ОПФ представляется следующим образом:

$$\hat{Y} = e^{0,370 + 0,113X}. \quad (1.1)$$

Представим (1.1) в следующем эквивалентном виде:

$$\hat{Y} = e^{0,370 + 0,113X} = e^{0,370} \cdot e^{0,113X} = 1,447 e^{0,113X} = 1,447 \left(e^{0,113} \right)^X = 1,447 \cdot 1,12^X \quad (1.2)$$

2. Показательная функция $Y = \beta_0 \beta_1^X \varepsilon$ ($\beta_0 > 0$, $\beta_1 > 0$).

Приведем уравнение к линейному виду:

$$\ln Y = \ln \beta_0 + X \ln \beta_1 + \ln \varepsilon.$$

Оцениваем параметры:

$$\ln b_1 = \frac{n \sum X_i \cdot \ln Y_i - \sum X_i \sum \ln Y_i}{n \sum X_i^2 - (\sum X_i)^2} = \frac{17 \cdot 1403,321 - 426 \cdot 54,451}{17 \cdot 11018,46 - 426^2} = 0,113,$$

$$b_1 = e^{0,113} = 1,12$$

$$\ln b_0 = \frac{1}{n} \sum \ln Y_i - b_1 \frac{1}{n} \sum X_i = \frac{1}{17} \cdot 54,451 - 0,113 \cdot \frac{426}{17} = 0,370,$$

$$b_0 = e^{0,370} = 1,447$$

$$\hat{Y} = 1,447 \cdot 1,12^X. \quad (1.3)$$

Видим, что уравнения (1.2) и (1.3) идентичны.

Практическое применение экспоненты (показательной кривой) возможно, если все значения зависимой переменной Y положительны. Поэтому, если исследуется, например, финансовые результаты деятельности предприятий, среди которых наряду с прибыльными есть и убыточные, то данная функция не может быть использована.

Коэффициент $b_1 = 1,12$ является *средним коэффициентом роста*. Его значение показывает, что при увеличении стоимости ОПФ (X) на 1 млн. руб. величина чистой прибыли (Y) возрастает в среднем в 1,12 раза, то есть на 12 %.

3. Степенная функция: $Y = \beta_0 X^{\beta_1} \varepsilon$, ($\beta_0 > 0$).

$$\ln Y = \ln \beta_0 + \beta_1 \ln X + \ln \varepsilon.$$

Формулы для расчета оценок

$$b_1 = \frac{n \sum \ln X_i \cdot \ln Y_i - \sum \ln X_i \sum \ln Y_i}{n \sum \ln X_i^2 - (\sum \ln X_i)^2} = \frac{17 \cdot 176,116 - 54,460 \cdot 54,452}{17 \cdot 175,099 - 54,46^2} = 2,651,$$

$$\ln b_0 = \frac{1}{n} \sum \ln Y_i - b_1 \frac{1}{n} \sum \ln X_i = \frac{54,452}{17} - 2,651 \cdot \frac{54,460}{17} = 5,291,$$

$$b_0 = e^{\ln b_0} = e^{5,291} = 0,005.$$

$$\hat{Y} = 0,05 X^{2,651}$$

Коэффициент регрессии (показатель степени) является средним коэффициентом эластичности и показывает, что рост стоимости ОПФ на 1% вызывает рост чистой прибыли в среднем на 2,651%.

4. Линейно-логарифмическая модель $Y = \beta_0 + \beta_1 \ln X + \varepsilon$.

$$b_1 = \frac{n \sum \ln X_i \cdot Y_i - \sum \ln X_i \cdot \sum Y_i}{n \sum \ln X_i^2 - (\sum \ln X_i)^2} = \frac{17 \cdot 1553,158 - 54,46 \cdot 54,451}{17 \cdot 175,099 - 54,46^2} = 57,274$$

$$\ln b_0 = \frac{1}{n} \sum Y_i - b_1 \frac{1}{n} \sum \ln X_i = \frac{54,451}{17} - 57,274 \cdot \frac{54,46}{17} = -155,626$$

$$\hat{Y} = -155,626 + 57,274 \ln X \quad (1.4)$$

Эта модель обычно используется, когда необходимо исследовать влияние процентного изменения объясняющей переменной X на абсолютное изменение зависимой переменной. Здесь коэффициент β_1 определяет изменение переменной Y при единичном приросте X (например, на 1%), т.е. характеризует отношение абсолютного изменения Y к относительному изменению X .

Из (1.4) следует, что если стоимость ОПФ увеличится на 1%, то величина чистой прибыли в среднем вырастет на $0,01\beta_1 = 0,5727$ млн.руб.

5. Гиперболическая зависимость $Y = \beta_0 + \beta_1 \frac{1}{X} + \varepsilon$

Формулы для расчета оценок:

$$b_1 = \frac{n \sum_{i=1}^n \frac{1}{X_i} Y_i - \sum_{i=1}^n \frac{1}{X_i} \sum_{i=1}^n Y_i}{n \sum_{i=1}^n \left(\frac{1}{X_i}\right)^2 - \left(\sum_{i=1}^n \frac{1}{X_i}\right)^2} = \frac{17 \cdot 18,038 - 0,704 \cdot 473,51}{17 \cdot 0,03 - 0,704^2} = -1241,84$$

$$b_0 = \bar{Y} - b_1 \left(\overline{\frac{1}{X_i}} \right) = \frac{473,51}{17} - (-1241,84) \cdot \frac{0,704}{17} = 79,289$$

$$\hat{Y} = 79,287 - 1241,84 \frac{1}{X}.$$

Величина $b_1 = -1241,8 < 0$ говорит о том, с ростом X (стоимости ОПФ) имеем возрастающую функцию (рис. 1.3) с верхней асимптотой (при $X \rightarrow \infty$), т.е. с максимальным предельным уровнем Y (чистой прибыли) в размере $b_0 = 79,287$ млн. руб.

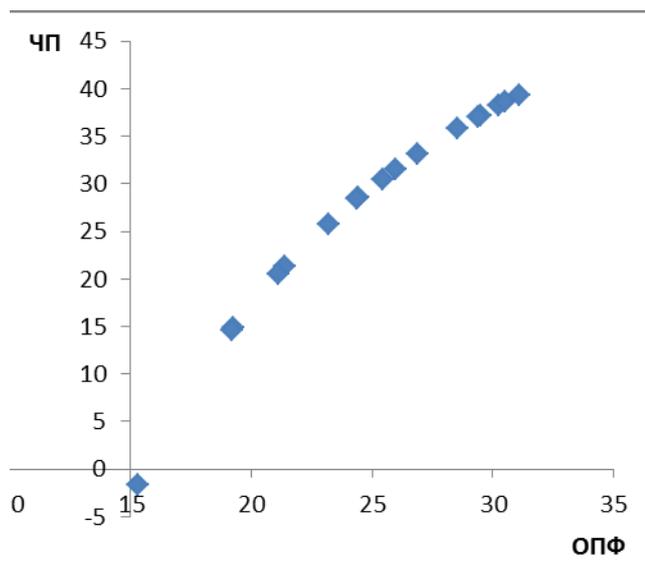


Рис. 1.3. Гиперболическая функция (расчетные значения Y)

Такие модели находят широкое применение как на микроуровне (например, при описании связи удельных расходов ресурсов с объемом выпускаемой продукции), так и на макроуровне (например, кривая Филлипса, характеризующая нелинейное соотношение между нормой безработицы X и процентом прироста заработной платы Y ; кривые Энгеля, моделирующие взаимосвязь доходов (или расходов) X и долей расходов на непродовольственные товары Y).

Таблица 1.4 – Расчетные данные

i	Y _i	X _{ii}	Y _i X _{ii}	X _{ii} ²	Y _i ²	lnY _i	lnX _{ii}	X _{ii} lnY _i	Y _i lnX _{ii}	lnY _i lnX _{ii}	(lnX _{ii}) ²	1/X _{ii}	Y _i (1/X _{ii})	(1/X _{ii}) ²
1	13,2	21,4	282,48	457,960	174,240	2,580	3,063	55,217	40,437	7,904	9,384	0,047	0,617	0,002
2	8,16	19,27	157,2432	371,333	66,586	2,099	2,959	40,452	24,142	6,211	8,753	0,052	0,423	0,003
3	15,14	19,21	290,8394	369,024	229,220	2,717	2,955	52,200	44,745	8,031	8,735	0,052	0,788	0,003
4	7,45	15,34	114,283	235,316	55,503	2,008	2,730	30,806	20,342	5,483	7,455	0,065	0,486	0,004
5	18,79	24,45	459,4155	597,803	353,064	2,933	3,197	71,720	60,065	9,377	10,218	0,041	0,769	0,002
6	21,52	23,21	499,4792	538,704	463,110	3,069	3,145	71,231	67,671	9,651	9,888	0,043	0,927	0,002
7	25,33	21,12	534,9696	446,054	641,609	3,232	3,050	68,260	77,262	9,858	9,304	0,047	1,199	0,002
8	26,78	24,4	653,432	595,360	717,168	3,288	3,195	80,219	85,551	10,503	10,205	0,041	1,098	0,002
9	28,72	25,44	730,6368	647,194	824,838	3,358	3,236	85,417	92,947	10,866	10,474	0,039	1,129	0,002
10	30,57	26,87	821,4159	721,997	934,525	3,420	3,291	91,896	100,606	11,255	10,831	0,037	1,138	0,001
11	33,1	25,98	859,938	674,960	1095,610	3,500	3,257	90,918	107,818	11,399	10,610	0,038	1,274	0,001
12	38,65	29,5	1140,175	870,250	1493,823	3,655	3,384	107,809	130,807	12,368	11,454	0,034	1,310	0,001
13	40,54	28,57	1158,228	816,245	1643,492	3,702	3,352	105,774	135,905	12,411	11,238	0,035	1,419	0,001
14	39,62	29,41	1165,224	864,948	1569,744	3,679	3,381	108,209	133,968	12,441	11,433	0,034	1,347	0,001
15	41,3	30,23	1248,499	913,853	1705,690	3,721	3,409	112,482	140,785	12,684	11,620	0,033	1,366	0,001
16	41,95	30,5	1279,475	930,250	1759,803	3,736	3,418	113,963	143,374	12,770	11,681	0,033	1,375	0,001
17	42,69	31,1	1327,659	967,210	1822,436	3,754	3,437	116,748	146,734	12,903	11,814	0,032	1,373	0,001
Сумма	473,51	426	12723,39	11018,46	15550,46	54,452	54,460	1403,321	1553,158	176,116	175,099	0,704	18,038	0,0304

Таблица 1.5 - Результаты оценивания моделей с использованием MicrosoftExcel(функции ЛИНЕЙН)

МОДЕЛИ	линейная		экспонента		линейно-логарифм.		степенная		показательная		гипербола	
b₀	-34,741		0,370		-155,626		0,005 e^{^(-5,291)}		1,447 e^{^(0,37)}		79,289	
b₁	2,498		0,113		57,274		2,651		1,12 e^{^(0,113)}		-1241,84	
коэффициенты (оценки)	2,498	-34,741	0,113	0,370	57,274	-155,626	2,651	-5,291	0,113	0,37	-1241,84	79,289
станд.ошибки оценок	0,206	5,248	0,011	0,291	5,486	17,60541	0,254	0,817	0,0114	0,291	146,174	6,185
R² s_e	0,907	3,820	0,867	0,212	0,879	4,364	0,879	0,202	0,867	0,212	0,828	5,205
F (n – 2)	146,861	15	97,804	15	109,011	15	108,599	15	97,804	15	72,176	15
сум.квадратов	2142,687	218,848	4,390	0,673	2075,891	285,644	4,449	0,615	4,390	0,673	1955,195	406,34
t-статистика	12,119	-6,62	9,89	1,27	10,441	-8,83983	10,421	-6,48	9,89	1,27	-8,496	12,82

Сравнение моделей

В табл.1.5 приведены результаты применения ППП Excel, функции ЛИНЕЙН для расчета оценок параметров моделей на основе данных табл. 1.3. Как видим, величина всех коэффициентов совпадает с их значениями, полученными на основе непосредственного применения формул МНК.

Кроме того, в табл. 1.5 для всех рассматриваемых функций приведены стандартные ошибки регрессии и коэффициенты детерминации, которые ЛИНЕЙН выводит в третью строку второго столбца и первого столбцов соответственно.

Итак, модели зависимости величины чистой прибыли от стоимости ОПФ имеют вид (см.табл.1.5):

линейная	$\hat{Y} = -34,741 + 2,498X$
показательная(экспонента)	$\hat{Y} = 1,447 \cdot 1,12^x$
линейно-логарифмическая	$\hat{Y} = -155,626 + 57,274 \ln X$
степенная	$\hat{Y} = 0,005 X^{2,651}$
гиперболическая	$\hat{Y} = 79,287 - \frac{1241,84}{X}$

Для выбора лучшей модели вначале следует привести характеристики в сопоставимый вид. Так как для оценивания некоторых моделей (показательной и степенной) выполнены преобразования зависимой переменной (логарифмирование), то сравнивать коэффициенты детерминации и стандартные ошибки регрессии для разных моделей напрямую нельзя (величина стандартных ошибок регрессии, вычисленная через логарифмы, оказалась существенно заниженной: сравните данные табл.1.5 и 1.7); значение коэффициента детерминации тоже должно быть скорректировано. Следует вначале записать модель в количественном виде (с использованием вычисленных оценок параметров) и найти теоретические (расчетные) значения исследуемого показателя с помощью этой модели, а затем определить требуемые характеристики.

Итак, выполним расчеты показателей качества моделей с использованием формул (см. табл. 1.6 – 1.7):

$$\text{стандартная ошибка регрессии (остатков)} s_e = \sqrt{\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n-2}},$$

$$\text{коэффициент детерминации} \quad R^2 = 1 - \frac{\sum (Y_i - \hat{Y})^2}{\sum (Y_i - \bar{Y})^2}$$

$$\text{скорректированный коэффициент детерминации} \quad \bar{R}^2 = 1 - \left[(1 - R^2) \frac{(n-1)}{(n-p-1)} \right],$$

$$\text{средняя ошибка аппроксимации} \quad \bar{A} = \frac{1}{n} \sum \left| \frac{Y_i - \hat{Y}}{Y_i} \right| \cdot 100\% .$$

Сравнение этих показателей (см. табл. 1.7) позволяет сделать вывод, что лучшими характеристиками по сравнению с другими моделями обладают линейная и степенная модели. Обе эти модели являются предпочтительными: линейная – в силу ее простоты, а степенная – из-за возможности использования показателя степени в качестве среднего коэффициента эластичности. Хотя средняя ошибка аппроксимации и стандартная ошибка остатков минимальны, а индекс детерминации принимает наибольшее значение для степенной функции, предпочтение можно отдать линейной модели, поскольку при сравнении коэффициента детерминации R^2 (рассчитанного для линейной регрессии) и индекса детерминации r^2 (рассчитанного для различных видов нелинейных регрессии) если окажется $|R^2 - r^2| < 0,1$, то линейная зависимость считается оправданной, и усложнять форму регрессионного уравнения не следует.

Из табл. 1.7 видим, что линейный коэффициент детерминации, равный 0,907, меньше только индекса детерминации степенной модели (0,909), но отличаются они друг от друга меньше, чем на 0,1 (что касается и других индексов детерминации).

Таблица 1.7 – Результаты расчетов характеристик регрессий

Модуль остатков					Модуль остатков / Y					Квадрат остатков				
линейная	показательная	линейно-логарифм.	степенная	гипербола	линейная	показательная	линейно-логарифм.	степенная	гипербола	линейная	показательная	линейно-логарифм	степенная	гипербола
5,514	3,070	6,626	3,769	8,059	0,418	0,233	0,502	0,286	0,611	30,405	9,426	43,904	14,207	64,954
5,234	4,628	5,661	4,691	6,685	0,641	0,567	0,694	0,575	0,819	27,390	21,417	32,051	22,005	44,691
1,896	2,439	1,497	2,395	0,496	0,125	0,161	0,099	0,158	0,033	3,596	5,947	2,242	5,735	0,246
3,873	0,750	6,692	0,431	9,115	0,520	0,101	0,898	0,058	1,224	15,002	0,563	44,782	0,186	83,085
7,543	4,180	8,667	5,370	9,708	0,401	0,222	0,461	0,286	0,517	56,893	17,474	75,120	28,833	94,252
1,715	1,555	2,956	0,475	4,265	0,080	0,072	0,137	0,022	0,198	2,942	2,417	8,739	0,225	18,189
7,315	9,567	6,258	8,943	4,840	0,289	0,378	0,247	0,353	0,191	53,514	91,524	39,166	79,979	23,425
0,572	3,939	0,560	2,751	1,614	0,021	0,147	0,021	0,103	0,060	0,327	15,518	0,314	7,569	2,606
0,086	3,029	1,011	1,879	1,755	0,003	0,105	0,035	0,065	0,061	0,007	9,176	1,021	3,531	3,080
1,808	0,371	2,293	0,459	2,503	0,059	0,012	0,075	0,015	0,082	3,268	0,137	5,256	0,211	6,264
2,945	5,792	2,167	4,722	1,610	0,089	0,175	0,065	0,143	0,049	8,676	33,544	4,694	22,296	2,594
0,297	2,008	0,439	1,097	1,457	0,008	0,052	0,011	0,028	0,038	0,088	4,032	0,193	1,203	2,123
3,916	3,940	4,164	4,030	4,717	0,097	0,097	0,103	0,099	0,116	15,334	15,526	17,337	16,240	22,252
0,898	0,626	1,584	0,194	2,556	0,023	0,016	0,040	0,005	0,065	0,806	0,392	2,509	0,038	6,532
0,529	2,856	1,689	1,108	3,090	0,013	0,069	0,041	0,027	0,075	0,280	8,158	2,853	1,228	9,550
0,505	3,575	1,830	1,470	3,377	0,012	0,085	0,044	0,035	0,080	0,255	12,781	3,348	2,160	11,402
0,254	6,031	1,454	3,032	3,331	0,006	0,141	0,034	0,071	0,078	0,064	36,369	2,114	9,190	11,097
Сумма					2,804	2,634	3,508	2,329	4,296	218,848	284,401	285,644	214,835	406,34
Средняя ошибка аппроксимации A=					16,496	15,495	20,634	13,698	25,270					
										Стандартные ошибки регрессии (остатков) s_e				
										3,820	4,354	4,364	3,784	5,205
										Коэффициент и индекс детерминации R^2				
										0,907	0,880	0,879	0,909	0,828
										Скорректированный коэффициент детерминации $\overline{R^2}$				
										0,901	0,872	0,871	0,903	0,816

На основе выбранных моделей (линейной и степенной) выполним **прогноз** величины чистой прибыли при увеличении средней стоимости ОПФ на 10%.

$$\bar{X} = 426 / 17 = 25,0588$$

$$X_0 = 25,0588 \cdot 1,1 = 27,565$$

Линейная модель

Точечный прогноз

$$\hat{Y}_0 = -34,741 + 2,498 \cdot 27,565 = 34,113 (\text{млн.руб.})$$

Интервальный прогноз:

$$\hat{Y}_0 - t_{1-\alpha;v} \cdot s_P \leq Y_0^* \leq \hat{Y}_0 + t_{1-\alpha;v} \cdot s_P$$

$$s_P = s_{\hat{Y}} \sqrt{1 + \frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\sum (X_i - \bar{X})^2}}$$

$$s_P = 3,81967 \sqrt{1 + \frac{1}{17} + \frac{(27,565 - 25,0588)^2}{343,4016}} = 3,81967 \cdot 1,0378 = 3,964$$

$$t_{1-\alpha;v} = t_{0,95;15} = 2,13$$

$$34,113 - 2,13 \cdot 3,964 \leq Y_0^* \leq 34,113 + 2,13 \cdot 3,964$$

$$25,669 \leq Y_0^* \leq 42,557$$

Таким образом, с вероятностью 0,95 величина чистой прибыли для некоторого предприятия, стоимость ОПФ которого составляет 27,565 млн. грн., должна находиться в пределах от 25,669 до 42,557 млн. грн.

Степенная модель

Точечный прогноз

$$\hat{Y} = 0,005 X^{2,651} = 0,005 \cdot 27,565^{2,651} = 33,203 (\text{млн.руб.})$$

Интервальный прогноз:

Интервальный прогноз сделаем для линеаризованной модели, а затем подвергнем концы интервала прогнозирования потенцированию.

Модель представим в виде:

$$\ln Y = \ln b_0 + b_1 \ln X = -5,291 + 2,651 \ln X.$$

Найдем вначале точечный прогноз по этой модели:

$$\ln Y = \hat{Y}_0 = -5,291 + 2,651 \cdot \ln(27,565) = 3,503.$$

Интервальный прогноз:

$$\hat{Y}_0 - t_{1-\alpha;v} \cdot s_p \leq \ln Y_0^* \leq \hat{Y}_0 + t_{1-\alpha;v} \cdot s_p$$

$$s_p = 0,202 \sqrt{1 + \frac{1}{17} + \frac{(27,565 - 25,0588)^2}{343,4016}} = 0,202 \cdot 1,0378 = 0,21$$

$$3,503 - 2,13 \cdot 0,21 \leq \ln Y_0^* \leq 3,503 + 2,13 \cdot 0,21$$

$$3,0552 \leq \ln Y_0^* \leq 3,950$$

$$e^{3,0552} \leq Y_0^* \leq e^{3,950}$$

$$21,226 \leq Y_0^* \leq 51,939$$

Таким образом, основываясь на прогнозе, сделанном по степенной модели, утверждаем, что с вероятностью 0,95 величина чистой прибыли для некоторого предприятия, стоимость ОПФ которого составляет 27,565 млн. грн., должна находиться в пределах от 21,226 до 51,939 млн. грн. Причем точечный прогноз по степенной модели (в отличие от линейной функции) не находится в середине интервала возможных значений величины чистой прибыли. Да и доверительный интервал оказался шире, чем при использовании линейной модели.

Пример 1.3 [19]. Для оценивания зависимости национального дохода (Y , ден. ед.) от денежной массы (X , ден. ед.) для некоторой гипотетической страны на основе данных за 10 лет получены следующие промежуточные результаты:

$$\bar{X} = 3,72; \quad \bar{Y} = 7,55; \quad \sum (Y_t - \bar{Y})^2 = 27,005;$$

$$\sum (X_t - \bar{X})^2 = 8,796; \quad \sum (X_t - \bar{X})(Y_t - \bar{Y}) = 15,09; \quad \sum e^2 = 1,117.$$

Используя эти данные, следует:

- 1) построить уравнение регрессии национального дохода на денежную массу; 2) построить 95%-й доверительный интервал для оцениваемых параметров;
- 3) проверить, можно ли отвергнуть гипотезы $\beta_1 = 0; \beta_1 = 1$.

Решение.

- 1) Найдем оценки параметров линейной регрессии:

$$b_1 = \frac{\sum (X_t - \bar{X})(Y_t - \bar{Y})}{\sum (X_t - \bar{X})^2} = \frac{15,09}{8,796} = 1,715;$$

$$b_0 = \bar{Y} - b_1 \bar{X} = 7,55 - 1,715 \cdot 3,72 = 1,168;$$

уравнение регрессии:

$$\hat{Y} = 1,168 + 1,715X.$$

Коэффициент регрессии позволяет считать, что при увеличении денежной массы X на одну ден. ед. национальный доход в среднем увеличится на 1,715 ден. ед.

- 2) Для построения доверительного интервала, с вероятностью 0,95 включающего неизвестный параметр β_1 , найдем стандартную ошибку регрессии и стандартную ошибку оценки b_1 :

$$s_e = \sqrt{\frac{\sum e^2}{n-2}} = \sqrt{\frac{1,117}{10-2}} = \sqrt{0,1397} = 0,3738;$$

Иногда удобно воспользоваться следующей формулой для вычисления остаточной суммы квадратов:

$$\sum e^2 = \sum Y^2 - b_0 \sum Y - b_1 \sum XY.$$

$$S_{b_1} = S_e \cdot \sqrt{\frac{1}{\sum (X_t - \bar{X})^2}} = 0,3738 \cdot \sqrt{\frac{1}{8,796}} = 0,126;$$

$$b_1 - t_{5\%,8} \cdot S_{b_1} \leq \beta_1 \leq b_1 + t_{5\%,8} \cdot S_{b_1}$$

$$t_{5\%,8} = 2,306;$$

$$1,715 - 2,306 \cdot 0,126 \leq \beta_1 \leq 1,715 + 2,306 \cdot 0,126$$

$$1,42 \leq \beta_1 \leq 2,01.$$

Таким образом, неизвестное истинное значение параметра β_1 с вероятностью 0,958 находится в интервале [1,42; 2,01].

3) Для проверки гипотез $\beta_1 = 0$; $\beta_1 = 1$ обратимся к построенному доверительному интервалу: так как значения 0 и 1 не принадлежат полученному интервалу, то обе гипотезы можно отвергнуть.

Тот же результат получим, рассчитав t -статистики:

$$t_1 = \frac{b_1 - 0}{S_{b_1}} = \frac{1,715 - 0}{0,126} = 13,61 > t_{5\%,8} = 2,306;$$

$$t_2 = \frac{b_1 - 1}{S_{b_1}} = \frac{1,715 - 1}{0,126} = 5,68 > t_{5\%,8} = 2,306;$$

Поскольку вычисленные t -статистики больше критического (табличного) значения, то обе гипотезы отвергаются, и коэффициент регрессии, количественно характеризующий влияние денежной массы на национальный доход, можно считать статистически значимой характеристикой неизвестного истинного значения параметра.

Вопросы и задания для самостоятельной работы

1. В чем состоит различие между теоретическим и эмпирическим уравнениями регрессии? Каковы основные причины наличия в регрессионной модели случайного отклонения?
2. В чем суть метода наименьших квадратов (МНК)? Приведите формулы расчета оценок параметров парного линейного уравнения регрессии по МНК?
3. Каков экономический смысл коэффициента регрессии?
4. Что представляет собой коэффициент детерминации? В каких пределах он может изменяться? Каков статистический смысл коэффициента детерминации?
5. Как соотносятся коэффициент корреляции и коэффициент детерминации для парной линейной регрессии?
6. На что направлена проверка статистической надежности (значимости) результатов оценивания модели парной линейной регрессии.
7. Представьте алгоритм проверки значимости коэффициентов.
8. Представьте алгоритм проверки значимости уравнения регрессии в целом.
9. Как осуществляется проверка значимости коэффициента парной корреляции
10. Дайте определение интервальных оценок коэффициентов.
11. Сформулируйте определение нелинейной регрессии. Перечислите основные типы. Чем следует руководствоваться при выборе типа модели?
12. Как построить точечный и интервальный прогнозы. Каково предназначение интервального прогноза?

Задание 1.1[19]. Кривая Филипса описывает связь темпа роста зарплаты и уровня безработицы:

$$\delta w_t = \beta_1 + \beta_2 \frac{1}{u_t} + \varepsilon_t, \quad (\beta_1 < 0 \text{ и } \beta_2 > 0)$$

где w_t – уровень заработной платы,

$$\delta w_t = \frac{100(w_t - w_{t-1})}{w_{t-1}} \text{ – темп роста зарплаты (в процентах),}$$

u_t – процент безработных в год t .

Используя данные табл.1.8, выполните следующие задания:

- 1) определите оценки параметров модели и проверьте наличие значимой связи между δw и u ;
- 2) найдите естественный уровень безработицы, т.е. такой уровень безработицы, при котором $\delta w = 0$;
- 3) определите, когда изменения в уровне безработицы оказывали наибольшее (наименьшее) влияние на темп изменения зарплаты;
- 4) Найдите 95%-е интервалы для β_1 и β_2 .

Таблица 1.8 – Изменение уровней заработной платы и безработицы в некоторой стране за 18 лет

Год t	w_t	u_t	Год t	w_t	u_t
1	1,62	1,0	10	2,66	1,8
2	1,65	1,4	11	2,73	1,9
3	1,79	1,1	12	2,80	1,5
4	1,94	1,5	13	2,92	1,4
5	2,03	1,5	14	3,02	1,8
6	2,12	1,2	15	3,13	1,1
7	2,26	1,0	16	3,28	1,5
8	2,44	1,1	17	3,43	1,3
9	2,57	1,3	18	3,58	1,4

Задание 1.2. По данным 15-летних наблюдений построена следующая эконометрическая модель зависимости ВВП – валового внутреннего продукта (млрд. руб.) от M_1 – денежной массы (млрд. руб.):

$$ВВП_t = -787.4723 + 8.0863M_{1t} + e_t$$

$$s_j = (\quad) \quad (0.2197)$$

$$t = (-10.0) \quad (\quad)$$

$$R^2 = 0.9912, \quad s_e = 2.58$$

Задание.

- 1) Найдите неизвестные значения и оцените статистическую значимость коэффициентов регрессии и общее качество уравнения.
- 2) По утверждениям монетаристов, денежная масса имеет существенное положительное влияние на ВВП. Подтверждается ли это моделью?

3) Предложение денег в следующем году планируется на уровне 550 млрд. руб. Каково прогнозное значение ВВП на данный год? В каком интервале будет находиться прогнозируемое значение ВВП с вероятностью 0,95?

Задание 1.3[12]. По 19 предприятиям оптовой торговли изучается зависимость объема реализации (Y) от размера торговой площади (X_1) и товарных запасов (X_2). Были получены следующие варианты уравнений регрессии:

1) $Y = 25 + 15 X_1$ $R^2 = 0,90$;

2) $Y = 42 + 27 X_2$ $R^2 = 0,84$;

3) $Y = 25 + 15 X_1 + 8 X_2$ $R^2 = 0,92$;

(2,5) (4,0)

4) $Y = 21 + 14 X_1 + 20 X_2 + 0,6 X_2^2$ $R^2 = 0,95$.

(5,0) (12) (0,2)

В скобках указаны значения стандартных ошибок для коэффициентов регрессии. Проанализируйте тесноту связи результата с каждым из факторов. Выберите лучшее из уравнений 3 и 4, обоснуйте принятое решение (используйте скорректированный коэффициент детерминации).

ТЕМА 2. ПОСТРОЕНИЕ И АНАЛИЗ МНОЖЕСТВЕННОЙ РЕГРЕССИОННОЙ МОДЕЛИ

Данная тема знакомит студентов с понятиями линейной модели множественной регрессии. Для оценки неизвестных параметров уравнения регрессии чаще всего используется метод наименьших квадратов (МНК), позволяющий получить несмещенные, состоятельные и эффективные оценки параметров уравнения. Особое внимание при изучении данной темы необходимо

также уделить статистическим свойствам оценок параметров регрессионной модели.

Для выполнения лабораторной работы № 2 студенты должны

знать:

- особенности модели множественной регрессии;
- применение МНК для оценивания параметров многофакторной модели.

уметь:

- получать оценки регрессионных моделей.
- применять статистические пакеты для решения практических задач.

Задание для выполнения лабораторной работы №2

На базе данных, характеризующих социально-демографическую ситуацию в 52-х странах и представленным в табл. 2.1, изучается линейная (в среднем) зависимость результативного признака Y от пяти факторных признаков – регрессоров X_1, X_2, X_3, X_4, X_5 .

Требуется в соответствии с выбранным вариантом (табл. 2.2) выполнить множественный линейный регрессионный анализ, опираясь на результаты применения встроенных функций MicrosoftExcel.

1) Записать модель множественной линейной регрессии показателя Y .

2) Построить матрицу парных коэффициентов корреляции между признаками и сделать вывод о силе линейной связи результативного признака (зависимой переменной) с каждым из факторов, а также между факторами.

3) Вычислить оценки параметров модели и их стандартные ошибки (s_j); записать уравнение регрессии и его стандартную ошибку ($s_{\hat{y}}$); используя «Остатки», вычислить среднюю относительную ошибку аппроксимации A .

4) Предположив выполнение условий линейного регрессионного анализа,

- а) оценить статистическую значимость уравнения регрессии;
- б) оценить статистическую значимость коэффициентов уравнения регрессии.

5) При наличии в уравнении регрессии хотя бы одного незначимого коэффициента исключить тот регрессор, при котором коэффициент незначим, а соответствующая этому коэффициенту величина Р-значения является наибольшей (или, иначе, значение соответствующей t-статистики является наименьшим).

6) Выполнить п.п. 3 – 5 с оставшимися регрессорами. Процедуру пошагового исключения регрессоров продолжать до тех пор, пока не будет получено значимое уравнение регрессии со значимыми коэффициентами.

Замечание. Если после исключения регрессора уравнение становится статистически незначимым или остается значимым, но его нормированный R-квадрат значительно уменьшается, то этот регрессор следует «возвратить» в уравнение и исключить очередной регрессор, коэффициент при котором незначим (конечно, при наличии такого регрессора).

7) Выбрать лучшее уравнение и, используя его, сделать прогноз (точечный и интервальный) исследуемого показателя.

Обозначения признаков

X_1 – численность населения (в тыс. чел.);

X_2 – рождаемость (на 1000 чел.);

X_3 – смертность (на 1000 чел.);

X_4 – младенческая смертность – число детей, умерших в возрасте до 1 года (на 1000 чел.);

X_5 – среднее число детей в семье;

X_6 – ожидаемая продолжительность жизни мужчины (в годах);

X_7 – ожидаемая продолжительность жизни женщины (в годах);

X_8 – ВВП на душу населения (в долл. США по покупательной способности валют);

X_9 – плотность населения (количество человек на кв. км);

X_{10} – процент городского населения;

X_{11} – процент грамотных;

X_{12} – прирост населения (% в год).

Таблица 2.1 - Значения признаков [5].

№п/п	Страна	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁	X ₁₂
1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	Австралия	17800	15	8	7,30	1,90	74	80	16848	2,3	85	100	1,38
2	Австрия	8000	12	11	6,70	1,50	73	79	18396	94,0	58	99	0,20
3	Аргентина	33900	20	9	25,60	2,80	68	75	3408	120	86	95	1,30
4	Бангладеш	125000	35	11	106,00	4,70	53	53	202	800,0	16	35	2,40
5	Беларусь	10300	13	11	19,00	1,88	66	76	6500	50,0	65	99	0,32
6	Бельгия	10100	12	11	7,20	1,70	73	79	17912	329,0	96	99	0,20
7	Бразилия	156600	21	9	66,00	2,70	57	67	2354	180	75	81	1,28
8	Буркина-Фасо	10000	47	18	118,00	6,94	47	50	357	360	15	18	2,81
9	Великобритания	58400	13	11	7,20	1,83	74	80	15974	237,0	89	99	0,20
10	Вьетнам	73100	27	8	46,00	3,33	63	68	230	218,0	20	88	1,78
11	Гаити	6500	40	19	109,00	5,94	43	47	383	231,0	29	53	1,63
12	Германия	81200	11	11	6,50	1,47	73	79	17539	227,0	85	99	0,36
13	Гондурас	5600	35	6	45,00	4,90	65	70	1030	460	44	73	2,73
14	Гонконг	5800	13	6	5,80	1,40	75	80	14641	5494,0	94	77	0,09
15	Египет	60000	29	9	76,40	3,77	60	63	748	57,0	44	48	1,95
16	Замбия	9100	46	18	85,00	6,68	44	45	573	11,0	42	73	2,80
17	Индия	911600	29	10	79,00	4,48	58	59	275	283,0	26	52	1,90
18	Ирландия	3600	14	9	7,40	1,99	73	78	12170	51,0	57	98	0,30
19	Испания	39200	11	9	6,90	1,40	74	81	13047	77,0	78	95	0,25
20	Италия	58100	11	10	7,60	1,30	74	81	17500	188,0	69	97	0,21
21	Канада	29100	14	8	6,80	1,80	74	81	19904	2,8	77	97	0,70
22	Китай	1205200	21	7	52,00	1,84	67	69	377	124,0	26	78	1,10
23	Колумбия	35600	24	6	28,00	2,47	69	75	1538	31,0	70	87	2,00
24	Коста-Рика	3300	26	4	11,00	3,10	76	79	2031	64,0	47	93	2,30
25	Куба	11100	17	7	10,20	1,90	74	78	1382	99,0	74	94	0,95
26	Малайзия	19500	29	5	25,60	3,51	66	72	2995	58,0	43	78	2,30
27	Марокко	28600	29	6	50,00	3,83	66	70	1062	63,0	46	50	2,12
28	Мексика	91800	28	5	35,00	3,20	69	77	3604	46,0	73	87	1,90
29	Нидерланды	15400	13	9	6,30	1,58	75	81	17245	366,0	89	99	0,58
30	Новая Зеландия	3524	16	8	8,90	2,03	73	80	14381	13,0	84	99	0,57
31	Норвегия	4300	13	10	6,30	2,00	74	81	17755	11,0	75	99	0,40
32	ОАЭ	2800	28	3	22,00	4,50	70	74	14193	32,0	81	68	4,80
33	Польша	38600	14	10	13,80	1,94	69	77	4429	123,0	62	99	0,30
34	Португалия	10500	12	10	9,20	1,50	71	78	9000	108,0	34	85	0,36
35	Россия	149200	13	11	27,00	1,83	64	74	6680	8,8	74	99	0,20
36	Саудовская Аравия	18000	38	6	52,00	6,67	66	70	6651	7,7	77	62	3,20
37	Северная Корея	23100	24	6	27,70	2,40	67	73	1000	189,0	60	99	1,83
38	Сингапур	2900	16	6	5,70	1,88	73	79	14990	4456,0	100	88	1,20
39	США	260800	15	9	8,11	2,06	73	79	23474	26,0	75	97	0,99
40	Таиланд	59400	19	6	37,00	2,10	65	72	1800	115,0	22	93	1,40
41	Турция	62200	26	6	49,00	3,21	69	73	3721	79,0	61	81	2,02
42	Украина	51800	12	13	20,70	1,82	65	75	2340	87,0	67	97	0,05
43	Филиппины	69800	27	7	51,00	3,35	63	68	867	221,0	43	90	1,92
44	Финляндия	5100	13	10	5,30	1,80	72	80	15877	39,0	60	100	0,30
45	Франция	58000	13	9	6,70	1,80	74	82	18944	105,0	73	99	0,47
46	Чили	14000	23	6	14,60	2,50	71	78	2591	18,0	85	93	1,70
47	Швейцария	7000	12	9	6,20	1,60	75	82	22384	170,0	62	99	0,70
48	Швеция	8800	14	11	5,70	2,10	75	81	16900	19,0	84	99	0,52
49	Эфиопия	55200	45	14	110,00	6,81	51	54	122	47,0	12	24	3,10
50	ЮАР	43900	34	8	47,10	4,37	62	68	3128	35,0	49	76	2,60
51	Южная Корея	45000	16	6	21,70	1,65	68	74	6627	447,0	72	96	1,00
52	Япония	125500	11	7	4,40	1,55	76	82	19860	330,0	77	99	0,30

Таблица 2.2 – Выбор варианта для выполнения лабораторной работы

Вариант	Номер результативного признака	Номера факторных признаков
1	6	1, 2, 3, 4, 8
2	6	1, 2, 3, 8, 9
3	6	1, 2, 3, 8, 10
4	6	1, 2, 3, 5, 10
5	6	1, 2, 3, 8, 11
6	6	1, 2, 5, 8, 11
7	6	1, 2, 5, 8, 9
8	6	1, 2, 4, 8, 11
9	6	1, 2, 5, 8, 11
7	7	1, 2, 3, 4, 8
11	7	1, 2, 3, 8, 9
12	7	1, 2, 3, 8, 10
13	7	1, 2, 3, 5, 10
14	7	1, 2, 3, 8, 11
15	7	1, 2, 5, 8, 11
16	7	1, 2, 5, 8, 9
17	7	1, 2, 4, 8, 11
18	7	1, 2, 5, 8, 11
19	8	1, 2, 3, 4, 5
20	8	1, 2, 3, 5, 9
21	8	1, 2, 3, 5, 10
22	8	1, 2, 3, 4, 11
23	8	1, 2, 3, 4, 12
24	8	1, 2, 3, 5, 11
25	8	1, 2, 3, 9, 11

Для выполнения работы следует изучить теоретический материал, представленный в теме 3 Учебного пособия по эконометрике [18].

РЕШЕНИЕ ТИПОВЫХ ЗАДАЧ

Пример 2.1. В табл. 2.3 представлена информация о 20 торговых фирмах по пяти показателям, где

y – количество посещений в месяц (тыс. чел.);

x_1 – расходы на рекламу (тыс. руб.);

x_2 – торговые площади (m^2);

x_3 – число потенциальных покупателей (тыс. чел.);

x_4 – число конкурирующих магазинов.

Таблица 2.3 – Исходная информация [12]

i	Y _i	X _{1i}	X _{2i}	X _{3i}	X _{4i}	i	Y _i	X _{1i}	X _{2i}	X _{3i}	X _{4i}
1	9,5	7,5	891	4,362	11	11	9,4	6	615	3,78	8
2	15	6,1	250	2,947	5	12	12	7	700	5,181	6
3	7,7	10	262	3,298	8	13	5,82	4	288	1,25	12
4	12	7	100	7,568	9	14	11,84	3,2	350	2,005	10
5	3,87	2,6	100	3,824	12	15	8,05	6,5	100	4,951	9
6	8	3,5	524	1,376	13	16	9	9	120	4,193	11
7	14,14	5	450	5,578	10	17	9	6,5	132	4,805	8
8	2,2	0,6	100	0,901	14	18	10	4	316	3,333	12
9	8,1	1,25	550	1,622	7	19	12	5	245	2,4	2
10	8,6	2,65	204	4,355	6	20	17	9,5	615	6,556	5

Требуется построить модель множественной линейной регрессии и выполнить анализ посещаемости торговых предприятий.

Решение

Рассчитаем для рассматриваемых переменных матрицу парных коэффициентов корреляции (**Excel / Анализ данных / Корреляция**) (табл. 2.4).

Таблица 2.4 - Матрица парных коэффициентов корреляции

	Y	X ₁	X ₂	X ₃	X ₄
Y	1,0000	0,4979	0,3568	0,5290	-0,6254
X ₁	0,4979	1,0000	0,1703	0,6145	-0,3252
X ₂	0,3568	0,1703	1,0000	0,0467	-0,1283
X ₃	0,5290	0,6145	0,0467	1,0000	-0,2929
X ₄	-0,6254	-0,3252	-0,1283	-0,2929	1,0000

Проверим статистическую значимость коэффициентов парной корреляции. Для этого рассчитаем величину t-статистик (табл. 2.5) и сравним их с критическим значением, которое при уровне значимости 5% и 18-ти степенях свободы равняется 2,101.

$$t_{y,x1} = \frac{r_{x1} \sqrt{n-2}}{\sqrt{1-r_{x1}^2}} = \frac{0,4979\sqrt{20-2}}{\sqrt{1-0,4979^2}} = 2,4356 > t_{0,05;15} = 2,101$$

$$t_{x1,x2} = \frac{r_{x1,x2} \sqrt{n-2}}{\sqrt{1-r_{x1,x2}^2}} = \frac{0,1703\sqrt{20-2}}{\sqrt{1-0,1703^2}} = 0,7332 < t_{0,05;15} = 2,101$$

и т.д.

Таблица 2.5 – Величина t –статистик для парных корреляций

	Y	X ₁	X ₂	X ₃	X ₄
Y		2,4356	1,6203	2,6448	-3,4003
X ₁	2,4356		0,7332	3,3049	-1,4591
X ₂	1,6203	0,7332		0,1983	-0,5489
X ₃	2,6448	3,3049	0,1983		-1,2995
X ₄	-3,4003	-1,4591	-0,5489	-1,2995	

Анализ полученных результатов показывает, что значимые влияние на Y оказывают три переменные: X₁, X₃ и X₄ (для них вычисленные значения t – статистик при уровне значимости 0,05 больше табличного);

Выполним оценивание параметров регрессионной модели (рис. 2.1).

<i>Регрессионная статистика</i>	
Множественный R	0,7777
R-квадрат	0,6047
Нормированный R-квадрат	0,4993
Стандартная ошибка	2,5220
Наблюдения	20

<i>Дисперсионный анализ</i>					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>
Регрессия	4	145,9783	36,49458	5,737523	0,005245069
Остаток	15	95,41027	6,360685		
Итого	19	241,3886			

	<i>Коэффициенты</i>	<i>Стандартная ошибка</i>	<i>t-статистика</i>	<i>P-Значение</i>
Y-пересечение	9,9128	2,8162	3,5200	0,0031
X ₁	0,1477	0,2873	0,5142	0,6146
X ₂	0,0040	0,0025	1,5917	0,1323
X ₃	0,6202	0,4116	1,5068	0,1526
X ₄	-0,5319	0,1990	-2,6722	0,0174

Рис. 2.1. Отчет Excel об уравнении регрессии

Получили модель:

$$\hat{Y} = 9,9128 + 0,1477X_1 + 0,0040X_2 + 0,6202X_3 - 0,5319X_4$$

Дисперсионный анализ регрессии показывает, что уравнение регрессии значимо (p -величина статистики Фишера равняется 0,0052, что меньше 0,05).

Коэффициенты множественной корреляции ($R = 0,7777$) и множественной детерминации ($R^2 = 0,6047$) показывают, что уравнение регрессии довольно

точно описывает зависимость Y от остальных переменных (на 60,47% изменение Y обусловлено изменением переменных, включенных в модель).

Различия в исходном и скорректированном нормированном) коэффициентах множественной детерминации ($\overline{R^2} = 0,499$) говорит о том, что в уравнении регрессии есть незначимые переменные (что совпадает с результатом корреляционного анализа).

На это же указывают расчетные уровни значимости оценок коэффициентов уравнения регрессии (b_j). Среди них только один (при X_4) имеет p -значение, меньшее 0,05 ($p = 0,0174$). Поэтому, если судить формально на основе этих показателей, то на количество посещений в месяц магазинов фирмы значимо влияет только число конкурирующих магазинов.

Противоречивый вывод о статистической надежности результатов оценивания (значимость уравнения в целом и незначимость отдельных коэффициентов регрессии) свидетельствует о возможности мультиколлинеарности факторов (исследованию мультиколлинеарности посвящена Лабораторная работа №3).

Коэффициенты в натуральном масштабе (b_j) в данном случае сравнивать нельзя, т. к. они имеют разные единицы измерения. Для сравнения степени влияния факторов на величину исследуемого показателя используем частные коэффициенты корреляции, бета-коэффициенты (β_j). Для выполнения этих построений удобно иметь результаты применения инструмент анализа данных **Описательная статистика**. Для его активизации надо выполнить следующие шаги:

1) в главном меню выбрать последовательно пункты **Данные / Анализ данных / Описательная статистика**, после чего щелкнуть по кнопке **ОК**;

2) заполнить диалоговое окно ввода данных и параметров вывода (рис. 2.2):

Входной интервал - диапазон, содержащий анализируемые данные, это может быть одна или несколько строк (столбцов);

Группирование - по столбцам или по строкам - необходимо указать дополнительно;

Метки - флажок, который указывает, содержит ли первая строка названия столбцов или нет;

Выходной интервал - достаточно указать левую верхнюю ячейку будущего диапазона;

Если необходимо получить дополнительную информацию *Итоговой статистики, Уровня надежности, k-го наибольшего и наименьшего значений*, установите соответствующие флажки в диалоговом окне. Кнопка **ОК**.

В табл. 2.6 помимо выведенных программой показателей приведены коэффициенты вариации, рассчитанные по формулам:

$$V_y = \frac{\sigma_y}{\bar{Y}}, \quad V_{x_j} = \frac{\sigma_{x_j}}{\bar{X}_j} \quad j = \overline{1, p}$$

Из табл. 2.6 видно, что соотношение стандартных отклонений и коэффициентов вариации существенно различаются. Так, если $S_{x1} = 2,636$ существенно меньше, чем $S_{x2} = 234,075$ (почти в 100 раз), то $V_{x1} = 49,3\%$ меньше $V_{x2} = 67,7\%$ не на много. Это объясняется тем, что исследуемые показатели имеют разные единиц измерения (тыс. чел. и тыс. руб.). Поэтому для того, чтобы составить реальное представление о сравнительной мере рассеяния вычисляют коэффициент вариации. Для одинаковых единиц измерения эти два показателя дают сходную информацию (сравните X_1 и X_3).

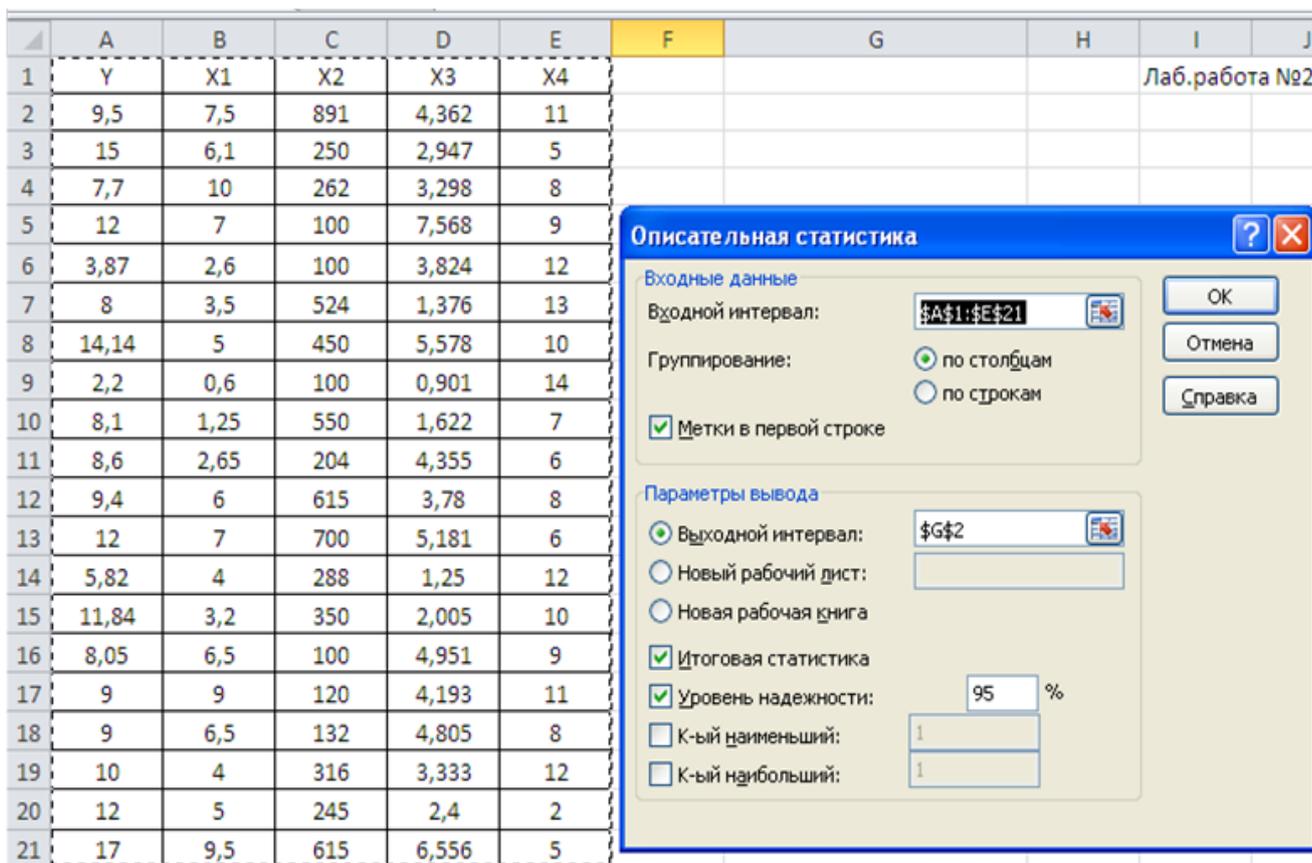


Рис. 2.2 - Диалоговое окно ввода параметров инструмента **Описательная статистика**

Результаты вычисления соответствующих показателей для каждого признака представлены на рис. 2.3 и в табл. 2.6.

Таблица 2.6 - Описательные статистики

Показатели	Y	X ₁	X ₂	X ₃	X ₄
Среднее \bar{Y}, \bar{X}_j	9,661	5,345	345,6	3,7148	8,9
Стандартное отклонение S_y, S_{xj}	3,564	2,636	234,075	1,803	3,1104
Дисперсия выборки D_y, D_{xj}	12,705	6,949	54791	3,24932	9,674
Коэффициент вариации V_y, V_{xj}	36,894	49,318	67,73	48,531	34,947
Количество наблюдений n	20	20	20	20	20

Определим **β -коэффициенты** по формуле (2.17) с использованием данных табл. 2.6 и рис. 2.3:

$$\beta_1 = b_1 \frac{S_{x1}}{S_y} = 0,1447 \frac{2,63603}{3,56436} = 0,109 \quad \beta_2 = b_2 \frac{S_{x2}}{S_y} = 0,004 \frac{2,63603}{3,56436} = 0,264$$

$$\beta_3 = b_3 \frac{S_{x3}}{S_y} = 0,6202 \frac{1,80255}{3,56436} = 0,314 \quad \beta_4 = b_4 \frac{S_{x4}}{S_y} = -0,5319 \frac{3,11025}{3,56436} = -0,464$$

G23					f_x						
	A	B	C	D	E	F	G	H	I	J	K
1	Y	X1	X2	X3	X4						
2	9,5	7,5	891	4,362	11						
3	15	6,1	250	2,947	5						
4	7,7	10	262	3,298	8						
5	12	7	100	7,568	9						
6	3,87	2,6	100	3,824	12		Описательная статистика				
7	8	3,5	524	1,376	13		Y	X1	X2	X3	X4
8	14,14	5	450	5,578	10						
9	2,2	0,6	100	0,901	14	Среднее	9,661	5,345	345,600	3,714	8,900
10	8,1	1,25	550	1,622	7	Стандартная ошибка	0,797	0,589	52,341	0,403	0,695
11	8,6	2,65	204	4,355	6	Медиана	9,200	5,500	275,000	3,802	9,000
12	9,4	6	615	3,78	8	Мода	12,000	7,000	100,000	#Н/Д	8,000
13	12	7	700	5,181	6	Стандартное отклонение	3,564	2,636	234,075	1,803	3,110
14	5,82	4	288	1,25	12	Дисперсия выборки	12,705	6,949	54790,989	3,249	9,674
15	11,84	3,2	350	2,005	10	Экссесс	0,357	-0,661	-0,229	-0,343	-0,325
16	8,05	6,5	100	4,951	9	Асимметричность	-0,003	0,024	0,799	0,266	-0,375
17	9	9	120	4,193	11	Интервал	14,800	9,400	791,000	6,667	12,000
18	9	6,5	132	4,805	8	Минимум	2,200	0,600	100,000	0,901	2,000
19	10	4	316	3,333	12	Максимум	17,000	10,000	891,000	7,568	14,000
20	12	5	245	2,4	2	Сумма	193,220	106,900	6912,000	74,285	178,000
21	17	9,5	615	6,556	5	Счет	20,000	20,000	20,000	20,000	20,000

Рис.2.3. Результат применения инструмента **Описательная статистика**

β -коэффициент показывает, на какую величину изменится среднеквадратическое отклонение (СКО) результативного признака, если СКО конкретного факторного признака изменится на 1 единицу, т.е. при увеличении на 1 единицу СКО расходов на рекламу (X_1) СКО количества посещений в месяц (Y) увеличится на 0,109 ($\beta_1 = 0,109$); при увеличении на 1 единицу СКО числа конкурирующих магазинов (X_4) СКО количества посещений уменьшится на 0,464 единицы ($\beta_4 = -0,464$).

Таким образом, судя по β -коэффициентам, наименьшее влияние на количество посещений в месяц магазинов фирмы имеет X_1 – расходы на рекламу, далее следует показатель, характеризующий величину торговых площадей X_2 , за ним – X_3 – число потенциальных покупателей, а наибольшее влияние на количество посещений в месяц магазинов фирмы оказывает X_4 – число конкурирующих магазинов.

По β -коэффициентам можно ранжировать факторы по силе влияния на Y , но количественно сравнить силу этого влияния можно используя коэффициенты эластичности модели

Коэффициенты эластичности показывают: на сколько процентов изменится среднее значение зависимой переменной Y при увеличении

факторного признака X_j на 1%. Вычислим коэффициенты эластичности для каждого фактора:

$$E_{x_1} = b_1 \frac{\bar{X}_1}{\bar{Y}} = 0,1447 \frac{5,345}{9,661} = 0,08 \quad E_{x_2} = b_2 \frac{\bar{X}_2}{\bar{Y}} = 0,004 \frac{345,6}{9,661} = 0,143$$

$$E_{x_3} = b_3 \frac{\bar{X}_3}{\bar{Y}} = 0,6202 \frac{3,71425}{9,661} = 0,239 \quad E_{x_4} = b_4 \frac{\bar{X}_4}{\bar{Y}} = -0,5319 \frac{8,9}{9,661} = -0,49$$

Таким образом, увеличение расходов на рекламу (X_1) на 1% при прочих равных условиях вызывает рост количества посещений магазина (Y) на 0,08%, т.е. практически не оказывает влияния. Чуть большему увеличению притока покупателей (на 0,143%) способствует рост торговой площади (X_2) на 1%. Приток покупателей уменьшится на 0,49% при увеличении на 1% числа конкурирующих магазинов.

Рассчитаем *частные коэффициенты корреляции*. Для этого найдем матрицу, обратную матрице парных корреляций (функция МОБР среди Математических) (см. табл. 2.7).

Для удобства представления информации в формулах припишем первой строке и первому столбцу в этой матрице нулевые индексы, а индексы остальных строк и столбцов соответствуют индексам объясняющих переменных.

Таблица 2.7 – Обратная матрица корреляций $C = R^{-1}$

Переменные		Y	X ₁	X ₂	X ₃	X ₄
	Индексы	0	1	2	3	4
Y	0	2,530006	-0,27637	-0,66786	-0,79348	1,174281
X ₁	1	-0,27637	1,743108	-0,14197	-0,88412	0,116906
X ₂	2	-0,66786	-0,14197	1,220098	0,321138	-0,21325
X ₃	3	-0,79348	-0,88412	0,321138	1,892974	-0,18819
X ₄	4	1,174281	0,116906	-0,21325	-0,18819	1,689924

Определение частных коэффициентов корреляции:

$$r_{jk} = \frac{-c_{jk}}{\sqrt{c_{jj} \cdot c_{kk}}}$$

где c_{jk} – элемент матрицы $C = R^{-1}$, расположенный в j -й строке и k -м столбце;

c_{jj} и c_{kk} - диагональные элементы матрицы C .

Проверка статистической значимости этих коэффициентов выполняется так же, как и парных коэффициентов корреляции – на основе t -статистики:

$$t_{jk} = \frac{r_{jk} \sqrt{n-p-1}}{\sqrt{1-r_{jk}^2}}.$$

$$r_{01} = \frac{-c_{01}}{\sqrt{c_{00} \cdot c_{11}}} = \frac{-(-0,27637)}{\sqrt{2,530006 \cdot 1,743108}} = 0,1316, \quad t_{01} = \frac{r_{01} \sqrt{20-4-1}}{\sqrt{1-r_{01}^2}} = \frac{0,1316 \cdot \sqrt{15}}{\sqrt{1-0,1316^2}} = 0,514$$

$$r_{02} = \frac{-c_{02}}{\sqrt{c_{00} \cdot c_{22}}} = \frac{-(-0,66786)}{\sqrt{2,530006 \cdot 1,220098}} = 0,3801, \quad t_{02} = \frac{r_{02} \sqrt{20-4-1}}{\sqrt{1-r_{02}^2}} = \frac{0,3801 \cdot \sqrt{15}}{\sqrt{1-0,3801^2}} = 1,592$$

$$r_{03} = \frac{-c_{03}}{\sqrt{c_{00} \cdot c_{33}}} = \frac{-(-0,79348)}{\sqrt{2,530006 \cdot 1,892974}} = 0,3626, \quad t_{03} = \frac{r_{03} \sqrt{20-4-1}}{\sqrt{1-r_{03}^2}} = \frac{0,3626 \cdot \sqrt{15}}{\sqrt{1-0,3626^2}} = 1,507$$

$$r_{04} = \frac{-c_{04}}{\sqrt{c_{00} \cdot c_{44}}} = \frac{-1,174281}{\sqrt{2,530006 \cdot 1,689924}} = -0,5679, \quad t_{04} = \frac{r_{04} \sqrt{20-4-1}}{\sqrt{1-r_{04}^2}} = \frac{-0,5679 \cdot \sqrt{15}}{\sqrt{1-(-0,5679)^2}} = -2,672$$

Расчеты показали, что частные (очищенные от влияния других факторов) коэффициенты корреляции меньше соответствующих парных коэффициентов. Сравнение их друг с другом позволяет ранжировать факторы по тесноте их связи с результирующим показателем. В целом, они подтвердили ранжирование факторов, выполненное с помощью β -коэффициентов и коэффициентов эластичности.

Табличное значение $t_{0,05;15} = 2,13$. Поскольку $|t_{04}| > 2,13$, то с вероятностью 0,95 делаем вывод о значимом влиянии на Y – количество посещений магазина в месяц – только фактора X_4 – число конкурирующих магазинов.

Рассчитаем Δ -коэффициенты. Из табл. 2.7 найдем коэффициенты парной корреляции:

$$r_{YX_1} = 0,4979 \quad r_{YX_2} = 0,3568 \quad r_{YX_3} = 0,529 \quad r_{YX_4} = -0,6254$$

Тогда

$$\Delta_1 = r_{YX_1} \frac{\beta_1}{R^2} = 0,4979 \frac{0,109}{0,6047} = 0,09, \quad \Delta_2 = r_{YX_2} \frac{\beta_2}{R^2} = 0,3568 \frac{0,264}{0,6047} = 0,156$$

$$\Delta_3 = r_{YX_3} \frac{\beta_3}{R^2} = 0,529 \frac{0,314}{0,6047} = 0,274, \quad \Delta_{41} = r_{YX_4} \frac{\beta_4}{R^2} = -0,6254 \frac{-0,464}{0,6047} = 0,480$$

Δ -коэффициент показывает удельный вес влияния конкретного факторного признака в совместном влиянии всех факторных признаков на результативный показатель. Это означает, что на 9% рост количества посещений в месяц можно объяснить увеличением расходов на рекламу, на 15,6% – расширением торговых площадей, на 27,4% – приростом числа потенциальных покупателей и на 48% – уменьшением числа конкурирующих магазинов.

Проведем теперь пошаговый регрессионный анализ, используя метод исключения. Вначале строится модель, включающая все переменные. Затем для каждой переменной вычисляют частную F -статистику, и ту переменную, для которой F -статистика минимальна, исключают из рассмотрения. Затем строится новая модель с оставшимися переменным, и, после вычисления частных F -статистик, вновь удаляют одну из переменных и т.д. Так поступают до тех пор, пока значения всех F -статистик не станут больше заданной величины.

Выполним построение моделей линейной регрессии методом наименьших квадратов (удобно в данном случае воспользоваться функцией ЛИНЕЙН MSExcel). Результаты приведены в табл. 2.8.

Таблица 2.8 – Результаты построения регрессии на всех факторах

Коэффициенты	Полная регрессия f(X ₁ ,X ₂ ,X ₃ ,X ₄)	Регрессия с исключенным фактором			
		X ₁	X ₂	X ₃	X ₄
<i>b</i> ₀	9,9128	10,3350	11,3661	11,3404	3,7178
<i>b</i> ₁	0,1477	-	0,2212	0,3981	0,2821
<i>b</i> ₂	0,0040	0,0042	-	0,0037	0,0046
<i>b</i> ₃	0,6202	0,7426	0,5643	-	0,7645
<i>b</i> ₄	-0,5319	-0,5498	-0,5600	-0,5713	-
<i>R</i>²	0,6047	0,5978	0,5380	0,5449	0,4166

Рассчитаем частные F-критерии:

$$F_{x_1} = \frac{R^2_{yx_1x_2x_3x_4} - R^2_{yx_2x_3x_4}}{1 - R^2_{yx_1x_2x_3x_4}} \cdot \frac{20 - 4 - 1}{1} = \frac{0,6047 - 0,5978}{0,6047} \cdot 15 = 0,1728 < F_{5\%,1;15}$$

$$F_{x_2} = \frac{R^2_{yx_1x_2x_3x_4} - R^2_{yx_1x_3x_4}}{1 - R^2_{yx_1x_2x_3x_4}} \cdot \frac{20 - 4 - 1}{1} = \frac{0,6047 - 0,5380}{0,6047} \cdot 15 = 1,6559 < F_{5\%,1;15}$$

$$F_{x_3} = \frac{R^2_{yx_1x_2x_3x_4} - R^2_{yx_1x_2x_4}}{1 - R^2_{yx_1x_2x_3x_4}} \cdot \frac{20 - 4 - 1}{1} = \frac{0,6047 - 0,5449}{0,6047} \cdot 15 = 1,4840 < F_{5\%,1;15}$$

$$F_{x_4} = \frac{R^2_{yx_1x_2x_3x_4} - R^2_{yx_1x_2x_3}}{1 - R^2_{yx_1x_2x_3x_4}} \cdot \frac{20 - 4 - 1}{1} = \frac{0,6047 - 0,4166}{0,6047} \cdot 15 = 4,6672 > F_{5\%,1;15}$$

Табличное значение F-критерия при 5%-м уровне значимости и числе степеней свободы $v_1 = 1$ и $v_2 = 15$ равно: $F_{5\%,1;15} = 4.543$. Сравнив вычисленные значения с табличным, делаем вывод о необходимости исключения из модели переменной X_1 (с минимальным значением F-статистики). Тогда модель имеет вид:

$$\hat{Y} = 10,3350 + 0,0042X_2 + 0,7426X_3 - 0,5498X_4.$$

Теперь строим регрессии $Y = f(X_3, X_4)$, $Y = f(X_2, X_4)$, $Y = f(X_2, X_3)$ сравниваем их с $Y = f(X_2, X_3, X_4)$ (табл. 2.9).

Таблица 2.9 – Результаты построения регрессии без учета фактора X_1

Коэффициенты	Полная регрессия $f(X_2, X_3, X_4)$	Регрессия с исключенным фактором		
		X_2	X_3	X_4
b_0	10,3350	12,1311	14,1920	4,1384
b_2	0,0042		0,0043	0,0051
b_3	0,7426	0,7481		1,01531
b_4	-0,5498	-0,5897	-0,6754	
R^2	0,5978	0,5219	0,4689	0,3904

Табличное значение F-критерия при 5%-м уровне значимости и числе степеней свободы $v_1 = 1$ и $v_2 = 16$ равно: $F_{5\%, 1; 16} = 4,494$.

$$F_{x_2} = \frac{R^2_{yx_2x_3x_4} - R^2_{yx_3x_4}}{1 - R^2_{yx_2x_3x_4}} \cdot \frac{20 - 3 - 1}{1} = \frac{0,5978 - 0,5219}{0,5978} \cdot 16 = 2,0297 < F_{5\%, 1; 16}$$

$$F_{x_3} = \frac{R^2_{yx_2x_3x_4} - R^2_{yx_2x_4}}{1 - R^2_{yx_2x_3x_4}} \cdot \frac{20 - 3 - 1}{1} = \frac{0,5978 - 0,4689}{0,5978} \cdot 16 = 3,4507 < F_{5\%, 1; 16}$$

$$F_{x_4} = \frac{R^2_{yx_2x_3x_4} - R^2_{yx_2x_3}}{1 - R^2_{yx_2x_3x_4}} \cdot \frac{20 - 3 - 1}{1} = \frac{0,5978 - 0,3904}{0,5978} \cdot 16 = 5,5516 > F_{5\%, 1; 16}$$

Минимальное значение частной F-статистики связано с фактором X_2 . Исключаем этот фактор из модели и еще раз выполним расчеты F-критерия, используя данные табл. 2.10.

Таблица 2.10 – Результаты построения регрессии без учета факторов X_1 и X_2

Коэффициенты	Полная регрессия $f(X_3, X_4)$	Регрессия с исключенным фактором	
		X_3	X_4
b_0	12,1311	16,0396	5,7757
b_3	0,7481		1,0461
b_4	-0,5897	-0,7167	
R^2	0,5219	0,3911	0,2799

Табличное значение F-критерия при 5%-м уровне значимости и числе степеней свободы $v_1 = 1$ и $v_2 = 17$ равно: $F_{5\%, 1; 17} = 4,4513$.

$$F_{x_3} = \frac{R^2_{yx_3x_4} - R^2_{yx_4}}{1 - R^2_{yx_3x_4}} \cdot \frac{20 - 2 - 1}{1} = \frac{0,5219 - 0,3911}{0,5219} \cdot 17 = 4,2616 < F_{5\%, 1; 17}$$

$$F_{x_4} = \frac{R^2_{yx_3x_4} - R^2_{yx_3}}{1 - R^2_{yx_3x_4}} \cdot \frac{20 - 2 - 1}{1} = \frac{0,5219 - 0,799}{0,5219} \cdot 17 = 7,852 > F_{5\%,1;17}$$

Окончательно модель регрессии после исключения незначимых переменных имеет вид:

$$\hat{Y} = 12,1311 + 0,7481X_3 - 0,5897X_4, \quad R^2 = 0,52195$$

ЛИНЕЙН

	-0,58973	0,74806	12,13109
	0,200988	0,346799	2,558897
	0,521947	2,605386	#Н/Д
	9,280444	17	#Н/Д
	125,992	115,3966	#Н/Д
t-статистика	-2,93414	2,157041	4,74075

$$\beta_3 = b_3 \frac{S_{x_3}}{S_y} = 0,74806 \frac{1,80255}{3,56436} = 0,378,$$

$$\beta_4 = b_4 \frac{S_{x_4}}{S_y} = -0,58973 \frac{3,11025}{3,56436} = -0,515$$

Коэффициент детерминации модели изменился незначительно (в первоначальной модели был равен 0,6047, а теперь составляет 0,522). Но уравнение регрессии при этом стало существенно проще: вместо четырех содержит всего две переменные.

Кроме того, вместо одной значимой переменной (X_4) теперь их стало две: X_3 и X_4 ($t_3 = 2,157$ и $|t_4| = 2,934$, что больше табличного $t_{5\%,17} = 2,11$), т. е. окончательно делаем вывод, что на количество посещений магазинов фирмы значимо влияют число потенциальных покупателей и число конкурирующих магазинов, и если судить по β -коэффициентам, то в меньшей мере – число потенциальных покупателей. Разные знаки при этих коэффициентах говорят о разной направленности таких влияний. А поскольку коэффициент парной корреляции между этими показателями является статистически незначимым (

$r_{x_3x_4} = -0,2929$, $|t_{3,4}| = 1,2995 < t_{\text{табл}}$ (т. е. эти переменные независимы), то можем привести интерпретацию их коэффициентов.

Так, коэффициент при X_3 , т. е. b_3 , равен 0,748, следовательно, изменение числа потенциальных покупателей на 1 тыс. приведет в среднем к изменению числа посещений магазинов фирмы на 748 человек (уменьшит при уменьшении и увеличит при увеличении). Коэффициент при X_4 , т. е. b_4 , равен (-0,59), следовательно, изменение числа конкурирующих магазинов на единицу приведет в среднем к изменению числа посещений магазинов фирмы на 590 человек (увеличит при уменьшении и уменьшит при увеличении).

При этом надо иметь в виду, что мы сумели описать последним уравнением изменение числа посещений магазинов фирмы только на 52,2%. Остальные 47,8% изменения числа посещений магазинов фирмы зависят от неучтенных в регрессии факторов, а также от ошибок наблюдений.

Выполним прогноз количества посещений в месяц (тыс. чел.) при X_3 – число потенциальных покупателей, равным 4,2 тыс. чел. и X_4 – число конкурирующих магазинов, равным 15, т.е. $X'_0 = (1 \ 4,2 \ 15)$.

Точечный прогноз:

$$\hat{Y}_0 = 12,1311 + 0,7481 \cdot 4,2 - 0,5897 \cdot 15 = 6,43 (\text{тыс.чел})$$

Интервальный прогноз:

1) среднего количества посещений в месяц для таких же магазинов, т.е. $M_x(Y)$:

$$t_{\alpha;v} = t_{0,05;17} = 2,11$$

$$6,43 - 2,11 \cdot 2,605 \sqrt{\mathbf{X}'_0 (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}_0} \leq M_x(Y) \leq 6,43 + 2,11 \cdot 2,605 \sqrt{\mathbf{X}'_0 (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}_0}$$

$$(\mathbf{X}'\mathbf{X}) = \begin{pmatrix} 20 & 74,285 & 178 \\ 74,285 & 337,6478 & 629,941 \\ 178 & 629,941 & 1768 \end{pmatrix}; \quad (\mathbf{X}'\mathbf{X})^{-1} = \begin{pmatrix} 0,9646 & -0,0926 & -0,0641 \\ -0,0926 & 0,0177 & 0,0030 \\ -0,0641 & 0,0030 & 0,0060 \end{pmatrix}$$

$$\mathbf{X}'_0 (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}_0 = (1 \ 4,2 \ 15) \begin{pmatrix} 0,9646 & -0,0926 & -0,0641 \\ -0,0926 & 0,0177 & 0,0030 \\ -0,0641 & 0,0030 & 0,0060 \end{pmatrix} \begin{pmatrix} 1 \\ 4,2 \\ 15 \end{pmatrix} = (0,414 \ 0,032 \ 0,039) \begin{pmatrix} 1 \\ 4,2 \\ 15 \end{pmatrix} = 0,3015$$

$$\sqrt{\mathbf{X}'_0(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}_0} = \sqrt{0,3015} = 0,549,$$

$$s_p = 2,605 \cdot 0,549 = 1,43 \text{ (тыс.чел.)}$$

$$6,43 - 2,11 \cdot 1,43 \leq M_x(Y) \leq 6,43 + 2,11 \cdot 1,43$$

$$3,41 \leq M_x(Y) \leq 9,45.$$

Итак, с вероятностью 0,95 среднее количество посещений в месяц для магазинов с числом потенциальных покупателей 4,2 тыс. чел. и числом конкурирующих магазинов, равным 15, находится в пределах от 3,41 до 9,45 тысяч человек.

2) для индивидуального значения Y_0^* при $X'_0 = (1 \ 4,2 \ 15)$:

$$s_p = s_{\bar{Y}} \sqrt{1 + \mathbf{X}'_0(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}_0} = 2,605 \sqrt{1 + 0,3015} = 2,972 \text{ (тыс.чел.)}$$

$$6,43 - 2,11 \cdot 2,972 \leq Y_0^* \leq 6,43 + 2,11 \cdot 2,972$$

$$0,159 \leq Y_0^* \leq 12,701.$$

Таким образом, с вероятностью 0,95 количество посещений в месяц для магазина с числом потенциальных покупателей 4,2 тыс. чел. и числом конкурирующих магазинов, равным 15, находится в пределах от 159 человек до 12701 человек.

Вопросы и задания для самостоятельной работы

1. Опишите методику построения и применения эконометрической модели.
2. Каковы условия применения МНК для расчета оценок параметров модели. Каким должен быть результат оценивания при выполнении этих условий?
3. В чем состоит и с какой целью проводится проверка значимости оценок параметров и уравнения в целом.
4. Как Вы считаете, для чего можно было бы применить следующую формулу:

$$P \left(b_j - t_{\frac{\alpha}{2}, v=n-p-1} \cdot S_{b_j} \leq \beta_j \leq b_j + t_{\frac{\alpha}{2}, v=n-p-1} \cdot S_{b_j} \right) = 1 - \alpha ?$$

5. Какие этапы включает схема определения интервальных оценок коэффициентов регрессии?
6. Как строится и что позволяет определить доверительный интервал для условного математического ожидания зависимой переменной?
7. В чем суть предсказания индивидуальных значений зависимой переменной?
8. Как выполняется пошаговый регрессионный анализ с исключением?

Задание 2.1. В кейнсианской теории спрос на деньги зависит от доходов и процентных ставок. Рассмотрите следующую модель:

$$m_t = \beta_1 + \beta_2 Y_t + \beta_3 r_t + \varepsilon_t, \quad (1)$$

где m_t – агрегат денежной массы М1 (млрд. долл.), Y_t – валовой национальный продукт (ВНП) (млрд.долл.), r_t – процентные ставки по 6-месячным государственным облигациям США (%). В таб. 2.11 представлены данные по этим переменным за период 1960 – 1983 гг. по экономике США.

Таблица 2.11 – Исходные данные

Год	m_t	Y_t	r_t	Год	m_t	Y_t	r_t
1960	141,8	506,5	3,247	1972	251,9	1185,9	4,466
1961	146,5	524,6	2,605	1973	265,8	1326,4	7,178
1962	149,2	565,0	2,908	1974	277,5	1434,2	7,926
1963	154,7	596,7	3,253	1975	291,1	1549,2	6,122
1964	161,8	637,7	3,686	1976	310,4	1718,0	5,266
1965	169,5	691,1	4,055	1977	335,5	1918,3	5,510
1966	173,7	756,0	5,082	1978	363,2	2163,9	7,572
1967	185,1	799,6	4,630	1979	389,0	2417,8	10,017
1968	199,4	873,4	5,470	1980	414,1	2631,7	11,374
1969	205,8	944,0	6,853	1981	440,6	2954,1	13,776
1970	216,5	992,7	6,562	1982	478,2	3073,0	11,084
1971	230,7	1077,6	4,511	1983	521,1	3309,5	8,750

Задание:

- найдите оценки коэффициентов регрессии (1). Интерпретируйте знаки коэффициентов;
- рассчитайте прогноз спроса на деньги при значениях: 1) $Y_t = 1000$, $r_t = 10$; 2) $Y_t = 2500$, $r_t = 5$;
- рассчитайте эластичность спроса на деньги (m_t) по доходам (Y_t) и по процентным ставкам (r_t) в двух точках 1) и 2) из п.б). Сравните результаты;
- рассмотрите модель:

$$\ln m_t = \beta_1 + \beta_2 \ln Y_t + \beta_3 \ln r_t + \varepsilon_t, \quad (2)$$

Повторите п.п. б) и с) и сравните результаты, полученные по разным моделям. Сравните модели (1) и (2). Какая из них вам представляется более предпочтительной? Почему?

Задание 2.2. В табл. 2.12 представлены результаты бюджетного обследования пяти случайно выбранных семей (в тыс. руб.).

Таблица 2.12 – Исходные данные

Семья, i	Накопления, Y	Доход, X_1	Имущество, X_2
1	3,0	40	60
2	6,0	55	36
3	5,0	45	36
4	3,5	30	15
5	1,5	30	90

Задание.

- 1) Оцените регрессию Y на X_1 и X_2 .
- 2) Постройте 95%-й доверительный интервал для β_1 ; β_2 .
- 3) Проверьте с 5%-м уровнем значимости следующие гипотезы:
 - a) $\beta_2 = 0$ (стоимость имущества незначительна);
 - b) $\beta_{21} = 0$ (величина дохода незначительна);
 - c) $\beta_1 = 1,57$ (такое значение коэффициента β_1 могло бы быть с высокой степенью надежности установлено для другой страны, и вас интересует вопрос, верно ли это для вашей страны).
- 4) Спрогнозируйте накопления семьи, имеющей доход 30 тыс. руб. и имущество стоимостью 52,5 тыс. руб. Постройте 95%-й доверительный интервал для прогнозной величины накоплений этой семьи.

Задание 2.3. Зависимость темпов прироста выпуска продукции от темпов прироста производственных факторов (основных фондов и живого труда) определяются следующим уравнением:

$$\hat{Y} = 8,587 + 0,145X_1 + 0,618X_2.$$

Приведите экономическую интерпретацию количественного выражения взаимосвязи.

Сделайте вывод о согласованности модели и опытных данных, используя следующую информацию для определения коэффициентов корреляции и детерминации:

$$\sum_{i=1}^{150} (Y_i - \hat{Y}_i)^2 = 43,57, \quad \sum_{i=1}^{150} (\hat{Y}_i - \bar{Y})^2 = 215,65.$$

Проверьте статистическую надежность коэффициента детерминации.

Задание 2.4. В результате исследования факторов, определяющих экономический рост по 73 странам, получено следующее уравнение регрессии:

$$\hat{G} = 1,4 - 0,52 P + 0,17 S + 11,16 I - 0,38 D - 4,75 In, \quad R^2 = 0,6$$

(−5,9) (4,34) (3,91) (−0,79) (−2,7)

где G – темпы экономического роста (темпы роста среднедушевого ВВП в % к базисному периоду);

P – реальный среднедушевой ВВП, %;

S – бюджетный дефицит, % к ВВП;

I – объем инвестиций, % к ВВП;

D – внешний долг, % к ВВП;

In – уровень инфляции, %.

В скобках указаны фактические значения t -критерия для коэффициентов множественной регрессии.

Задание.

1. Выполните содержательную интерпретацию модели.
2. Проверьте гипотезу о достоверности полученной модели в целом и коэффициентов регрессии. Сделайте выводы.

3. До получения результатов этого исследования ваш однокурсник заключил с вами пари, что эмпирические результаты по данной модели докажут наличие обратной связи между темпами экономического роста и объемом внешнего долга страны (% к ВВП). Выиграл ли это пари ваш однокурсник?

ТЕМА 3. ТЕСТИРОВАНИЕ И УСТРАНЕНИЕ МУЛЬТИКОЛЛИНЕАРНОСТИ

Данная тема знакомит студентов с требованиями (ограничениями), накладываемыми на факторы, включаемые в модель регрессии; с последствиями неправильного подбора факторов; со способами тестирования возникающих проблем и методами их устранения.

Цель выполнения лабораторной работы: научиться выявлять нарушения классической регрессии путем тестирования гипотезы об отсутствии мультиколлинеарности; описывать последствия этих нарушений для МНК-оценщика; выбирать методы оценивания параметров модели, позволяющие уменьшить отрицательные последствия, вызванные этими нарушениями.

Лабораторная работа включает проверку одной из основных предпосылок классической регрессии: об отсутствии мультиколлинеарности и получение «хорошей» модели.

Для выполнения лабораторной работы № 3 студенты должны

знать:

- содержание основных предпосылок классической регрессии;
- последствия нарушения предпосылок с точки зрения качества построенной методом наименьших квадратов модели;
- основные способы тестирования предпосылок и способы устранения негативных последствий нарушения предпосылок;

уметь:

- пользоваться таблицами для проверки статистических гипотез;
- пользоваться пакетом EXCEL:
 - строить таблицы; составлять и копировать формулы;
 - пользоваться МАСТЕРОМ ФУНКЦИЙ: статистическими и математическими функциями.

Задание для выполнения лабораторной работы №3

Для выполнения данной работы следует использовать исходные данные, использовавшиеся при выполнении лабораторной работы № 2.

Применяя критерий Феррара-Глоубера, выполните проверку факторов, включенных в модель, на мультиколлинеарность.

В случае необходимости (при наличии мультиколлинеарности) методом пошаговой регрессии постройте модель, которая должна обладать лучшими характеристиками.

Для выполнения работы следует изучить теоретический материал, представленный в теме 4 Учебного пособия по эконометрике [18].

РЕШЕНИЕ ТИПОВЫХ ЗАДАЧ

Пример 3.1. [2] По данным 20 сельскохозяйственных районов области исследуется зависимость переменной Y – урожайности зерновых культур (в ц/га) от ряда переменных – факторов сельскохозяйственного производства:

X_1 – число тракторов (приведенной мощности на 100 га);

X_2 – число зерноуборочных комбайнов на 100 га;

X_3 – число орудий поверхностной обработки почвы на 100 га;

X_4 – количество удобрений, расходуемых на 1га (т/га);

X_5 – количество химических средств защиты растений, расходуемых на 1га (т/га).

Исходные данные приведены в табл. 3.1.

Задание

1. Для построения модели, характеризующей зависимость урожайности зерновых культур от факторов сельскохозяйственного производства выполнить проверку факторов на мультиколлинеарность.

2. В случае обнаружения мультиколлинеарности принять меры по ее устранению (уменьшению), используя пошаговую процедуру отбора наиболее информативных переменных (пошаговую регрессию с включением факторов).

Таблица 3.1 – Исходные данные

i	Y	X1	X2	X3	X4	X5
1	9,7	1,59	0,26	2,05	0,32	0,14
2	8,4	0,34	0,28	0,46	0,59	0,66
3	9	2,53	0,31	2,46	0,3	0,31
4	9,9	4,63	0,4	6,44	0,43	0,59
5	9,6	2,16	0,26	2,16	0,39	0,16
6	8,6	2,16	0,3	2,69	0,32	0,17
7	12,5	0,68	0,29	0,73	0,42	0,23
8	7,6	0,35	0,26	0,42	0,231	0,08
9	6,9	0,52	0,24	0,49	0,2	0,08
10	13,5	3,42	0,31	3,02	1,37	0,73
11	9,7	1,78	0,3	3,19	0,73	0,17
12	10,7	2,4	0,32	3,3	0,25	0,14
13	12,1	9,36	0,4	11,51	0,39	0,38
14	9,7	1,72	0,28	2,26	0,82	0,17
15	7	0,59	0,29	0,6	0,13	0,35
16	7,2	0,28	0,26	0,3	0,09	0,15
17	8,2	1,64	0,29	1,44	0,2	0,08
18	8,4	0,09	0,22	0,05	0,43	0,2
19	13,1	0,08	0,25	0,03	0,73	0,2
20	8,7	1,36	0,26	0,17	0,99	0,42

Решение.

Построим модель линейной регрессии с помощью МНК (функция ЛИНЕЙН) (см.рис. 3.1 – фрагмент Листа Excel).

-2,89454	4,46446	0,111278	15,41705	-0,00488	3,535219	
3,090128	1,548034	0,834193	21,53855	0,933496	5,429066	
0,515414	1,602134	#Н/Д	#Н/Д	#Н/Д	#Н/Д	
2,978132	14	#Н/Д	#Н/Д	#Н/Д	#Н/Д	
38,22184	35,93566	#Н/Д	#Н/Д	#Н/Д	#Н/Д	
<hr/>						
-0,93671	2,883956	0,133397	0,715789	-0,00523	0,651165	t
<hr/>						
t табл = 2,14						
Fтабл(5%;v=5, v = 14) = 2,96						
2,978 > 2,96						

Рис.3.1. Результаты построения модели линейной регрессии с помощью МНК

Выводы, вытекающие из уравнения регрессии

$$\hat{Y} = 3,535 - 0,0049X_1 + 15,417X_2 + 0,111X_3 + 4,4645X_4 - 2,895X_5,$$

полученного по всей совокупности факторов, противоречат сущности рассматриваемой проблемы: отрицательные коэффициенты при X_1 – число тракторов (приведенной мощности на 100 га) и X_5 – количество химических средств защиты растений, расходуемых на 1га (т/га) говорят о том, что с ростом соответствующего фактора на единицу и при прочих равных условиях урожайность зерновых культур будет снижаться. Такой результат может быть следствием мультиколлинеарности.

Из рис. 3.1 видим, что статистически значимым оказался только коэффициент регрессии b_4 при переменной X_4 - количество удобрений, расходуемых на 1га (т/га), т.к. $2,884 > 2,14$.

Вычисленный коэффициент детерминации урожайности зерновых культур по совокупности пяти факторов ($X_1 - X_5$) с/хозяйственного производства оказался равным 0,515, т.е. 51,5% вариации зависимой переменной объясняется включенными в модель факторами. Поскольку фактическое значение $F = 2,978$ больше табличного, то уравнение регрессии значимо на уровне 5%. Таким образом, выводы о статистической значимости результатов оценивания оказались противоречивыми. Это тоже, как правило, является следствием мультиколлинеарности.

Проверка мультиколлинеарности методом Феррара-Глоубера

На первом этапе проверки мультиколлинеарности вычислим критерий χ^2 :

$$\chi^2 = - \left[n - 1 - \frac{1}{6}(2p + 5) \right] \ln |R|$$

где $|R|$ – определитель матрицы парных корреляций между факторами (R),

n – число наблюдений;

p – количество объясняющих переменных.

С помощью «Пакета анализа» построим матрицу парных корреляций между всеми показателями (табл. 3.2).

Таблица 3.2 – Матрица парных корреляций:

	У	X1	X2	X3	X4	X5
У	1	0,43025	0,374079	0,403153	0,575473	0,332137
X1	0,43025	1	0,854254	0,977908	0,108248	0,341013
X2	0,374079	0,854254	1	0,88192	0,02474	0,459592
X3	0,403153	0,977908	0,88192	1	0,027648	0,277923
X4	0,575473	0,108248	0,02474	0,027648	1	0,568757
X5	0,332137	0,341013	0,459592	0,277923	0,568757	1

На пересечении строк и столбцов X1, X2, X3, X4, X5 табл. 3.2 находится матрица R – матрица парных корреляций между факторами (независимыми переменными). Определитель матрицы R равен 0,0031 (использовали функцию МОПРЕД Excel).

Вычислим критерий χ^2 :

$$\chi^2 = - \left[n - 1 - \frac{1}{6}(2p + 5) \right] \ln |R| = - \left[20 - 1 - \frac{1}{6}(2 \cdot 5 + 5) \right] \ln |0,0031| = 95,55$$

Табличное значение критерия χ^2 найдем с помощью функции **ХИ2ОБР.ПХ**

при $\nu = \frac{1}{2} p(p - 1) = \frac{1}{2} 5(5 - 1) = 10$ степенях свободы и уровне значимости $\alpha = 0,05$:

$$\chi^2_{\text{табл}} = \chi^2_{0,05;10} = 18,307$$

Поскольку $\chi^2 = 95,55 > \chi^2_{0,05;10} = 18,307$, то делаем вывод о наличии мультиколлинеарности и переходим ко второму этапу исследования.

Найдем матрицу С, обратную матрице корреляций, применяя функцию **МОБР** Excel:

Обратная матрица:

29,14587	3,287925	-30,7346	-1,08265	-2,2926
3,287925	7,189521	-8,84967	1,225458	-2,66293
-30,7346	-8,84967	37,77282	0,29267	3,883712
-1,08265	1,225458	0,29267	1,826115	-1,31397
-2,2926	-2,66293	3,883712	-1,31397	2,673622

Найдем величину F-критериев для каждой объясняющей переменной, используя элементы матрицы С: $F_j = (c_{jj} - 1) \frac{n-p}{p-1}$

$$F_1 = (29,14587 - 1) \frac{20-5}{5-1} = 105,547 > F_T; \quad F_2 = (7,189521 - 1) \frac{20-5}{5-1} = 23,2107 > F_T;$$

$$F_3 = (37,77282 - 1) \frac{20-5}{5-11} = 137,8981 > F_T; \quad F_4 = (1,826115 - 1) \frac{20-5}{5-1} = 3,0979 > F_T;$$

$$F_5 = (2,673622 - 1) \frac{20-5}{5-1} = 6,2761 > F_T. \quad \text{Здесь } F_T = F_{5\%;4;15} = 3,06.$$

То есть, каждый из факторов связан тесной линейной зависимостью с остальными.

Наконец, выявим *мультиколлинеарные пары*. Для этого:

1) определим *частные коэффициенты корреляции*, показывающие тесноту связи между двумя объясняющими переменными при условии, что остальные переменные не влияют на эту связь:

$$r_{kj} = \frac{-c_{kj}}{\sqrt{c_{kk} c_{jj}}}$$

$$r_{12} = \frac{-(3,2879)}{\sqrt{29,14587 \cdot 7,1895}} = -0,22713;$$

$$r_{23} = \frac{-(-8,8497)}{\sqrt{7,1895 \cdot 37,7728}} = 0,537016;$$

$$r_{13} = \frac{-(-30,7346)}{\sqrt{29,14587 \cdot 37,7728}} = 0,926294;$$

$$r_{24} = \frac{-(1,2254)}{\sqrt{7,1895 \cdot 1,8261}} = -0,33821;$$

$$r_{14} = \frac{-(-1,0826)}{\sqrt{29,14587 \cdot 1,8261}} = 0,148401; \quad r_{25} = \frac{-(-2,66293)}{\sqrt{7,1895 \cdot 2,6736}} = 0,607379.$$

$$r_{15} = \frac{-(-2,2926)}{\sqrt{29,14587 \cdot 2,6736}} = 0,25971; \quad r_{34} = \frac{-(0,2927)}{\sqrt{37,7728 \cdot 1,8261}} = -0,03524$$

$$r_{45} = \frac{-(-1,31397)}{\sqrt{1,8261 \cdot 2,6736}} = 0,594663; \quad r_{35} = \frac{-(3,8837)}{\sqrt{37,7728 \cdot 2,6736}} = -0,38646$$

2) определим t-критерии для определения взаимосвязанных пар, учитывая, что $t_{\text{табл}}=2,13$:

$$t_{kj} = \frac{r_{kj} \cdot \sqrt{n-p}}{\sqrt{1-r_{kj}^2}}$$

$$t_{12} = \frac{-0,22713 \sqrt{20-5}}{\sqrt{1-(-0,22713)^2}} = -0,9033; \quad t_{23} = \frac{0,537016 \sqrt{20-5}}{\sqrt{1-0,537016^2}} = 2,465532 > t_{\text{табл}};$$

$$t_{13} = \frac{0,966294 \sqrt{20-5}}{\sqrt{1-0,966294^2}} = 9,520995 > t_{\text{табл}}; \quad t_{24} = \frac{-0,33821 \sqrt{20-5}}{\sqrt{1-(-0,33821)^2}} = -1,3919;$$

$$t_{14} = \frac{0,148401 \sqrt{20-5}}{\sqrt{1-0,148401^2}} = 0,581188; \quad t_{25} = \frac{0,607379 \sqrt{20-5}}{\sqrt{1-0,607379^2}} = 2,961144 > t_{\text{табл}};$$

$$t_{15} = \frac{0,259711 \sqrt{20-5}}{\sqrt{1-0,259711^2}} = 1,041596 \quad t_{34} = \frac{-0,03524 \sqrt{20-5}}{\sqrt{1-(-0,03524)^2}} = -0,13567$$

$$t_{45} = \frac{0,594663 \sqrt{20-5}}{\sqrt{1-0,594663^2}} = 2,864666 > t_{\text{табл}} \quad t_{35} = \frac{-0,38646 \sqrt{20-5}}{\sqrt{1-(-0,38646)^2}} = -1,62285$$

Вычисленные t-критерии сравниваем с табличным значением, которое при числе степеней свободы $20-5=15$ и уровне значимости $\alpha = 5\%$ составляет 2,13. Поскольку $t_{13} > t_{\text{табл}}$, $t_{23} > t_{\text{табл}}$, $t_{25} > t_{\text{табл}}$ и $t_{45} > t_{\text{табл}}$, то факторы X_1 и X_3 , X_2 и X_3 , X_2 и X_5 , X_4 и X_5 составляют мультиколлинеарные пары.

Учитывая негативное влияние мультиколлинеарности на результаты эконометрического оценивания, попытаемся устранить мультиколлинеарность с применением **процедуры пошагового отбора (со включением) наиболее информативных переменных.**

1-й шаг. Из объясняющих переменных выберем X_4 , имеющую с зависимой переменной Y наибольший коэффициент детерминации $R^2(1) = 0,575473^2 = 0,332$. С учетом поправки на несмещенность:

$$\overline{R^2}(1) = 1 - (1 - R^2) \frac{n - 1}{n - p - 1} = 1 - (1 - 0,332) \frac{20 - 1}{20 - 1 - 1} = 0,294.$$

Нижняя доверительная граница $R^2_{\min}(1)$:

$$R^2_{\min}(1) = \overline{R^2}(k) - 2 \sqrt{\frac{2k(n - k - 1)}{(n - 1)(n^2 - 1)}} (1 - R^2(k)) = 0,294 - 2 \sqrt{\frac{2(20 - 1 - 1)}{(20 - 1)(400 - 1)}} (1 - 0,332) = 0,2018.$$

Итак, $\boxed{R^2_{\min}(1) = 0,2018}$.

2-й шаг. Среди всевозможных пар объясняющих переменных (X_4, X_j) выбирается пара (X_4, X_3), имеющая с зависимой переменной Y наиболее высокий коэффициент детерминации $R^2(2) = 0,481$.

Y=f(X4,X2)		
3,488548	15,546142	3,404583
1,096995	7,6922782	2,294624
0,460734	1,5337495	#Н/Д
7,262178	17	#Н/Д
34,16691	39,990586	#Н/Д
3,180094	2,0210062	1,483722

$\hat{Y} = 3,4046 + 15,546X_2 + 3,4885X_4$
t-статистика

Y=f(X4,X1)		
3,295237	0,346033	7,33568
1,095514	0,165358	0,663186
0,468166	1,523144	#Н/Д
7,482439	17	#Н/Д
34,71804	39,43946	#Н/Д
3,007937	2,09263	11,06128

$\hat{Y} = 7,3357 + 0,346X_1 + 3,295X_4$
t-статистика

Оценка уравнения регрессии урожайности зерновых культур Y по X_3 и X_4 имеет вид:

Y=f(X4,X3)		
3,477423	0,282729	7,283856
1,076019	0,127491	0,659751
0,48124	1,504306	#Н/Д
7,885231	17	#Н/Д
35,68757	38,46993	#Н/Д
3,23175	2,217634	11,04031

$\hat{Y} = 7,284 + 0,283X_3 + 3,477X_4$
t-статистика

Значения скорректированного коэффициента детерминации и его нижней доверительной границы соответственно равны

$$\overline{R^2(2)} = 0,4202 \quad \text{и} \quad \boxed{R^2_{\min}(2) = 0,3219}.$$

Целесообразность включения в модель второй объясняющей переменной X_3 подтверждается сравнением нижних доверительных границ: $R^2_{\min}(1) < R^2_{\min}(2)$.

3-й шаг. Среди всевозможных троек объясняющих переменных (X_3, X_4, X_5) наиболее информативной оказалась тройка (X_3, X_4, X_5), поскольку она имеет максимальный коэффициент детерминации $R^2(3) = 0,4966$.

Уравнение регрессии по этим переменным:

$$\hat{Y} = 7,3978 + 0,3132X_3 + 4,028X_4 - 1,6177X_5$$

-1,61766	4,0280139	0,313215	7,397837
2,311157	1,3462611	0,136579	0,689387
0,496652	1,5273957	#Н/Д	#Н/Д
5,262392	16	#Н/Д	#Н/Д
36,8305	37,327004	#Н/Д	#Н/Д
-0,69994	2,9920006	2,293284	10,73104

t - статистика

Находим $\overline{R^2(3)} = 0,4023$ и $\boxed{R^2_{\min}(3) = 0,289}$.

Сравнение нижних доверительных границ $R^2_{\min}(2)$ и $R^2_{\min}(3)$ говорит о том, что третью объясняющую переменную X_5 в модель включать нецелесообразно, так как $R^2_{\min}(3) < R^2_{\min}(2)$ (см. также рис.3.2).

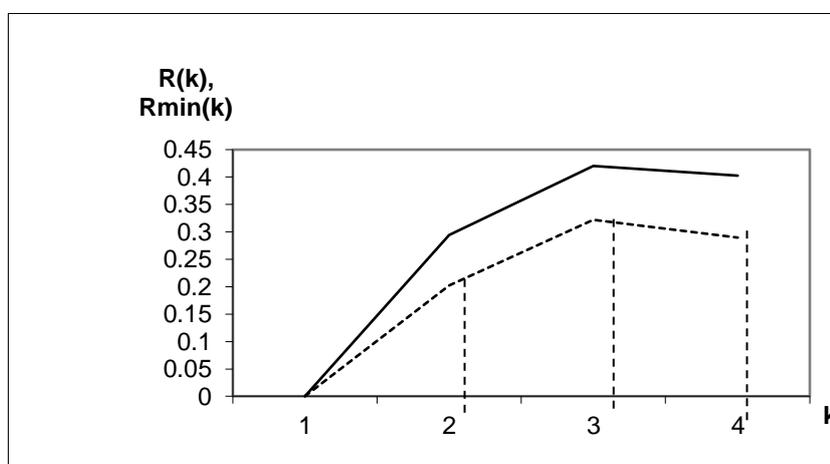


Рис. 3.2. График нижних доверительных границ коэффициентов детерминации

Этот вывод подтверждается также сравнением скорректированных коэффициентов детерминации $\overline{R^2(2)}$ и $\overline{R^2(3)}$: добавление X_5 в модель в качестве третьей объясняющей переменной не только не повышает значение $\overline{R^2}$, а даже несколько снижает его. (Кроме того, переменные X_4 и X_5 мультиколлинеарны).

Окончательно получили модель:

$$\hat{Y} = 7,284 + 0,283X_3 + 3,477X_4.$$

Заметим, что все три коэффициента регрессии статистически значимы (все t -статистики больше 2,2).

Проверка этой модели показала, что мультиколлинеарность отсутствует.

Таким образом, при увеличении X_3 – число орудий поверхностной обработки почвы на 100 га – на 1 единицу и при неизменном X_4 – количество удобрений, расходуемых на 1га (т/га) урожайность зерновых культур (Y) в среднем повышается на 0,283 ц/га; если же количество число орудий поверхностной обработки почвы остается постоянным, а количество удобрений, расходуемых на 1га, возрастет на 1т, то в среднем рост урожайности зерновых культур повысится на 3,477 ц/га.

Вопросы и задания для самостоятельной работы

1. В чем состоит проблема мультиколлинеарности (влияние на результат эконометрического моделирования, способы снижения влияния).
2. Приведите алгоритм выявления мультиколлинеарности с помощью критерия Фаррара-Глобера.
3. Охарактеризуйте методы «смягчения» проблемы мультиколлинеарности.
4. Приведите алгоритм пошагового отбора наиболее информативных факторов для включения в модель.

Задание 3.1. [3] В табл.3.4 приведены данные по реальному ВВП (X_1), объему потребления (X_2) и объему инвестиций (Y) для некоторой страны.

Таблица 3.4 – Исходные данные

X_1	240	248	261	274	273	269	283	296	312	319
X_2	149	154	162	169	167	171	180	188	196	200
Y	38.2	41.9	46.5	52.1	48.1	38.3	45.4	52.1	56.8	57.5

X_1	318	325	317	327	350	361	372	385	402	412	
X_2	200	202	205	215	225	235	245	252	261	266	
Y		50.9	54.5	44.7	50.4	65.8	63.7	64.0	76.4	71.6	71.8

- а) Постройте уравнение регрессии $Y = b_0 + b_1X_1 + b_2X_2 + e$
- б) Оцените качество построенного уравнения.
- в) Можно ли было ожидать при построении данного уравнения наличия мультиколлинеарности? Ответ поясните.
- г) Имеет ли место мультиколлинеарность для построенного вами уравнения? Как вы это определили?
- д) Постройте уравнения регрессии Y на X_1 и Y на X_2 . Какие выводы можно сделать по построенным моделям?
- е) Постройте уравнение регрессии X_2 на X_1 . Что обнаруживает построенная модель?
- ж) Как можно решить проблему мультиколлинеарности для первоначальной модели?

Задание 3.2 [8]. В табл. 3.3 представлены данные о потреблении некоторого продукта Y (усл. ед.) в зависимости от уровня урбанизации (доли городского населения) X_1 , относительного образовательного уровня X_2 и относительного заработка X_3 для девяти географических районов.

Таблица 3.3 – Исходные данные

i	Y	X_1	X_2	X_3	i	Y	X_1	X_2	X_3
1	167,1	42,2	11,2	31,9	6	174,6	44,5	10,8	8,5
2	174,4	48,6	10,6	13,2	7	163,7	39,1	10,7	24,3
3	160,8	42,6	10,6	28,7	8	174,5	40,1	10,0	18,6
4	162,0	39,0	10,4	26,1	9	185,7	45,9	12,0	20,4
5	140,8	34,7	9,3	30,1					

Используя пошаговую процедуру отбора наиболее информативных объясняющих переменных, постройте подходящую регрессионную модель, исключив при этом мультиколлинеарность. Оцените статистическую значимость коэффициентов регрессии и модели в целом.

Задание 3.3. Используем матричную запись для дисперсионно-ковариационной матрицы: $Var(b) = \sigma^2(X'X)^{-1}$

Что случится с этой матрицей, если:

- 1) существует полная корреляция;
- 2) коллинеарность высока, но не полная?

Задание 3.4. Известна матрица корреляций между переменными X_1, X_2, \dots, X_p

$$R = \begin{bmatrix} 1 & r_{12} & \dots & r_{1p} \\ r_{21} & 1 & \dots & r_{2p} \\ \dots & \dots & \dots & \dots \\ r_{p1} & r_{p2} & \dots & 1 \end{bmatrix}$$

Как определить для этой матрицы:

- а) имеется ли полная мультиколлинеарность?
- б) имеется ли меньшая, чем полная, коллинеарность?
- в) X не коррелированы?

Используйте для ответа на вопрос определитель матрицы R .

ТЕМА 4. ИСПОЛЬЗОВАНИЕ ФИКТИВНЫХ ПЕРЕМЕННЫХ В ЭКОНОМЕТРИЧЕСКОМ АНАЛИЗЕ

Изучение данной темы направлено на формирование у студентов представлений о возможности и способах модельного отражения влияния на изменчивость исследуемого показателя качественных факторов (т.е. не имеющих количественного описания) с помощью так называемых «фиктивных» переменных.

Цель лабораторной работы: научиться использовать фиктивные переменные в эконометрическом анализе.

Для выполнения лабораторной работы № 4 студенты должны

знать:

- сущность и назначение фиктивных переменных;
- способы введения фиктивных переменных в модель;
- методы построения и исследования модели, включающей фиктивные переменные;
- построение прогнозов на основе построенной модели с учетом особенностей фиктивных переменных;

уметь:

- пользоваться таблицами для проверки статистических гипотез;
- пользоваться пакетом EXCEL:
 - строить таблицы; составлять и копировать формулы;
 - пользоваться МАСТЕРОМ ФУНКЦИЙ: статистическими и математическими функциями.

Задание для выполнения лабораторной работы №4

На основе анализа временного ряда, выбранного в соответствии со своим вариантом (номером предприятия) из табл. 4.1, составьте прогноз потребления электроэнергии на следующий год, используя *фиктивные переменные*.

Таблица 4.1 – Потребление электроэнергии (тыс. кВт.ч)

№ пред приятия	2008 год				2009 год			
	1 кв.	2 кв.	3 кв.	4 кв.	1 кв.	2 кв.	3 кв.	4 кв.
1	452	127	338	1148	511	126	371	1096
2	402	114	401	1067	419	145	589	944
3	285	63	195	731	371	63	289	754
4	103	25	29	178	121	13	44	135
5	231	92	220	529	216	72	282	393
6	169	29	80	404	191	45	142	358
7	123	20	96	338	149	22	104	287
8	242	57	174	473	226	54	192	456
9	359	71	398	764	330	109	442	726
10	159	50	293	350	133	75	348	339
11	164	46	218	410	163	68	250	393
12	284	110	322	812	349	109	386	679
13	123	46	104	339	121	28	153	298
14	224	65	202	486	250	67	271	552
15	222	68	204	611	255	62	181	616
16	126	45	208	336	116	38	188	290
17	75	21	98	294	92	24	117	223
18	165	66	196	510	167	73	182	445
19	29	20	34	92	28	17	40	46
20	415	79	364	577	364	88	415	557
21	604	200	837	1761	582	257	1008	1615
22	361	96	368	982	351	116	521	942
23	116	23	91	304	103	31	114	348
24	89	27	48	366	109	22	63	278
25	161	48	68	482	175	25	66	616
№ предпр.	2010 год				2011 год			
	1 кв.	2 кв.	3 кв.	4 кв.	1 кв.	2 кв.	3 кв.	4 кв.
1	532	126	286	1065	495	98	398	1299
2	426	159	506	972	431	152	491	964
3	307	43	253	629	325	68	298	569
4	92	17	49	137	67	11	24	143
5	179	51	211	337	157	53	163	354
6	176	30	111	331	172	46	131	327
7	111	21	77	261	134	16	88	308
8	197	51	210	414	200	55	174	500
9	339	114	397	548	230	119	389	614
10	171	95	304	563	161	112	314	211
11	160	59	251	336	136	100	253	351
12	316	130	313	658	297	138	342	584
13	184	33	131	316	97	43	192	320
14	190	122	246	427	183	100	286	409
15	274	66	199	495	237	85	330	630
16	127	63	149	238	116	64	166	231
17	87	32	155	277	62	33	168	212
18	143	57	190	406	144	44	131	330
19	27	29	25	55	29	34	52	68
20	266	121	299	500	225	128	262	392
21	781	315	1030	1293	707	413	917	1252
22	378	153	461	732	368	172	501	723
23	117	43	100	203	109	40	88	304
24	114	21	126	244	95	44	103	205
25	198	20	111	546	182	31	188	450

№ предпр.	2012 год				2013 год			
	1 кв.	2 кв.	3 кв.	4 кв.	1 кв.	2 кв.	3 кв.	4 кв.
1	587	125	430	1184	610	139	307	1121
2	464	254	531	676	481	250	412	697
3	274	48	299	670	300	71	223	647
4	87	5	12	150	85	11	14	150
5	148	49	187	380	192	60	146	299
6	185	48	94	279	160	39	80	250
7	169	31	78	257	128	17	44	217
8	265	69	183	497	279	55	170	483
9	263	163	344	454	195	165	262	396
10	163	132	207	193	137	130	178	175
11	194	105	226	294	191	125	194	276
12	341	183	362	648	337	146	308	548
13	153	68	205	263	145	107	156	213
14	198	126	255	397	173	147	240	327
15	303	96	378	630	279	122	363	598
16	116	75	141	132	101	72	97	148
17	119	61	167	162	85	68	143	159
18	125	65	143	276	96	71	136	204
19	31	23	34	68	28	29	22	53
20	214	137	287	365	256	126	271	348
21	731	483	753	1065	693	501	591	695
22	400	253	456	609	365	286	447	595
23	156	65	111	235	101	66	99	176
24	115	81	143	160	94	78	126	133
25	180	76	228	297	163	110	195	240
№ предпр.	2014 год				2015 год			
	1 кв.	2 кв.	3 кв.	4 кв.	1 кв.	2 кв.	3 кв.	4 кв.
1	574	131	333	1203	562	137	396	1061
2	437	227	415	660	461	190	393	612
3	293	36	140	673	293	29	130	375
4	67	9	7	137	66	8	11	84
5	154	50	104	255	147	48	82	207
6	166	47	79	221	140	47	83	161
7	94	17	26	174	104	15	27	123
8	226	82	236	509	291	84	172	442
9	172	113	254	400	227	135	231	286
10	166	125	190	217	200	168	217	141
11	131	92	169	224	133	98	168	156
12	291	149	305	673	272	75	179	386
13	94	64	104	204	78	44	197	210
14	210	139	183	291	233	127	175	279
15	244	87	316	582	303	74	260	452
16	94	66	83	121	97	61	63	109
17	102	70	111	147	84	75	76	124
18	84	39	85	213	85	40	85	168
19	18	22	22	56	29	23	30	41
20	223	130	236	325	247	115	244	282
21	581	361	561	962	630	373	556	831
22	380	233	435	494	370	234	344	465
23	111	47	107	146	189	70	101	163
24	136	79	93	174	102	89	104	132
25	185	82	192	304	165	124	179	209

№ предпр.	2016 год				2017 год			
	1 кв.	2 кв.	3 кв.	4 кв.	1 кв.	2 кв.	3 кв.	4 кв.
1	546	119	381	1142	536	129	371	1132
2	369	176	356	780	398	166	346	781
3	224	261	122	368	254	251	112	358
4	68	5	18	92	70	6	8	90
5	111	33	96	253	114	37	86	243
6	112	27	56	164	132	29	46	154
7	84	15	34	141	74	14	24	131
8	272	64	135	468	252	68	125	458
9	146	79	153	290	156	75	143	280
10	166	140	189	214	176	148	179	204
11	113	90	128	191	123	94	118	181
12	291	106	256	424	281	103	246	414
13	95	34	172	289	85	37	162	279
14	163	89	154	218	153	87	144	208
15	230	56	111	496	235	58	101	486
16	70	49	63	117	75	46	53	107
17	54	48	67	123	50	44	60	119
18	66	27	48	158	68	28	41	152
19	28	32	41	69	38	35	31	62
20	205	85	232	316	215	89	222	310
21	555	329	501	798	675	327	511	791
22	423	281	394	582	443	288	384	589
23	109	61	75	188	119	65	71	180
24	111	56	97	156	121	59	90	149
25	177	84	162	219	187	87	152	209

Для выполнения работы следует изучить теоретический материал, представленный в теме 5 Учебного пособия по эконометрике [18].

РЕШЕНИЕ ТИПОВЫХ ЗАДАЧ

Пример 1. В табл. 4.2 представлены расходы потребителей на газ и электричество в США в постоянных ценах с 1 квартала 2011г. по 4 квартал 2016г. Необходимо построить модель, описывающую процесс потребления электроэнергии во времени и составить прогноз на следующий год, используя *фиктивные переменные*.

Решение.

На рис. 4.1 представлено графическое изображение рассматриваемого временного ряда, который характеризуется небольшой тенденцией к повышению и сильными сезонными колебаниями. Как и следовало предполагать, расходы такого рода всегда значительно выше зимой, чем летом.

Таблица 4.2 – Расходы потребителей на газ и электричество

Год	Квартал	Y	t	Z ₂	Z ₃	Z ₄
2011	1	7,33	1	0	0	0
	2	4,7	2	1	0	0
	3	5,1	3	0	1	0
	4	5,46	4	0	0	1
2012	1	7,65	5	0	0	0
	2	4,92	6	1	0	0
	3	5,15	7	0	1	0
	4	5,55	8	0	0	1
2013	1	7,96	9	0	0	0
	2	5,01	10	1	0	0
	3	5,05	11	0	1	0
	4	5,59	12	0	0	1
2014	1	7,74	13	0	0	0
	2	5,1	14	1	0	0
	3	5,67	15	0	1	0
	4	5,92	16	0	0	1
2015	1	8,04	17	0	0	0
	2	5,27	18	1	0	0
	3	5,51	19	0	1	0
	4	6,04	20	0	0	1
2016	1	8,26	21	0	0	0
	2	5,51	22	1	0	0
	3	5,41	23	0	1	0
	4	5,83	24	0	0	1

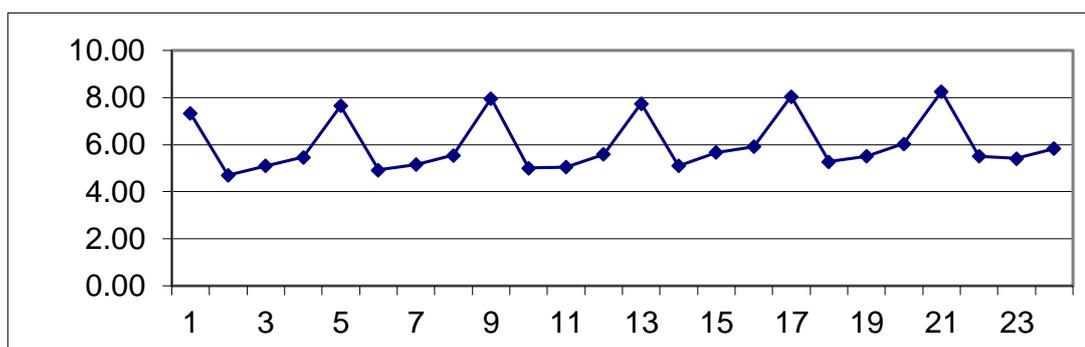


Рис. 4.1 График изменения расходов потребителей на газ и электричество

Учитывая отличие в объеме потребления электроэнергии в отдельные периоды времени (по кварталам), для описания этих отличий введем фиктивные переменные. Здесь качественный признак – время года – имеет 4 градации

(квартал 1, 2, 3, 3), поэтому введем 3 фиктивные переменные (на единицу меньше, чем число градаций признака).

Возьмем, например, первый квартал в качестве базового (эталонного) и будем использовать фиктивные переменные для оценки различий между ним и другими кварталами. Запишем модель в следующем виде:

$$Y = b_0 + b_1t + a_2Z_2 + a_3Z_3 + a_4Z_4 + \varepsilon$$

где

$$Z_2 = \begin{cases} 1 - \text{для второго квартала} \\ 0 - \text{для остальных кварталов} \end{cases}$$

$$Z_3 = \begin{cases} 1 - \text{для третьего квартала} \\ 0 - \text{для остальных кварталов} \end{cases}$$

$$Z_4 = \begin{cases} 1 - \text{для четвертого квартала} \\ 0 - \text{для остальных кварталов} \end{cases}$$

Коэффициенты a_2 , a_3 , a_4 дают численную оценку эффекта, вызываемого сменой сезонов. Коэффициент a_2 показывает дополнительное потребление газа и электроэнергии во 2-м квартале по сравнению с первым кварталом, связанное с изменением времени года. Аналогично, a_3 и a_4 показывают соответствующие дополнительные количества в 3 и 4 кварталах относительно 1-го квартала (базового).

Оценив регрессионную зависимость расходов от времени и фиктивных переменных методом наименьших квадратов (с помощью функции ЛИНЕЙН), получаем:

	-2,1931	-2,57818	-2,77659	0,031589	7,482518
	0,084642	0,084087	0,083752	0,004329	0,075929
	0,986654	0,144869	#Н/Д	#Н/Д	#Н/Д
	351,1631	19	#Н/Д	#Н/Д	#Н/Д
	29,47954	0,398754	#Н/Д	#Н/Д	#Н/Д
t	-25,9102	-30,6608	-33,1524	7,297479	98,54636

$$\hat{Y} = 7,482 + 0,0316t - 2,777Z_2 - 2,578Z_3 - 2,193Z_4 \quad (4.1)$$

(t) (98,5) (7,3) (-33,2) (-30,7) (-25,9)

$$R^2 = 0,987, \quad F = 351.$$

Все t -тесты, относящиеся к коэффициентам при фиктивных переменных показывают высокую значимость, как и F -тест для их совместной объясняющей способности. Из этого результата выводим отдельные уравнения, характеризующие динамику уровней ряда, для каждого квартала:

$$\left. \begin{aligned} \hat{Y} &= 7,482 + 0,0316t && - \text{1-й квартал,} \\ \hat{Y} &= 4,705 + 0,0316t && - \text{2-й квартал,} \\ \hat{Y} &= 4,904 + 0,0316t && - \text{3-й квартал,} \\ \hat{Y} &= 5,289 + 0,0316t && - \text{4-й квартал.} \end{aligned} \right\} \begin{aligned} & (7,482 - 2,777 = 4,705) \\ & (7,482 - 2,578 = 4,904) \\ & (7,482 - 2,193 = 5,289) \end{aligned} \quad (4.2)$$

Результаты моделирования представлены на рис. 4.2.

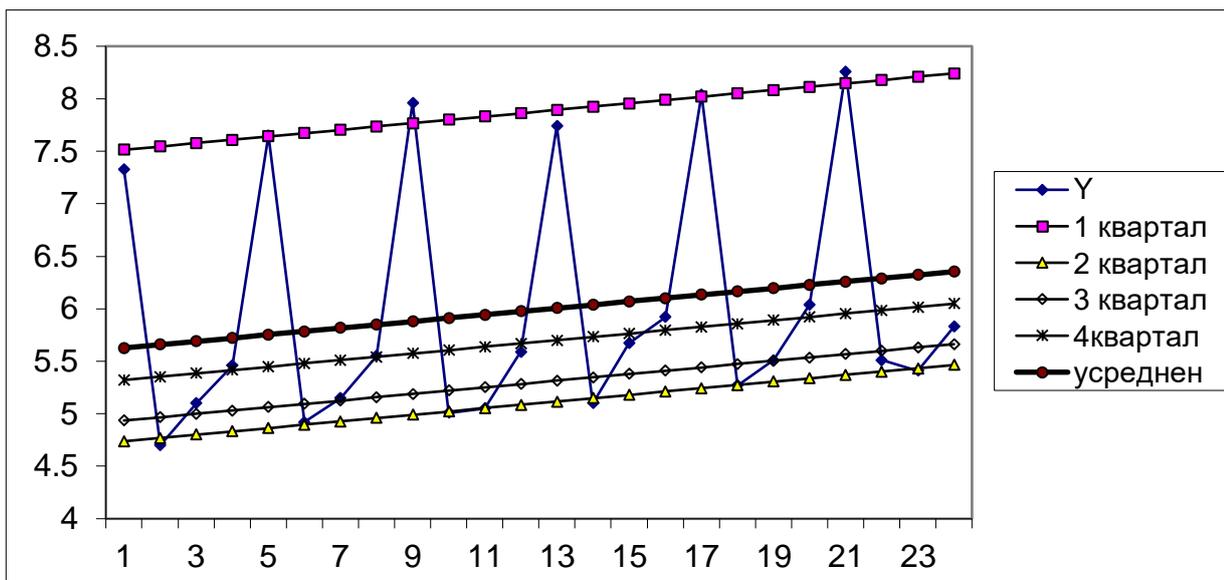


Рис. 4.2. Результаты моделирования

При необходимости можно использовать оцененную регрессию для получения оценки сезонных колебаний в каждом квартале. Выражение (4.2) дает 4 отдельные линии регрессии. Усредняя их, получаем:

$$\frac{7,482 + 4,705 + 4,904 + 5,289}{4} = 5,595,$$

или $\hat{Y} = 5,595 + 0,0316t$

Расстояние между отдельной линией регрессии для любого квартала и усредненной линией (см. рис. 4.2), представленное разностью значений свободного члена в уравнении регрессии, дает оценку сезонных отклонений в этом квартале, которая составила:

$$\begin{aligned} \text{для 1 квартала} & \quad 7,482 - 5,595 = 1,887, \\ \text{для 2 квартала} & \quad 4,705 - 5,595 = -0,89, \\ \text{для 3 квартала} & \quad 4,904 - 5,595 = -0,691, \\ \text{для 4 квартала} & \quad 5,289 - 5,595 = -0,306. \end{aligned}$$

(Проверка: сумма сезонных отклонений должна равняться нулю, и в данном случае это правило выполняется).

Пример 2 [11]. В табл. 4.3 приведена информация о совокупном располагаемом доходе DPI (X) и объемах продаж SALES (Y) лыжного инвентаря в США (квартальные данные за период с 1964 по 1972 г.; DPI – в млрд. долл., SALES – в млн. долл., в ценах 1972 г.

Таблица 4.3 – Совокупный располагаемый доход DPI и объемы продаж SALES лыжного инвентаря в США (в ценах 1972 г.)

Год, квартал	Доход (X)	Объем продаж (Y)	Год, квартал	Доход (X)	Объем продаж (Y)
1964 1	157,4034	53,43051	1968 3	184,8915	47,97185
2	165,5068	48,21284	4	185,3775	55,73602
3	161,5369	44,02953	1969 1	185,6923	54,49402
4	165,2709	53,99799	2	186,6742	49,77978
1965 1	166,4503	52,75629	3	188,8089	51,92244
2	168,7129	44,42772	4	189,7213	55,98525
3	170,4	47,48852	1970 1	190,0968	56,91452
4	172,0651	52,8684	2	192,7143	49,4746
1966 1	172,5916	54,791	3	195,0632	46,63579
2	173,6561	47,48408	4	193,6364	56,43117
3	175,2532	47,04873	1971 1	196,1125	56,65473
4	175,725	53,51625	2	196,9447	49,45025
1967 1	175,9565	55,43292	3	197,6725	49,78809
2	178,6025	45,90745	4	197,3382	57,025
3	181,4444	51,01481	1972 1	197,0121	58,8431
4	181,8293	56,75671	2	197,6394	50,12981
1968 1	182,3892	56,37485	3	196,304	50,3677
2	185,2722	50,91834	4	196	58,8

Задание.

1. Постройте модель зависимости объема продаж от дохода с помощью фиктивных переменных.

2. Для более обоснованных выводов о возможном различии результатов оценивания примените тест Чоу.

Решение

С помощью МАСТЕРА ДИАГРАММ в Excel построим график (рис. 4.3), который показывает динамику изменения совокупного располагаемого дохода DPI (X) и объемов продаж SALES (Y) лыжного инвентаря в США (квартальные данные за период с 1964 по 1972 г.; DPI — в млрд долл., SALES — в млн долл., в ценах 1972 г., см. табл. 4.3).

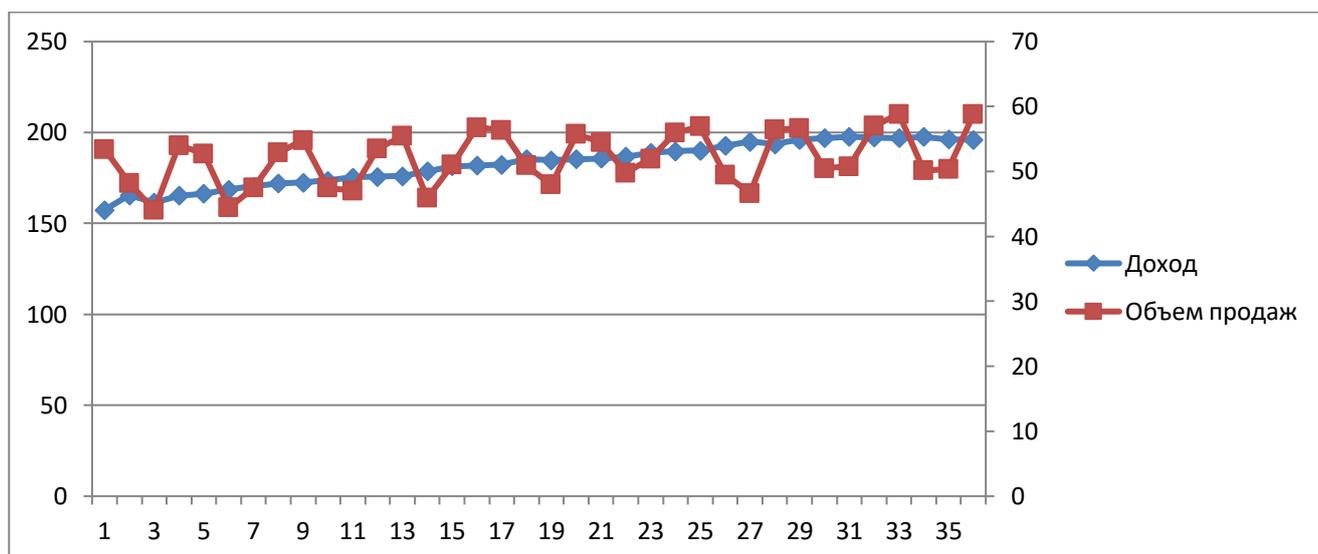


Рис.4.3. Динамика изменения совокупного располагаемого дохода и продаж

Оценивание линейной модели

$$Y_t = \beta_0 + \beta_1 X_t + \varepsilon_t, \quad t = 1, \dots, 36$$

с помощью МНК дает следующие результаты:

0,11978	30,19365	
0,05600	10,25282	
0,11861	3,93296	
4,57528	34	
70,77104	525,91681	
2,13899	2,94491	t

$$\hat{Y} = 30,19365 + 0,11978 X_t(4.1)$$

Оценка коэффициента при переменной X в (4.1) статистически значима, о чем свидетельствует $t = 2,139 > t_{5\%, v=34} = 2,032$. Но коэффициент детерминации, равный величине $0,1186$ говорит о том, что модель только на $11,9\%$ характеризует изменчивость исследуемого показателя. Величина F -статистики Фишера, равная $4,575$, незначительно превышает табличное значение $F_{5\%, 1;34} = 4,13$.

График изменения объема продаж (рис. 4.3) отражает тот факт, что спрос на зимний спортивный инвентарь возрастает в осенне-зимний период и снижается в весенне-летний период года, т.е. имеет *сезонный* характер.

Построенная модель (4.1) не учитывает фактора сезонности спроса и потому оказывается неадекватной, поэтому она не может быть использована для прогнозирования объема спроса в зависимости от величины совокупного располагаемого дохода.

Дополним модель фиктивной переменной (Z), которой придадим значение 1 для I и IV кварталов и 0 – для II и III кварталов. Добавление такой переменной в качестве объясняющей позволяет учесть сезонные колебания спроса.

Оценивание расширенной модели

$$Y_t = \beta_0 + \beta_1 X_t + \alpha Z + \varepsilon_t, \quad t = 1, \dots, 36$$

по данным табл.4.4. дает результаты, представленные в табл. 4.5.

Таблица 4.4 – Исходные данные

Год, квартал	Продажи	Доход	Фиктивная переменная	Продажи (расчетн.)	
	Y	X	Z	\hat{Y}	
1	2	3	4	5	
1964	1	53,43051	157,4034	1	52,3267
	2	48,21284	165,5068	0	46,2318
	3	44,02953	161,5369	0	45,7089
	4	53,99799	165,2709	1	53,3630
1965	1	52,75629	166,4503	1	53,5183
	2	44,42772	168,7129	0	46,6541

Продолжение табл. 4.4.

1	2	3	4	5	
	3	47,48852	170,4	0	46,8763
	4	52,8684	172,0651	1	54,2578
1966	1	54,791	172,5916	1	54,3272
	2	47,48408	173,6561	0	47,3051
	3	47,04873	175,2532	0	47,5155
	4	53,51625	175,725	1	54,7399
1967		55,43292	175,9565	1	54,7703
	2	45,90745	178,6025	0	47,9566
	3	51,01481	181,4444	0	48,3309
	4	56,75671	181,8293	1	55,5438
1968	1	56,37485	182,3892	1	55,6176
	2	50,91834	185,2722	0	48,8351
	3	47,97185	184,8915	0	48,7849
	4	55,73602	185,3775	1	56,0112
1969	1	54,49402	185,6923	1	56,0526
	2	49,77978	186,6742	0	49,0197
	3	51,92244	188,8089	0	49,3009
	4	55,98525	189,7213	1	56,5833
1970	1	56,91452	190,0968	1	56,6327
	2	49,4746	192,7143	0	49,8153
	3	46,63579	195,0632	0	50,1246
	4	56,43117	193,6364	1	57,0989
1971	1	56,65473	196,1125	1	57,4251
	2	49,45025	196,9447	0	50,3724
	3	49,78809	197,6725	0	50,4683
	4	57,025	197,3382	1	57,5865
1972	1	58,8431	197,0121	1	57,5435
	2	50,12981	197,6394	0	50,4639
	3	50,3677	196,304	0	50,2880
	4	58,8	196	1	57,4102

Таблица 4.5 – Результаты расчетов (ЛИНЕЙН)

7,162237595	0,131708	24,43316	
0,467958592	0,019989	3,676299	
0,891166284	1,402808	#Н/Д	
135,1074303	33	#Н/Д	
531,7480985	64,93976	#Н/Д	
15,30528066	6,589093	6,646131	t

Из табл. 4.5 видим, что оцененное на основе данных табл. 4.4 уравнение регрессии имеет вид:

$$\hat{Y}_t = 24,4332 + 0,1317X_t + 7,1622 Z \quad (4.2)$$

Величина t-статистик (см. табл.4.5) свидетельствует о том, что при уровне значимости 5% статистически значимыми можно считать оценки всех параметров (критическое значение $t_{0,1; 33} = 2,034$).

Коэффициент детерминации по сравнению с моделью (4.1) существенно вырос: с 0,119 до 0,89; величина F-статистики Фишера увеличилась с 4,575 в (4.1) до 135,107. Таким образом, введение фиктивной переменной значительно повысило качество модели (теперь модель на 89% объясняет изменчивость исследуемого показателя при высокой степени статистической значимости).

Модель позволяет сделать следующие выводы. Увеличение совокупного располагаемого дохода на 1 млрд. дол. при прочих равных условиях способствует росту объемов продаж лыжного инвентаря в США в среднем на 0,1317млн. дол.; под влиянием всех факторов (кроме цены) ежеквартальный рост объема продаж увеличивается в среднем на 13,089 тыс. руб.

Фактор сезонности играет в рассматриваемом процессе значимую роль. Коэффициенты при фиктивных переменных дают численную оценку эффекта, вызываемого сменой сезонов. Так, значение коэффициента при переменной Z ($a = 7,1622$) свидетельствует о том, что спрос на лыжный инвентарь в I и IV кварталах возрастает по сравнению со спросом во II и III кварталах в среднем примерно на 7,16 млн долл. (в ценах 1972 г.).

Найдем расчетные значения Y (объема продаж), подставив в уравнение (4.2) значения X и Z(гр.5 табл. 4.4).

График (рис. 4.4), построенный по данным гр.5 табл. 4.4, иллюстрирует качество подобранной расширенной модели.



Рис.4.4. Фактический и расчетный объем продаж

По существу, мы получили две различные модели линейной связи между объемом продаж и доходом:

для весенне-летнего периода (при $Z=0$) $\hat{Y}_t = 24,4332 + 0,1317X_t$,

для осенне-зимнего периода (при $Z=1$) $\hat{Y}_t = (24,4332 + 7,1622) + 0,1317X_t$.

При этом (одинаковая для обеих моделей) предельная склонность к закупке лыжного инвентаря оценивается величиной 0,1317.

На основе полученных уравнений можно прогнозировать величину рассматриваемого показателя в зависимости от дохода с учетом сезонных изменений. Так, если совокупный располагаемый доход в следующем году достигнет, например, 200 млн. дол., то получим прогноз объема продаж:

в весенне-летнем периоде

$$\hat{Y}_t = 24,4332 + 0,1317 \cdot 200 = 50,7732 \text{ (млрд. дол.)},$$

в осенне-зимнем периоде

$$\hat{Y}_t = (24,4332 + 7,1622) + 0,1317 \cdot 200 = 57,9354 \text{ (млрд. дол.)}$$

Вместо подбора отдельных моделей для осенне-зимнего и весенне-летнего периодов можно было бы заняться подбором отдельных моделей для каждого из четырех кварталов года. С этой целью в качестве дополнительных объясняющих переменных можно взять, например, переменные Z1, Z2, Z4, принимающие значение 1 соответственно в I, II и IV кварталах и равные 0 в остальных кварталах.

Применим тест Чоу для рассматриваемой задачи. Разобьем исходный массив данных на 2 подвыборки (табл. 4.6) и построим три модели (табл. 4.7).

Таблица 4.6 – Разбиение массива данных на 2 подвыборки

Осень-зима		Весна-лето	
Доход	Объем продаж	Доход	Объем продаж
157,4034	53,43051	165,5068	48,21284
165,2709	53,99799	161,5369	44,02953
166,4503	52,75629	168,7129	44,42772
172,0651	52,8684	170,4	47,48852
172,5916	54,791	173,6561	47,48408
175,725	53,51625	175,2532	47,04873
175,9565	55,43292	178,6025	45,90745
181,8293	56,75671	181,4444	51,01481
182,3892	56,37485	185,2722	50,91834
185,3775	55,73602	184,8915	47,97185
185,6923	54,49402	186,6742	49,77978
189,7213	55,98525	188,8089	51,92244
190,0968	56,91452	192,7143	49,4746
193,6364	56,43117	195,0632	46,63579
196,1125	56,65473	196,9447	50,45025
197,3382	57,025	197,6725	50,78809
197,0121	58,8431	197,6394	50,12981
196	58,8	196,304	50,3677

Таблица 4.7 – Результаты применения функции ЛИНЕЙН

По общей выборке		Подвыборка осень-зима		Подвыборка весна-лето	
0,11978071	30,19365405	0,12774394	32,31795608	0,135896515	23,66601346
0,05599873	10,25282367	0,02000786	3,654326372	0,03563174	6,539669081
0,11860647	3,932955835	0,71813246	1,006406661	0,476199257	1,743780216
4,57527752	34	40,7642518	16	14,54596663	16
70,7710406	525,9168144	41,2882504	16,20566987	44,23093085	48,65231109

В последней строке табл. 4.7 выделены суммы квадратов остатков по рассматриваемым моделям.

Вычислим статистику Чоу

$$F = \frac{S_0 - S_1 - S_2}{S_1 + S_2} \cdot \frac{n_1 + n_2 - 2p - 2}{p + 1},$$

где S_0 , S_1 и S_2 – суммы квадратов остатков для моделей, полученных по общей выборке, 1-й и 2-й подвыборкам, соответственно;

n_1 и n_2 – размер первой и второй подвыборок, соответственно, $n = n_1 + n_2$.

$$F = \frac{525,92 - 16,21 - 48,65}{16,21 + 48,65} \cdot \frac{18 + 18 - 2 \cdot 1 - 2}{1 + 1} = 113,74$$

Табличное значение F-статистики: $F_{\alpha; p+1; n-2p-2} = F_{5\%; 2; 32} = 3,2945$.

Поскольку $F > F_{\text{табл}}$, то с точки зрения улучшения качества модели разбиение на подинтервалы целесообразно.

Таким образом, необходимость построения различных регрессий для осенне-зимнего и весенне-летнего периодов подтверждается тестом Чоу.

Вопросы и задания для самостоятельной работы

1. В каком случае применяются фиктивные переменные?
2. В чем преимущества фиктивных переменных?
3. Как фиктивные переменные включаются в модель регрессии?
4. В чем суть «ловушки» фиктивных переменных?
5. Как сезонные переменные применяются для устранения сезонного фактора?
6. Какова идея теста Чоу?

Задание 1. В таблице 4.8 представлены данные, характеризующие объем продаж (Y_t) некоторого продукта, дополненные ценой этого продукта (X_t).

Таблица 4.8 – Объем продаж и цена продукта

Период	Y_t	X_t	Период	Y_t	X_t
2011г., 1 кв.	239	10	2014г., 1 кв.	481	25,5
2 кв.	201	12	2 кв.	450	22,9
3 кв.	182	12,5	3 кв.	423	23,4
4 кв.	297	13	4 кв.	511	22,8
2012г., 1 кв.	324	15	2015г., 1 кв.	592	23
2 кв.	278	14	2 кв.	498	22,9
3 кв.	257	16	3 кв.	472	24
4 кв.	384	17,3	4 кв.	536	23,8
2013г., 1 кв.	401	18	2017г., 1 кв.	668	28
2 кв.	360	19,5	2 кв.	523	26,5
3 кв.	335	21	3 кв.	545	26,8
4 кв.	462	25	4 кв.	686	27

- 1) Постройте модель зависимости объема продаж от дохода с помощью фиктивных переменных.
- 2) Проверьте качество модели.
- 3) Сделайте прогноз на 2018 г.
- 4) Для более обоснованных выводов о возможном различии результатов оценивания примените тест Чоу.

Задача 2. В табл. 4.9 представлены совокупный объем внутренних инвестиций (Y) и валовой внутренний продукт США (млрд. долл.) за период с 1939 по 1954 гг.

Таблица 4.9 – Объем инвестиций и ВВП США (1939 - 1954 гг.)

Годы	Y	X	Годы	Y	X
1939	9,3	90,8	1947	34,0	232,8
1940	13,1	100,0	1948	45,9	259,1
1941	17,9	124,9	1949	35,3	258,0
1942	9,9	158,3	1950	53,8	286,2
1943	5,8	192,0	1951	59,2	330,2
1944	7,2	210,5	1952	52,1	347,2
1945	10,6	212,3	1953	53,3	366,1
1946	30,7	209,3	1954	52,7	366,3

- 1) Напишите и оцените уравнения, позволяющие ответить на вопрос, изменилась ли зависимость инвестиций от валового внутреннего продукта во время войны (1942 – 1945 гг.) по сравнению с мирным временем.
- 2) Для более обоснованных выводов о возможном различии результатов оценивания примените тест Чоу.

Задание 3. Изучается зависимость уровня годового дохода (переменная $ГД$ в сотнях тыс. руб.) финансового аналитика в зависимости от опыта работы $ОР$ (годы), пола (переменная $Пол$ равная 1 для мужчин и 0 для женщин), владения английским языком (переменная FE равна 1, если аналитик свободно владеет английским языком и 0 в противном случае) и наличия сертификата CFA (переменная CFA равна 1, если сертификат есть). Оцененная модель имеет вид:

$$ГД = 1200 + 200 ОР - 10 ОР^2 + 200 Пол + 400 FE + 150 CFE$$

Все параметры модели значимы на 5%.

- 1) Рассчитайте годовой доход мужчины, свободно владеющего английским, со стажем работы 2 года без сертификата CFA.
- 2) При прочих равных условиях на сколько отличаются годовые доходы мужчин и женщин в отрасли?

Задание 4. По данным для 570 индивидуумов оценили зависимость почасовой заработной платы $EARN$ от длительности обучения S и от способностей индивидуума, описываемых обобщенной переменной $ASVABC$: по общей выборке

$$EARN = -9,96 + 0,93 S + 0,21 ASVABC \qquad RSS_1 = 32189,36$$

$$(2,02) \quad (0,16) \quad (0,04)$$

а также отдельно

$$\text{для мужчин } EARN = -7,23 + 1,01 S + 0,35 ASVABC \quad RSS_1 = 5223,7$$

(2,63) (0,27) (0,06)

$$\text{и женщин } EARN = -11,4 + 0,81 S + 0,14 ASVABC \quad RSS_1 = 10231,24$$

(3,24) (0,19) (0,03)

Можно ли считать, что эта зависимость одинакова для мужчин и женщин?

Задание 5. Оценивалась зависимость расходов на питание в расчете на одного человека от относительного индекса цен на питание и располагаемого дохода:

$$\ln Q = \beta_0 + \beta_1 \ln P + \beta_2 \ln I_n + \varepsilon$$

Были получены следующие результаты:

Оценки	1927-1941г.г. (1)	1948-1962г.г. (2)	Все наблюдения
β_0	4,555	5,052	4,058
β_1	-0,235	-0,237	-0,123
β_2	0,243	0,141	0,242
$RSS \cdot 100$	0,1151	0,0544	0,2866

Можно ли считать зависимость единой для довоенных и послевоенных лет?

ТЕМА 5. ОБОБЩЕННАЯ МОДЕЛЬ ЛИНЕЙНОЙ РЕГРЕССИИ С ГЕТЕРОСКЕДАСТИЧНЫМИ ОСТАТКАМИ (ТЕСТИРОВАНИЕ И УСТРАНЕНИЕ ГЕТЕРОСКЕДАСТИЧНОСТИ)

Данная тема посвящена изучению одной из разновидностей обобщенной модели линейной регрессии (эконометрической модели), которая отличается от классической тем, что возмущающие переменные не имеют постоянной дисперсии.

Цель выполнения лабораторной работы: научиться выявлять нарушения одной из основных предпосылок классической регрессионной модели (существования гетероскедастичности); описывать последствия этих нарушений с точки зрения качества модели; применять методы оценивания параметров модели, позволяющие уменьшить отрицательные последствия, вызванные этими нарушениями.

Для выполнения лабораторной работы № 5 студенты должны

знать:

- содержание основных предпосылок классической регрессии;
- последствия нарушения предпосылок с точки зрения качества построенной методом наименьших квадратов модели;
- основные способы тестирования предпосылок;
- способы устранения негативных последствий нарушения предпосылок;

уметь:

- пользоваться таблицами для проверки статистических гипотез;
- пользоваться пакетом EXCEL:
 - строить таблицы;
 - составлять и копировать формулы;
 - пользоваться МАСТЕРОМ ФУНКЦИЙ: статистическими и математическими функциями.

Задание для выполнения лабораторной работы № 5

Для выборочных данных (табл. 5.1) постройте эконометрическую модель, в которой зависимой переменной будет фактор Y , а независимой – X . Для этого необходимо:

- 1) методом наименьших квадратов (МНК) вычислить оценки параметров модели, проверить статистическую надежность оценок и модели в целом;
- 2) проанализировать целесообразность применения МНК, проверив регрессионные остатки на гетероскедастичность с применением различных тестов;
- 3) рассчитать оценки параметров модели с учетом проведенного тестирования.

Таблица 5.1 – Исходные данные [20]

Вариант 1			Вариант 2			Вариант 3		
№ п/п	Прибыль банка, Y	Вложения в ЦБ, X	№ п/п	Прибыль предприятия, Y	Кредиторская задолженность, X	№ п/п	Затраты на охр. внешней среды, Y	Доход предприятия, X
1	53,78	620,81	1	45,0	127,1	1	5,3	220,5
2	20,00	456,93	2	57,7	133,2	2	5,5	229,7
3	19,63	75,35	3	67,3	138,7	3	6,2	242,1
4	12,91	420,16	4	71,2	138,3	4	6,0	251,0
5	8,76	36,46	5	62,9	150,4	5	6,4	260,3
6	7,24	106,38	6	73,9	156,4	6	7,0	271,0
7	7,13	132,74	7	76,0	162,5	7	7,0	281,3
8	5,97	0,01	8	84,5	169,8	8	7,9	294,9
9	5,50	56,77	9	73,8	165,7	9	7,9	289,8
10	4,76	55,69	10	77,3	170,0	10	8,5	288,7
11	4,69	14,35	11	79,0	175,1	11	7,9	294,0
12	4,60	4,63	12	99,8	185,8	12	8,3	320,8
13	6,48	0,36	13	61,8	182,9	13	7,0	308,2
14	4,04	17,26	14	73,9	187,5	14	7,1	316,4
15	3,38	23,66	15	83,4	196,4	15	7,9	330,2
16	2,79	14,68	16	85,4	197,9	16	7,7	340,9
17	2,58	16,43	17	81,8	211,2	17	8,2	351,6
18	2,35	45,52	18	90,2	219,0	18	8,0	369,5
19	2,08	30,00	19	93,2	219,7	19	9,0	373,7
20	1,92	22,06	20	100,4	221,7	20	9,3	382,8
21	1,82	50,77	21	92,4	228,5	21	8,6	396,4
22	1,42	10,39	22	10,2	235,0	22	9,7	407,5
23	0,94	7,61	23	10,0	236,3	23	9,4	415,1
24	3,80	4,98	24	12,2	254,8	24	10,3	455,6
25	6,70	13,57	25	10,4	249,5	25	11,7	436,2
26	0,53	21,02	26	11,9	258,4	26	13,2	450,3
27	0,28	30,41	27	13,7	271,0	27	15,0	473,5
28	0,24	14,80	28	14,8	268,5	28	17,0	483,9
29	0,23	15,60	29	13,6	279,2	29	14,4	488,1
30	0,11	24,58	30	14,4	292,9	30	14,9	512,0

Вариант 4			Вариант 5			Вариант 6		
№ п/п	Потребление семьи, Y	Доход семьи, X	№ п/п	Прибыль банка Y	Среднегодовые остатки на текущих счетах X	№ п/п	Среднедушевые затраты на ср. образование региона, Y	Среднедушевой доход в регионе, X
1	325	330	1	9,8	20,4	1	295	6247
2	597	630	2	38,9	70,6	2	275	6183
3	690	870	3	20,9	32	3	531	8914
4	700	890	4	5,1	14,3	4	316	7505
5	950	1020	5	31,3	303,4	5	304	6813
6	980	1090	6	1,7	10	6	431	7873
7	985	1107	7	2,2	9,3	7	326	6640
8	987	1278	8	32,2	59,7	8	427	8063
9	1336	1467	9	2,6	5,5	9	259	5736
10	1300	1647	10	38,1	69,5	10	294	7391
11	1504	1827	11	8,3	21,6	11	423	8818
12	1560	2000	12	28,5	266,1	12	335	6607
13	1632	2200	13	10,8	18,3	13	320	6951
14	1875	2795	14	29,6	47,3	14	342	7526
15	2014	2982	15	10,5	240,3	15	288	6489
16	2400	2847	16	41,6	86,5	16	343	6541
17	2300	3027	17	3,1	8,3	17	320	6456
18	2500	3207	18	29,7	397,8	18	821	10851
19	2590	3398	19	34,2	251,5	19	387	8850
20	2570	3567	20	49,6	541,7	20	424	8604
21	2490	3747	21	51,1	325,8	21	265	6700
22	2657	3927	22	48,8	392,7	22	437	8745
23	2750	4107	23	57,6	410,7	23	355	8001
24	2860	4287	24	54,3	428,7	24	322	6333
25	2870	5467	25	50,4	546,7	25	466	8442
26	2900	5647	26	55,3	564,7	26	274	7342
27	3100	5827	27	84,9	682,7	27	359	9032
28	3115	6007	28	62,6	600,8	28	388	6505
29	3250	6687	29	77,5	668,8	29	311	7478
30	3300	7220	30	93,3	722,1	30	397	7839

Вариант 7			Вариант 8			Вариант 9		
№ п/п	Капитало вложения в регион, Y	Промышленная продукция (среднедуш.), X	№ п/п	Прибыль банка Y	Объем выданных кредитов, X	№ п/п	Прибыль банка Y	Высоколиквидные активы банка X
1	161	526	1	83,7	26,3	1	28,4	243,6
2	117	791	2	78,7	26,2	2	27,7	241,6
3	98	376	3	67,7	21,5	3	26,8	209,5
4	330	2712	4	65,4	21,3	4	19,3	153,2
5	254	2508	5	64,8	21,4	5	14,9	106,3
6	135	661	6	64,4	20,0	6	14,7	94,2
7	94	250	7	48,3	18,0	7	7,0	62,0
8	299	2795	8	32,3	17,1	8	6,8	52,7
9	152	992	9	22,3	16,5	9	6,7	41,2
10	280	1247	10	20,6	15,6	10	5,4	40,5
11	146	504	11	20,4	15,8	11	5,0	40,1
12	224	1607	12	19,6	14,9	12	7,9	39,9
13	196	751	13	18,1	14,1	13	4,6	28,5
14	163	1187	14	11,7	13,5	14	3,6	27,9
15	158	748	15	8,5	12,6	15	3,0	27,0

16	391	1766	16	8,1	12,3	16	2,7	25,2
17	176	1133	17	7,6	11,2	17	2,0	24,8
18	206	1241	18	7,3	11,0	18	1,8	20,1
19	127	476	19	7,1	10,6	19	1,7	19,6
20	215	1034	20	6,7	10,0	20	1,3	17,5
21	107	489	21	6,6	10,0	21	1,2	15,8
22	190	854	22	6,5	8,7	22	1,1	14,9
23	118	832	23	6,2	8,2	23	0,6	14,7
24	112	402	24	6,0	7,9	24	0,6	14,6
25	146	823	25	3,1	6,6	25	0,6	12,1
26	141	773	26	3,1	5,5	26	0,6	11,6
27	221	684	27	1,7	5,2	27	0,5	9,1
28	217	662	28	1,3	5,5	28	0,3	8,8
29	315	845	29	1,4	5,1	29	0,2	5,6
30	200	909	30	1,2	5,3	30	0,1	3,1

Вариант 10			Вариант 11			Вариант 12		
№ п/п	Прибыль банка, Y	Величина КИП банка, X	№ п/п	Прибыль предприятия, Y	Задолженность перед юр. и физ. лицами X	№ п/п	Рыночная цена акции Y	Прибыль АО X
1	37,7	754	1	37,2	2173,6	1	0,315	740
2	57,0	851	2	26,1	1386,6	2	0,290	850
3	31,2	194	3	22,0	1383,6	3	0,261	930
4	44,3	617	4	26,1	886,1	4	0,058	111
5	41,7	152	5	15,6	875,5	5	0,048	133
6	12,1	66	6	16,6	586,8	6	0,041	106
7	31,4	125	7	8,2	285,4	7	0,032	77
8	1,9	36	8	10,9	243,9	8	0,058	146
9	4,9	46	9	5,7	133,4	9	0,067	168
10	6,3	54	10	10,6	235,0	10	0,072	184
11	15,9	70	11	7,9	214,2	11	0,158	351
12	3,5	40	12	5,5	158,7	12	0,205	559
13	7,5	58	13	5,9	250,1	13	0,620	1234
14	13,4	67	14	6,1	194,3	14	0,548	6555
15	1,1	26	15	7,0	195,3	15	0,761	8224
16	5,3	44	16	0,8	85,8	16	0,800	8672
17	5,5	47	17	7,1	258,7	17	0,947	9733
18	3,1	29	18	3,6	108,0	18	1,058	12450
19	2,8	40	19	2,4	96,0	19	1,245	15736
20	8,1	58	20	2,8	99,2	20	1,584	21437
21	24,0	135	21	6,6	122,8	21	1,420	17445
22	1,6	32	22	5,8	138,2	22	1,711	23111
23	3,6	52	23	0,4	64,2	23	1,945	28673
24	3,8	46	24	8,4	108,6	24	2,503	35204
25	4,7	37	25	1,7	92,7	25	2,689	32668
26	1,0	24	26	6,0	132,9	26	3,007	39577
27	13,5	67	27	3,1	104,0	27	3,100	40010
28	0,4	10	28	1,8	95,1	28	3,880	43995
29	1,3	22	29	1,3	90,5	29	4,300	45742
30	3,2	43	30	1,1	87,2	30	4,100	43588

Вариант 13			Вариант 14			Вариант 15		
№ п/п	Затраты предприятия на охр. труда, Y	Объем выпуска продукции, X	№ п/п	Затраты предприятия на инновац. продукцию Y	Объем реализованной продукции X	№ п/п	Сумма выплаченных налогов, Y	Доход предприятия, X
1	7,5	289,0	1	1,1	25,3	1	1576	2927
2	7,8	299,3	2	2,1	78,0	2	876	3639
3	8,6	313,2	3	1,2	35,5	3	302	855
4	8,4	323,1	4	0,7	16,0	4	327	1217
5	8,8	333,5	5	4,2	487,0	5	283	983
6	9,6	345,5	6	0,5	11,3	6	1017	5640
7	9,6	357,1	7	0,5	10,5	7	302	1125
8	10,7	372,3	8	1,5	66,0	8	180	683
9	10,7	366,6	9	0,3	6,4	9	680	3554
10	11,4	365,3	10	1,9	76,8	10	161	652
11	10,7	371,3	11	0,9	24,1	11	360	1967
12	11,1	401,3	12	4,0	293,0	12	958	3755
13	9,6	387,2	13	1,0	20,4	13	377	2231
14	9,7	396,4	14	1,1	52,3	14	103	587
15	10,7	411,9	15	1,8	264,6	15	74	193
16	10,4	423,8	16	2,9	95,5	16	63	349
17	11,0	435,8	17	0,4	9,4	17	60	256
18	10,8	455,8	18	12,7	638,1	18	60	125
19	12,0	460,5	19	3,4	277,0	19	157	693
20	12,4	470,7	20	13,7	868,3	20	373	1937
21	11,5	486,0	21	8,0	8,0	21	152	733
22	12,8	498,4	22	10,6	10,6	22	81	284
23	12,5	506,9	23	13,0	13,0	23	243	849
24	13,6	552,3	24	12,4	12,4	24	66	256
25	15,3	530,5	25	13,8	13,8	25	78	257
26	17,1	546,3	26	15,6	15,6	26	68	357
27	19,3	572,3	27	18,0	18,0	27	73	457
28	21,7	584,0	28	16,7	16,7	28	57	137
29	18,5	588,7	29	15,9	15,9	29	153	811
30	19,1	615,4	30	19,7	19,7	30	349	1569

Вариант 16			Вариант 17			Вариант 18		
№ п/п	Прибыль банка, Y	Вложения в ЦБ, X	№ п/п	Прибыль предприятия, Y	Кредиторская задолженность, X	№ п/п	Затраты на охр. внешней среды, Y	Доход предприятия, X
1	10,30	21,38	1	59,5	66,9	1	7,2	295,4
2	8,40	21,71	2	99,1	109,6	2	7,4	306,4
3	5,85	21,83	3	85,7	78,1	3	8,2	321,3
4	5,84	22,63	4	81,8	78,4	4	8,0	332,0
5	8,16	23,46	5	88,5	127,3	5	8,5	343,2
6	9,65	24,29	6	88,5	96,2	6	9,2	356,0
7	6,14	28,05	7	72,3	73,0	7	9,2	368,0
8	7,53	29,09	8	88,4	105,5	8	10,3	384,6

9	8,99	30,69	9	77,5	90,2	9	10,3	378,5
10	5,72	31,61	10	91,9	109,8	10	11,6	377,2
11	10,58	7,04	11	93,6	114,9	11	10,3	383,5
12	11,41	7,39	12	76,3	122,7	12	11,3	415,8
13	10,21	11,66	13	114,4	125,6	13	9,2	400,6
14	9,81	12,01	14	90,6	102,3	14	9,3	410,5
15	7,55	14,64	15	98,0	136,2	15	10,2	427,1
16	7,03	17,42	16	99,9	137,6	16	10,1	439,9
17	6,31	20,60	17	96,4	151,0	17	11,2	506,3
18	7,69	37,03	18	94,7	158,8	18	11,0	533,2
19	5,89	37,44	19	107,0	168,3	19	12,3	539,5
20	14,37	43,49	20	101,9	161,5	20	12,9	553,0
21	7,96	52,55	21	117,5	189,3	21	11,8	573,5
22	7,43	57,80	22	106,4	174,8	22	13,4	590,2
23	10,37	62,72	23	130,2	176,1	23	13,0	601,5
24	11,11	63,80	24	117,5	159,5	24	14,4	662,2
25	25,24	82,38	25	222,6	194,6	25	16,5	633,2
26	12,85	110,05	26	235,1	198,1	26	18,7	654,4
27	12,74	136,41	27	328,4	219,0	27	21,3	689,1
28	18,52	423,83	28	276,0	210,9	28	24,4	704,7
29	25,61	460,69	29	109,9	208,3	29	20,5	711,1
30	59,39	624,48	30	562,2	232,7	30	21,2	745,8

Вариант 19			Вариант 20			Вариант 21		
№ п/п	Потребление семьи, Y	Доход семьи, X	№ п/п	Прибыль банка Y	Среднегодовые остатки на текущих счетах X	№ п/п	Среднедушевые затраты на ср.образование региона, Y	Среднедушевой доход в регионе, X
1	341	372	1	11,0	22,5	1	320	6951
2	423	510	2	42,9	77,8	2	342	7526
3	519	668	3	23,1	35,3	3	314	6489
4	530	690	4	6,0	15,8	4	343	6541
5	760	855	5	34,5	333,8	5	322	6456
6	1019	1211	6	2,3	11,2	6	821	10851
7	1024	1230	7	2,9	10,5	7	377	8850
8	1027	1419	8	35,5	65,8	8	468	8604
9	1823	1628	9	3,3	6,4	9	265	6700
10	1774	2082	10	42,0	76,6	10	457	8745
11	2049	2307	11	9,4	23,9	11	325	8001
12	2125	2523	12	31,5	292,8	12	305	6333
13	2222	2773	13	12,1	20,1	13	476	8442
14	2550	3517	14	32,7	52,1	14	294	7342
15	2738	3751	15	11,7	264,4	15	349	9032
16	3259	3582	16	45,9	95,3	16	388	6505
17	3127	3807	17	3,8	9,4	17	311	7478
18	3394	4032	18	32,8	437,7	18	397	7839
19	3516	4271	19	37,7	276,8	19	295	6242

20	3489	4482	20	54,7	596,0	20	315	7697
21	3381	4707	21	56,3	358,5	21	356	7624
22	3606	4932	22	53,8	432,1	22	339	7374
23	3732	5157	23	63,5	451,9	23	459	8001
24	3880	5382	24	59,8	471,7	24	428	10022
25	3894	6857	25	55,5	601,5	25	403	8380
26	3934	7082	26	60,9	621,3	26	345	7696
27	4204	7307	27	93,5	751,1	27	280	6615
28	4224	7532	28	69,0	661,0	28	473	8306
29	4107	8382	29	85,4	735,8	29	477	7847
30	4874	9048	30	102,7	794,4	30	423	7051

Вариант 22			Вариант 23			Вариант 24		
№ п/п	Капитало вложения в регион Y	Промышленная продукция (среднедуш.) X	№ п/п	Прибыль банка Y	Объем выданных кредитов, X	№ п/п	Прибыль банка Y	Высоко ликвидные активы банка X
1	157	510	1	92,9	29,7	1	30,5	263,9
2	109	770	2	87,4	29,6	2	29,8	231,7
3	103	363	3	75,3	24,5	3	28,8	227,1
4	520	3401	4	72,7	24,2	4	20,9	166,3
5	325	3146	5	72,1	24,3	5	16,2	115,6
6	136	643	6	71,5	22,8	6	16,0	102,5
7	88	240	7	53,9	20,3	7	7,8	67,8
8	588	3505	8	36,3	19,6	8	7,6	57,7
9	152	967	9	21,5	19,0	9	7,5	45,3
10	299	1217	10	20,0	18,1	10	6,1	44,5
11	121	489	11	19,7	18,2	11	5,4	44,1
12	211	2020	12	18,9	14,4	12	5,2	43,9
13	184	731	13	17,5	13,6	13	4,9	31,6
14	161	1158	14	11,3	13,1	14	3,9	30,9
15	178	728	15	8,3	12,2	15	3,3	30,0
16	624	2219	16	7,9	11,9	16	3,0	28,0
17	179	1105	17	7,4	10,9	17	2,3	26,7
18	194	1211	18	7,1	10,7	18	2,1	21,8
19	112	461	19	6,9	10,4	19	2,0	21,3
20	202	1103	20	6,5	9,7	20	1,5	9,1
21	105	474	21	6,4	9,7	21	1,4	17,3
22	179	832	22	6,3	8,5	22	1,3	16,3
23	130	810	23	6,0	8,0	23	0,8	16,1
24	95	389	24	5,9	7,7	24	0,8	16
25	137	802	25	3,1	6,4	25	0,8	13,4
26	132	753	26	3,1	5,4	26	0,8	12,9
27	208	665	27	1,7	5,1	27	0,7	10,3
28	179	644	28	1,3	5,4	28	0,5	9,9
29	146	823	29	1,4	5,0	29	0,4	6,6
30	161	886	30	1,3	5,2	30	0,3	4,0

Вариант 25			Вариант 26			Вариант 27		
№ п/п	Прибыль банка, Y	Величина КИП банка, X	№ п/п	Прибыль предприятия, Y	Задолженность перед юр. и физ. лицами X	№ п/п	Рыночная цена акции Y	Прибыль АО X
1	72,8	808,0	1	40,6	2326,5	1	0,393	871
2	109,5	911,8	2	28,7	1484,4	2	0,365	994
3	60,5	208,8	3	24,3	1481,5	3	0,332	1084
4	85,4	661,4	4	28,7	948,8	4	0,105	166
5	80,4	163,8	5	17,5	937,5	5	0,094	191
6	24,2	71,8	6	18,6	628,6	6	0,086	161
7	60,9	135,0	7	8,8	306,1	7	0,076	128
8	3,6	37,8	8	12,5	261,7	8	0,105	206
9	7,5	48,0	9	6,2	143,4	9	0,115	230
10	13,2	59,0	10	12,1	252,2	10	0,121	248
11	31,4	76,1	11	8,5	229,9	11	0,217	435
12	5,7	41,9	12	6,0	170,5	12	0,270	668
13	15,5	63,3	13	6,4	268,3	13	0,734	1424
14	26,7	72,9	14	6,6	208,6	14	0,654	7384
15	2,5	27,6	15	7,5	209,7	15	0,892	11056
16	11,3	46,0	16	1,2	87,8	16	0,936	11598
17	11,7	49,0	17	7,6	277,5	17	1,101	12882
18	5,1	30,7	18	4,1	110,5	18	2,385	16170
19	4,7	41,9	19	2,8	98,2	19	2,611	20146
20	16,6	63,3	20	3,3	101,5	20	3,022	27044
21	46,8	145,7	21	7,1	132,1	21	2,823	22213
22	3,2	33,7	22	6,3	148,6	22	3,175	29069
23	5,8	56,8	23	0,8	65,8	23	3,458	38799
24	6,0	48,0	24	9,0	111,1	24	4,134	43702
25	7,2	38,8	25	2,1	94,9	25	4,359	40633
26	2,4	25,6	26	6,5	142,9	26	4,743	48993
27	26,9	72,9	27	3,6	106,4	27	4,856	49517
28	1,6	11,3	28	2,2	97,3	28	5,800	54339
29	2,8	23,5	29	1,7	92,6	29	6,308	56453
30	5,3	45,0	30	1,5	89,2	30	6,066	53846

Вариант 28			Вариант 29			Вариант 30		
№ п/п	Затраты предприятия на охрану труда, Y	Объем выпуска продукции, X	№ п/п	Затраты предприятия на инновац. продукцию Y	Объем реализованной продукции X	№ п/п	Сумма выплаченных налогов, Y	Доход предприятия, X
1	6,0	242,9	1	2,5	50,7	1	355	2342
2	6,2	253,0	2	4,6	156,0	2	614	2912
3	7,0	266,6	3	2,7	71,1	3	212	642
4	6,8	276,4	4	1,5	32,2	4	229	913
5	7,2	286,6	5	8,9	1120,4	5	199	738
6	7,9	298,4	6	1,2	22,7	6	712	4512
7	7,0	309,7	7	1,1	21,2	7	212	844

8	8,8	324,7	8	3,2	132,0	8	118	513
9	8,8	319,1	9	0,8	12,8	9	477	2844
10	9,5	317,9	10	4,1	153,6	10	105	490
11	8,8	323,7	11	2,0	48,2	11	253	1574
12	9,3	514,9	12	8,4	674,1	12	671	3004
13	7,9	494,7	13	2,2	41,0	13	264	1785
14	8,0	507,8	14	2,3	104,8	14	68	441
15	8,8	530,1	15	3,8	608,8	15	49	145
16	8,6	547,0	16	6,2	191,0	16	42	262
17	9,2	564,2	17	1,0	19,0	17	40	193
18	8,9	592,8	18	31,1	1467,8	18	40	94
19	10,0	599,5	19	7,2	637,2	19	103	520
20	10,3	614,1	20	33,6	1997,3	20	262	1550
21	9,6	635,8	21	19,7	1202,8	21	100	550
22	10,8	653,6	22	26,1	1449,0	22	53	214
23	10,5	665,8	23	31,7	1515,3	23	171	637
24	17,3	730,6	24	30,4	1581,5	24	44	193
25	19,4	699,5	25	33,6	2015,7	25	51	193
26	21,7	722,1	26	37,9	2082,0	26	45	268
27	24,5	759,2	27	43,7	2516,2	27	48	343
28	27,5	775,8	28	40,6	2214,8	28	38	103
29	23,5	782,6	29	38,7	2465,1	29	100	609
30	24,3	820,8	30	47,9	2661,2	30	245	1256

Для выполнения работы следует изучить теоретический материал, представленный в темах 6, 8 Учебного пособия по эконометрике [18].

РЕШЕНИЕ ТИПОВЫХ ЗАДАЧ

Пример 1[4]. *Тестирование гетероскедастичности*

В табл. 5.2 приведена информация о величине добавленной стоимости в обрабатывающей промышленности (ДС) и валового внутреннего продукта (ВВП) в различных странах. Выборка включает как малые страны (Словакия, Словения), так и большие (Франция, Великобритания, Италия). Выпуск продукции обрабатывающей промышленности обычно составляет 15-52% ВВП. Очевидно, что при большом ВВП относительное его изменение на 1% в абсолютном значении будет выражаться значительно большими цифрами, чем при малом. Поэтому при моделировании зависимости между величиной ДС и ВВП может иметь место гетероскедастичность.

Таблица 5.2 – Добавленная стоимость в обрабатывающей промышленности и ВВП (млн.долл.США) для межстрановой выборки, 1994

Страна	ДС (Y)	ВВП (X)	Страна	ДС (Y)	ВВП (X)
Бельгия	44517	232006	Малайзия	18874	72505
Канада	112617	547203	Мексика	55073	420788
Чили	13096	50919	Нидерланды	48595	334286
Дания	25927	151266	Норвегия	13484	122926
Финляндия	21581	97624	Португалия	17025	87352
Франция	256316	1330998	Сингапур	20648	71039
Греция	9392	98861	Словакия	2720	13746
Гонконг	11758	130823	Словения	4520	14386
Венгрия	7227	41506	Испания	80104	483652
Ирландия	17572	52662	Швеция	34806	198432
Израиль	11349	74121	Швейцария	57503	261388
Италия	145013	1016286	Сирия	3317	44753
Южная Корея	161318	380820	Турция	31115	135961
Кувейт	2797	24848	Великобритания	244397	1024609

На рис. 5.1 представлена диаграмма рассеяния, демонстрирующая различие в отклонениях точек с координатами ВВП и ДС, соответствующих разным странам.

По данным, приведенным в табл. 5.2, с помощью МНК оценим зависимость ДС от ВВП (рис. 5.2):

$$\hat{Y} = 603,9 + 0,194X. \quad (5.1)$$

Из полученного уравнения видим, что в каждом дополнительном млн. долл. ВВП выпуск продукции обрабатывающей промышленности составляет в среднем 194 тыс. долл. Коэффициент при X, равный 0,194, является статистически значимой характеристикой соответствующего истинного значения параметра, о чем свидетельствует t-статистика, равная 14,52.

На основе модели (5.1) найдем расчетные (теоретические) значения ДС (\hat{Y}) и остатки, как разность фактических и расчетных значений Y (рис. 5.2).

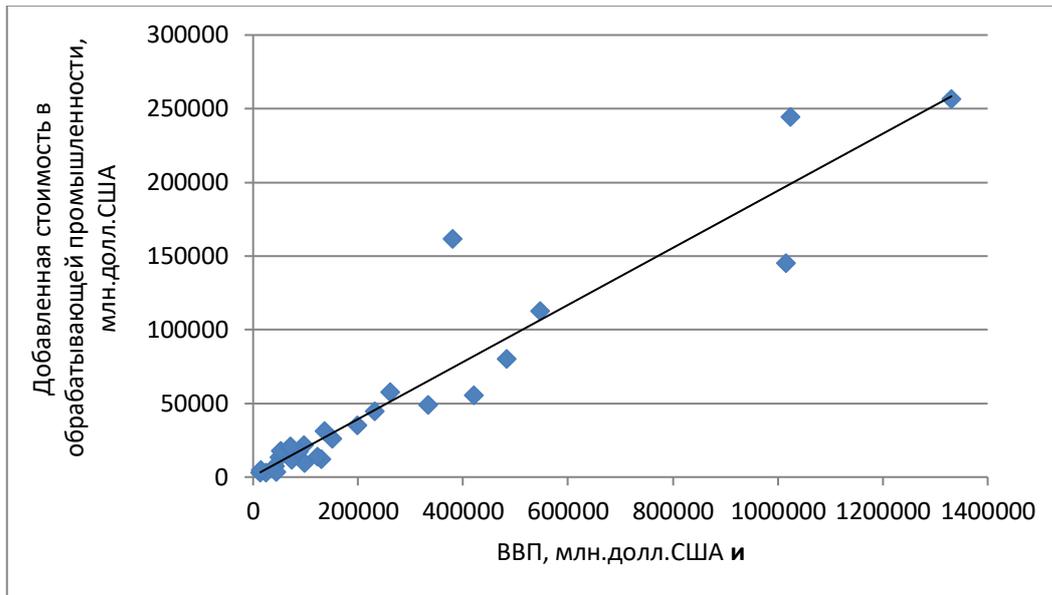


Рис.5.1. Выпуск продукции обрабатывающей промышленности и ВВП

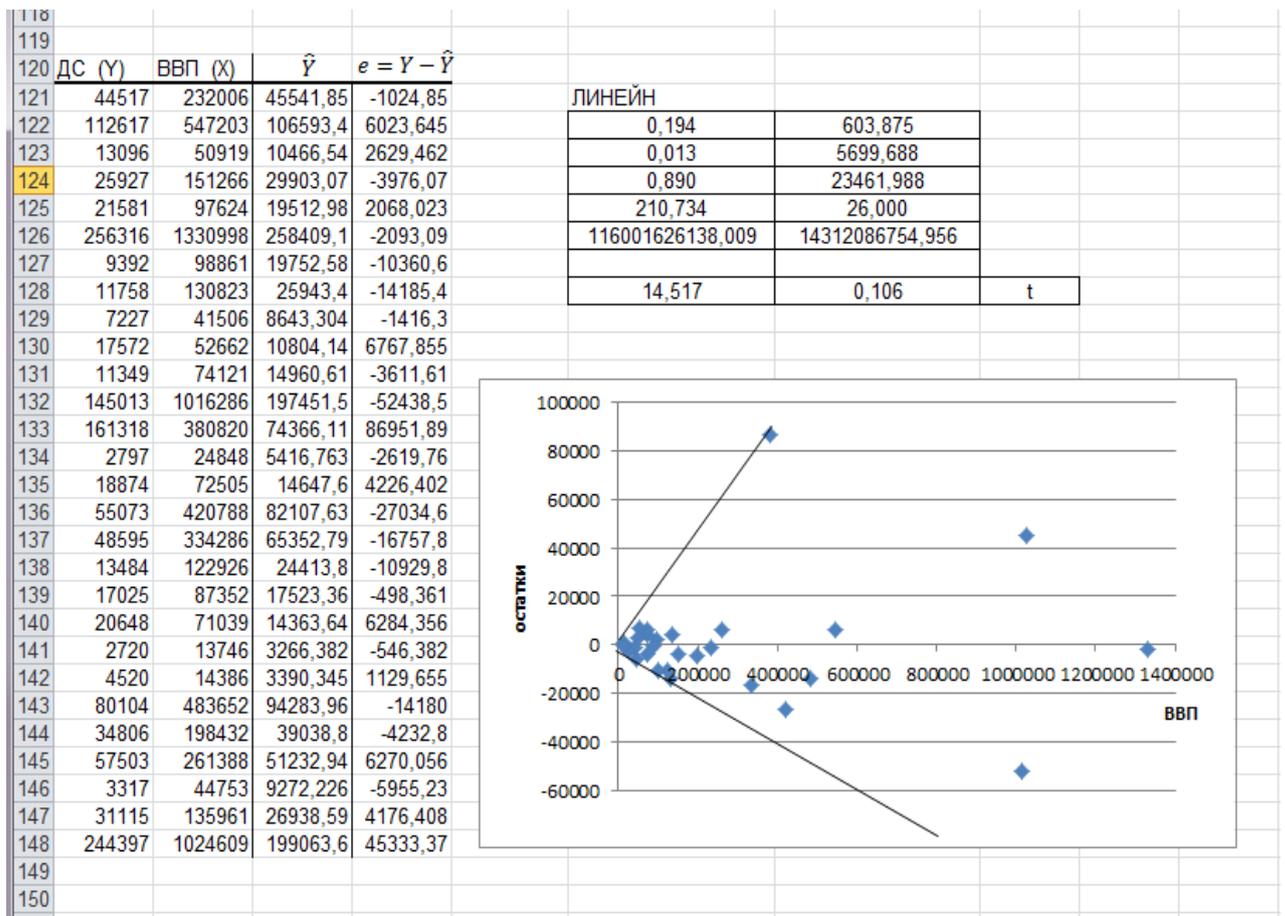


Рис.5.2. Результаты расчетов

Выполним проверку гетероскедастичности с помощью различных критериев.

1. Графический анализ остатков

На рис. 5.1 и 5.2 хорошо видна неоднородность выборочной совокупности и зависимость остатков от изменения независимой переменной (ВВП).

2. Тест ранговой корреляции Спирмена.

Прежде всего, определим ранги значений ВВП (столбца X) и ранги модулей остатков $|e_i|$. Для этого воспользуемся функцией РАНГ, которая в Excel находится среди Статистических и имеет следующий синтаксис: **РАНГ.СП(число;ссылка;[порядок])**. Эта функция возвращает ранг числа в списке чисел, то есть его величину относительно других значений в списке. Если несколько значений имеют одинаковый ранг, возвращается среднее.

Аргументы функции РАНГ.СП:

- **Число** Обязательный. Число, для которого определяется ранг.
- **Ссылка** Обязательный. Массив чисел или ссылка на список чисел. Нечисловые значения в ссылке игнорируются.
- **Порядок** Необязательный. Число, определяющее способ упорядочения:

если значение аргумента "порядок" равно 0 (нулю) или опущено, ранг числа определяется в Excel так, как если бы ссылка была списком, отсортированным по убыванию.

если значение аргумента "порядок" — любое число, кроме нуля, то ранг числа определяется в Excel так, как если бы ссылка была списком, отсортированным по возрастанию.

Порядок расчета рангов представлен на рис. 5.3 и 5.4.

Найдем коэффициент ранговой корреляции (рис. 5.5):

$$r_{x,e} = 1 - 6 \cdot \frac{\sum d_i^2}{n(n^2 - 1)},$$

где d_i — разность между рангами значений x_i и $|e_i|$ в наблюдении i .

РАНГ.СР								
=РАНГ(B161;\$B\$161:\$B\$188;0)								
	A	B	C	D	E	F	G	H
159								
160	ДС (Y)	ВВП (X)	\hat{Y}	$e = Y - \hat{Y}$	$ e $	Формула для расчета ранга X	Ранг X	
161	44517	232006	45541,85	-1024,85	1024,85	=РАНГ(B161;\$B\$161:\$B\$188;0)	10	
162	112617	547203	106593,4	6023,64	6023,64		4	
163	13096	50919	10466,54	2629,46	2629,46		23	
164	25927	151266	29903,07	-3976,07	3976,07		12	
165	21581	97624	19512,98	2068,02	2068,02		17	
166	256316	1330998	258409,1	-2093,09	2093,09		1	
167	9392	98861	19752,58	-10360,58	10360,58		16	
168	11758	130823	25943,4	-14185,40	14185,40		14	
169	7227	41506	8643,304	-1416,30	1416,30		25	
170	17572	52662	10804,14	6767,86	6767,86		22	
171	11349	74121	14960,61	-3611,61	3611,61		19	
172	145013	1016286	197451,5	-52438,53	52438,53		3	
173	161318	380820	74366,11	86951,89	86951,89		7	
174	2797	24848	5416,763	-2619,76	2619,76		26	
175	18874	72505	14647,6	4226,40	4226,40		20	
176	55073	420788	82107,63	-27034,63	27034,63		6	
177	48595	334286	65352,79	-16757,79	16757,79		8	
178	13484	122926	24413,8	-10929,80	10929,80		15	
179	17025	87352	17523,36	-498,36	498,36		18	
180	20648	71039	14363,64	6284,36	6284,36		21	
181	2720	13746	3266,382	-546,38	546,38		28	
182	4520	14386	3390,345	1129,65	1129,65		27	
183	80104	483652	94283,96	-14179,96	14179,96		5	
184	34806	198432	39038,8	-4232,80	4232,80		11	
185	57503	261388	51232,94	6270,06	6270,06		9	
186	3317	44753	9272,226	-5955,23	5955,23		24	
187	31115	135961	26938,59	4176,41	4176,41		13	
188	244397	1024609	199063,6	45333,37	45333,37		2	
189								

Рис.5.3. Расчет рангов

РАНГ.СР								
=РАНГ(E161;\$E\$161:\$E\$188;0)								
	A	B	C	D	E	F	G	H
159								
160	ДС (Y)	ВВП (X)	\hat{Y}	$e = Y - \hat{Y}$	$ e $	Формула для расчета ранга модуля e	Ранг модуля e	
161	44517	232006	45541,85	-1024,85	1024,85	=РАНГ(E161;\$E\$161:\$E\$188;0)	26	
162	112617	547203	106593,4	6023,64	6023,64		13	
163	13096	50919	10466,54	2629,46	2629,46		20	
164	25927	151266	29903,07	-3976,07	3976,07		18	
165	21581	97624	19512,98	2068,02	2068,02		23	
166	256316	1330998	258409,1	-2093,09	2093,09		22	
167	9392	98861	19752,58	-10360,58	10360,58		9	
168	11758	130823	25943,4	-14185,40	14185,40		6	
169	7227	41506	8643,304	-1416,30	1416,30		24	
170	17572	52662	10804,14	6767,86	6767,86		10	
171	11349	74121	14960,61	-3611,61	3611,61		19	
172	145013	1016286	197451,5	-52438,53	52438,53		2	
173	161318	380820	74366,11	86951,89	86951,89		1	
174	2797	24848	5416,763	-2619,76	2619,76		21	
175	18874	72505	14647,6	4226,40	4226,40		16	
176	55073	420788	82107,63	-27034,63	27034,63		4	
177	48595	334286	65352,79	-16757,79	16757,79		5	
178	13484	122926	24413,8	-10929,80	10929,80		8	
179	17025	87352	17523,36	-498,36	498,36		28	
180	20648	71039	14363,64	6284,36	6284,36		11	
181	2720	13746	3266,382	-546,38	546,38		27	
182	4520	14386	3390,345	1129,65	1129,65		25	
183	80104	483652	94283,96	-14179,96	14179,96		7	
184	34806	198432	39038,8	-4232,80	4232,80		15	
185	57503	261388	51232,94	6270,06	6270,06		12	
186	3317	44753	9272,226	-5955,23	5955,23		14	
187	31115	135961	26938,59	4176,41	4176,41		17	
188	244397	1024609	199063,6	45333,37	45333,37		3	
189								

Рис.5.4. Расчет рангов

	A	B	C	D	E	F	G	H	I
160	ДС (Y)	ВВП (X)	\hat{Y}	$e = Y - \hat{Y}$	$ e $	Ранг X	Ранг модуля e	D	D^2
161	44517	232006	45541,85	-1024,85	1024,85	10	26	-16	256
162	112617	547203	106593,4	6023,64	6023,64	4	13	-9	81
163	13096	50919	10466,54	2629,46	2629,46	23	20	3	9
164	25927	151266	29903,07	-3976,07	3976,07	12	18	-6	36
165	21581	97624	19512,98	2068,02	2068,02	17	23	-6	36
166	256316	1330998	258409,1	-2093,09	2093,09	1	22	-21	441
167	9392	98861	19752,58	-10360,58	10360,58	16	9	7	49
168	11758	130823	25943,4	-14185,40	14185,40	14	6	8	64
169	7227	41506	8643,304	-1416,30	1416,30	25	24	1	1
170	17572	52662	10804,14	6767,86	6767,86	22	10	12	144
171	11349	74121	14960,61	-3611,61	3611,61	19	19	0	0
172	145013	1016286	197451,5	-52438,53	52438,53	3	2	1	1
173	161318	380820	74366,11	86951,89	86951,89	7	1	6	36
174	2797	24848	5416,763	-2619,76	2619,76	26	21	5	25
175	18874	72505	14647,6	4226,40	4226,40	20	16	4	16
176	55073	420788	82107,63	-27034,63	27034,63	6	4	2	4
177	48595	334286	65352,79	-16757,79	16757,79	8	5	3	9
178	13484	122926	24413,8	-10929,80	10929,80	15	8	7	49
179	17025	87352	17523,36	-498,36	498,36	18	28	-10	100
180	20648	71039	14363,64	6284,36	6284,36	21	11	10	100
181	2720	13746	3266,382	-546,38	546,38	28	27	1	1
182	4520	14386	3390,345	1129,65	1129,65	27	25	2	4
183	80104	483652	94283,96	-14179,96	14179,96	5	7	-2	4
184	34806	198432	39038,8	-4232,80	4232,80	11	15	-4	16
185	57503	261388	51232,94	6270,06	6270,06	9	12	-3	9
186	3317	44753	9272,226	-5955,23	5955,23	24	14	10	100
187	31115	135961	26938,59	4176,41	4176,41	13	17	-4	16
188	244397	1024609	199063,6	45333,37	45333,37	2	3	-1	1
189								Сумма	1608

Рис.5.5. Результаты расчета

$$r_{x,e} = 1 - 6 \cdot \frac{1608}{28(28^2 - 1)} = 0,56.$$

Для проверки гипотезы о гомоскедастичности рассчитаем величину t-статистики:

$$t = \frac{r_{x,e} \sqrt{n-2}}{\sqrt{1-r_{x,e}^2}} = \frac{0,56 \sqrt{28-2}}{\sqrt{1-0,56^2}} = 3,446$$

Табличное значение t-статистики при уровне значимости 5% и 26 степенях свободы составляет 2,056. Поскольку фактическое значение $t = 3,446 > t_{5\%,26} = 2,056$, то нулевая гипотеза о гомоскедастичности отклоняется в пользу альтернативной. Таким образом, делаем вывод о **гетероскедастичности** остатков.

3. Тест Голдфелда-Квандта

Для ранжирования значений фактора X и установления соответствующего Y можно воспользоваться **Данные / Сортировка списков** (по возрастанию). В диалоговом окне «Сортировка диапазона» в строке «Сортировать по» выберите фактор, по которому следует упорядочить данные, в нашем случае X . Рекомендуется данные для сортировки поместить на новое место в листе, т.к. ячейки переупорядочиваются в соответствии с заданным порядком сортировки.

Выполненные в соответствии с тестом Голдфелда-Квандта расчеты представлены на рис. 5.6.

	A	B	C	D	E	F	G
58							
59				Упорядоченные по возрастанию ВВП			
60	ДС (Y)	ВВП (X)		ДС (Y)	ВВП (X)		
61	44517	232006	1	2720	13746	0,231	-664,609
62	112617	547203	2	4520	14386	0,053	2916,749
63	13096	50919	3	2797	24848	0,680	4175,929
64	25927	151266	4	7227	41506	19,092	9
65	21581	97624	5	3317	44753	332934354,719	156945457,827
66	256316	1330998	6	13096	50919		
67	9392	98861	7	17572	52662		
68	11758	130823	8	20648	71039		
69	7227	41506	9	18874	72505		
70	17572	52662	10	11349	74121		
71	11349	74121	11	17025	87352	Ftabl= 3,1789	
72	145013	1016286	1				
73	161318	380820	2	F= 13517977523/156945457,8 =		86,13169002	>3,1789
74	2797	24848	3	гетероскедстичность			
75	18874	72505	4				
76	55073	420788	5				
77	48595	334286	6				
78	13484	122926	1	34806	198432	0,186	7593,655
79	17025	87352	2	44517	232006	0,032	21632,448
80	20648	71039	3	57503	261388	0,788	38755,613
81	2720	13746	4	48595	334286	33,366	9
82	4520	14386	5	161318	380820	50115423020,568	13517977523,068
83	80104	483652	6	55073	420788		
84	34806	198432	7	80104	483652		
85	57503	261388	8	112617	547203		
86	3317	44753	9	145013	1016286		
87	31115	135961	10	244397	1024609		
88	244397	1024609	11	256316	1330998		
89							
90							
91	$c \approx \frac{28 \cdot 4}{15} = 7,47 \approx 6$						
92							
93							

Рис.5.6. Результаты расчетов с использованием теста Голдфелда-Квандта

Воспользовавшись функцией ЛИНЕЙН, нашли суммы квадратов остатков по двум моделям, построенным на основе полученных подвыборок (рис. 5.6). В таблицах результатов, которые выводит функция ЛИНЕЙН, эти величины располагаются в последней (пятой) строке второго столбца.

Отношение сумм квадратов остатков оказалось больше табличного значения F-Фишера при 5%-м уровне значимости, что позволило сделать вывод о гетероскедастичности остатков.

4. Тест Парка.

В соответствии с тестом Парка выполним следующие шаги (рис. 5.7):

ДС (Y)	ВВП (X)	\hat{Y}	$e = Y - \hat{Y}$	e^2	$\ln(e^2)$	$\ln(X)$
44517	232006	45541,85	-1024,85	1050320,34	13,865	12,355
112617	547203	106593,36	6023,64	36284296,28	17,407	13,213
13096	50919	10466,54	2629,46	6914072,56	15,749	10,838
25927	151266	29903,07	-3976,07	15809096,11	16,576	11,927
21581	97624	19512,98	2068,02	4276720,11	15,269	11,489
256316	1330998	258409,09	-2093,09	4381016,64	15,293	14,101
9392	98861	19752,58	-10360,58	107341518,59	18,492	11,501
11758	130823	25943,40	-14185,40	201225461,87	19,120	11,782
7227	41506	8643,30	-1416,30	2005916,59	14,512	10,634
17572	52662	10804,14	6767,86	45803864,36	17,640	10,872
11349	74121	14960,61	-3611,61	13043700,47	16,384	11,213
145013	1016286	197451,53	-52438,53	2749798920,04	21,735	13,832
161318	380820	74366,11	86951,89	7560631919,72	22,746	12,850
2797	24848	5416,76	-2619,76	6863158,92	15,742	10,121
18874	72505	14647,60	4226,40	17862472,13	16,698	11,191
55073	420788	82107,63	-27034,63	730871438,15	20,410	12,950
48595	334286	65352,79	-16757,79	280823461,14	19,453	12,720
13484	122926	24413,80	-10929,80	119460553,62	18,598	11,719
17025	87352	17523,36	-498,36	248363,28	12,423	11,378
20648	71039	14363,64	6284,36	39493129,98	17,492	11,171
2720	13746	3266,38	-546,38	298532,91	12,607	9,529
4520	14386	3390,35	1129,65	1276119,80	14,059	9,574
80104	483652	94283,96	-14179,96	201071295,08	19,119	13,089
34806	198432	39038,80	-4232,80	17916571,51	16,701	12,198
57503	261388	51232,94	6270,06	39313603,66	17,487	12,474
3317	44753	9272,23	-5955,23	35464711,34	17,384	10,709
31115	135961	26938,59	4176,41	17442387,56	16,674	11,820
244397	1024609	199063,63	45333,37	2055114132,20	21,444	13,840
				ЛИНЕЙН	1,303	1,768213
					0,32947	3,915783
					0,3757852	2,088631
					15,652326	26
					68,281409	113,4219
				t	3,956302	0,45156

Рис 5.7. Результаты расчетов с использованием теста Парка

- 1). На основе построенного уравнения регрессии $\hat{y}_i = b_0 + b_1 x_i$ найдены расчетные значения \hat{Y} и остатки $e_i = y_i - \hat{y}_i$.
- 2). Определены логарифмы квадратов остатков: $\ln e_i^2 = \ln(y_i - \hat{y}_i)^2$
- 3). Построено уравнение регрессии : $\ln e_i^2 = \ln(1,768) + 1,303 \ln x_i + u_i$.
- 4). Для проверки статистической значимости коэффициента $\beta = 1,303$ определена t-статистика Стьюдента: $t = \frac{1,303}{0,32947} = 3,96$.

Поскольку вычисленная t-статистика больше табличного значения, равного 2,056, то делаем вывод о том, что коэффициент β статистически значим, что означает наличие связи между $\ln e_i^2$ и $\ln x_i$, т. е. **существование гетероскедастичности.**

5. Тест Бреуша – Пагана

В соответствии с тестом Бреуша – Пагана:

- 1) оценили регрессию Y на X с помощью обычного МНК и нашли остатки $e_i = Y_i - \hat{Y}_i$, квадраты остатков e_i^2 и $\hat{\sigma}^2 = (1/n) \sum e_i^2 = 511145955,5$ (см. рис. 5.8).

- 2) построили регрессию

$$\frac{e_i^2}{\hat{\sigma}^2} = 0,2285 + 0,00000287 x_i,$$

и нашли для нее регрессионную сумму квадратов $RSS = 25,546$ (рис. 5.8);

- 3) вычислили статистику $\chi^2 = RSS/2 = 12,773$ и сравнили ее с критическим значением $\chi^2(1) = 3,84$ при выбранном уровне значимости 5% и числе степеней свободы $p = 1$. Поскольку вычисленная статистика $\chi^2 = 12,773$ больше критического значения, равного 3,84, то делаем вывод о **гетероскедастичности.**

B125 fx {=ЛИНЕЙН(G90:G117;C90:C117;1;1)}						
A	B	C	D	E	F	G
	ДС (Y)	ВВП (X)	\hat{Y}	$e = Y - \hat{Y}$	e^2	$e^2/\hat{\sigma}^2$
1	44517	232006	45541,85	-1024,85	1050320,34	0,00205
2	112617	547203	106593,36	6023,64	36284296,28	0,07099
3	13096	50919	10466,54	2629,46	6914072,56	0,01353
4	25927	151266	29903,07	-3976,07	15809096,11	0,03093
5	21581	97624	19512,98	2068,02	4276720,11	0,00837
6	256316	1330998	258409,09	-2093,09	4381016,64	0,00857
7	9392	98861	19752,58	-10360,58	107341518,59	0,21000
8	11758	130823	25943,40	-14185,40	201225461,87	0,39368
9	7227	41506	8643,30	-1416,30	2005916,59	0,00392
10	17572	52662	10804,14	6767,86	45803864,36	0,08961
11	11349	74121	14960,61	-3611,61	13043700,47	0,02552
12	145013	1016286	197451,53	-52438,53	2749798920,04	5,37967
13	161318	380820	74366,11	86951,89	7560631919,72	14,79153
14	2797	24848	5416,76	-2619,76	6863158,92	0,01343
15	18874	72505	14647,60	4226,40	17862472,13	0,03495
16	55073	420788	82107,63	-27034,63	730871438,15	1,42987
17	48595	334286	65352,79	-16757,79	280823461,14	0,54940
18	13484	122926	24413,80	-10929,80	119460553,62	0,23371
19	17025	87352	17523,36	-498,36	248363,28	0,00049
20	20648	71039	14363,64	6284,36	39493129,98	0,07726
21	2720	13746	3266,38	-546,38	298532,91	0,00058
22	4520	14386	3390,35	1129,65	1276119,80	0,00250
23	80104	483652	94283,96	-14179,96	201071295,08	0,39337
24	34806	198432	39038,80	-4232,80	17916571,51	0,03505
25	57503	261388	51232,94	6270,06	39313603,66	0,07691
26	3317	44753	9272,23	-5955,23	35464711,34	0,06938
27	31115	135961	26938,59	4176,41	17442387,56	0,03412
28	244397	1024609	199063,63	45333,37	2055114132,20	4,02060
				Сумма	14312086754,96	
				$\hat{\sigma}^2 = (1/n) \sum e_i^2 =$	Сумма / 28 =	511145955,5
	ЛИНЕЙН					
	2,87438E-06	0,228459618				
	1,62828E-06	0,695555856				
	0,107027561	2,863160742				
	3,116240158	26				
	RSS	25,54596901	213,1399253			
	RSS/2=	12,77298451	> 3,84			
	хи-квадрат табл.	3,84				

Рис.5.8. Результаты расчетов с использованием теста Бреуша – Пагана

6. Тест Уайта

1). При применении теста Уайта оценивается регрессия Y на все X , их квадраты и попарные произведения с помощью обычного МНК. Поскольку в рассматриваемой задаче только одна независимая переменная, от оцениваем регрессию Y на имеющийся X и его квадрат (рис. 5.9); находим остатки

$$e_i = Y_i - \hat{Y}_i \text{ и квадраты остатков } e_i^2 = (Y_i - \hat{Y}_i)^2;$$

Как отмечалось ранее, не выявление методом Уайта гетероскедастичности, может означать лишь отсутствие определенного вида зависимости дисперсии возмущений от значений факторов, тем более, что применение других тестов указывает на гетероскедастичность.

7. Тест Глейзера

1) Оцениваем регрессионную зависимость Y от X с помощью обычного МНК и находим остатки для каждого наблюдения:

$$e_i = Y_i - \hat{Y}_i;$$

2) Оцениваем регрессионные зависимости модулей остатков $|e_i|$ от независимой переменной x , меняя величину показателя степени γ (см. рис. 5.10) в уравнении

$$|e_i| = \alpha_0 + \alpha_1 x_i^\gamma + u_i.$$

Результаты расчетов свели в табл. 5.3.

Наиболее подходящей формой зависимости $|e_i|$ от X оказалась модель вида:

$$|\hat{e}_i| = -2069,79 + 1640,05X^{0.25}$$

так как она характеризуется наилучшими показателями t_1 , R^2 и F .

Коэффициент при $X^{0.25}$ является статистически значимым, следовательно, тест Глейзера позволил сделать вывод о наличии гетероскедастичности и о ее форме, что позволит составить представление о способе преобразования модели с целью преодоления гетероскедастичности.

	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	
						Степени X											
ДС (Y)	ВВП (X)	Y	rach	e	mod e	-1	-0,5	-0,25	0,25	0,5	1	1,5	2				
44517	232006	45541,85	-1024,85	1024,85	4,31E-06	0,002	0,046	21,947	481,670	232006	111750333	5,383E+10					
112617	547203	106593,36	6023,64	6023,64	1,827E-06	0,001	0,037	27,198	739,732	547203	404783410	2,994E+11					
13096	50919	10466,54	2629,46	2629,46	1,964E-05	0,004	0,067	15,022	225,652	50919	11489994	2,593E+09					
25927	151266	29903,07	-3976,07	3976,07	6,611E-06	0,003	0,051	19,721	388,929	151266	58831779	2,288E+10					
21581	97624	19512,98	2068,02	2068,02	1,024E-05	0,003	0,057	17,676	312,448	97624	30502462	9,53E+09					
256316	1330998	258409,09	-2093,09	2093,09	7,513E-07	0,001	0,029	33,966	1153,689	1330998	1,536E+09	1,772E+12					
9392	98861	19752,58	-10360,58	10360,58	1,012E-05	0,003	0,056	17,732	314,422	98861	31084043	9,773E+09					
11758	130823	25943,40	-14185,40	14185,40	7,644E-06	0,003	0,053	19,018	361,695	130823	47317976	1,711E+10					
7227	41506	8643,30	-1416,30	1416,30	2,409E-05	0,005	0,070	14,273	203,730	41506	8456026,3	1,723E+09					
17572	52662	10804,14	6767,86	6767,86	1,899E-05	0,004	0,066	15,149	229,482	52662	12084982	2,773E+09					
11349	74121	14960,61	-3611,61	3611,61	1,349E-05	0,004	0,061	16,500	272,252	74121	20179570	5,494E+09					
145013	1016286	197451,53	-52438,53	52438,53	9,84E-07	0,001	0,031	31,751	1008,110	1016286	1,025E+09	1,033E+12					
161318	380820	74366,11	86951,89	86951,89	2,626E-06	0,002	0,040	24,842	617,106	380820	235006364	1,45E+11					
2797	24848	5416,76	-2619,76	2619,76	4,024E-05	0,006	0,080	12,555	157,632	24848	3916852	617423104					
18874	72505	14647,60	4226,40	4226,40	1,379E-05	0,004	0,061	16,409	269,268	72505	19523242	5,257E+09					
55073	420788	82107,63	-27034,63	27034,63	2,376E-06	0,002	0,039	25,469	648,682	420788	272957492	1,771E+11					
48595	334286	65352,79	-16757,79	16757,79	2,991E-06	0,002	0,042	24,045	578,175	334286	193275712	1,117E+11					
13484	122926	24413,80	-10929,80	10929,80	8,135E-06	0,003	0,053	18,725	350,608	122926	43098844	1,511E+10					
17025	87352	17523,36	-498,36	498,36	1,145E-05	0,003	0,058	17,192	295,554	87352	25817208	7,63E+09					
20648	71039	14363,64	6284,36	6284,36	1,408E-05	0,004	0,061	16,326	266,531	71039	18934126	5,047E+09					
2720	13746	3266,38	-546,38	546,38	7,275E-05	0,009	0,092	10,828	117,243	13746	1611626,9	188952516					
4520	14386	3390,35	1129,65	1129,65	6,951E-05	0,008	0,091	10,952	119,942	14386	1725480,6	206956996					
80104	483652	94283,96	-14179,96	14179,96	2,068E-06	0,001	0,038	26,371	695,451	483652	336356235	2,339E+11					
34806	198432	39038,80	-4232,80	4232,80	5,04E-06	0,002	0,047	21,106	445,457	198432	88392937	3,938E+10					
57503	261388	51232,94	6270,06	6270,06	3,826E-06	0,002	0,044	22,611	511,261	261388	133637539	6,832E+10					
3317	44753	9272,23	-5955,23	5955,23	2,234E-05	0,005	0,069	14,545	211,549	44753	9467454,6	2,003E+09					
31115	135961	26938,59	4176,41	4176,41	7,355E-06	0,003	0,052	19,202	368,729	135961	50132750	1,849E+10					
244397	1024609	199063,63	45333,37	45333,37	9,76E-07	0,001	0,031	31,816	1012,230	1024609	1,037E+09	1,05E+12					
МОДЕЛИ зависимости Y от X в разных степенях																	
						-1	-0,5	-0,25	0,25	0,5	1	1,5	2				
-351301148	17272,53	-4213242,6	25855,01	-562438	42863,308	1640,1	-20697,9	36,1088	-3660,082	0,026129	5262,2178	1,983E-05	8190,952	1,49E-08	9551,732		
194386037	4467,744	1712137,32	6457,167	200179,8	11364,477	519,74	10918,71	11,7512	6088,2198	0,009964	4256,4325	9,341E-06	3939,189	8,71E-09	3870,966		
0,1116	18570,33	0,18890876	17743,95	0,232907	17255,966	0,2769	16753,64	0,26641	16874,985	0,209161	17521,023	0,147745	18188,64	0,101462	18675,98		
3,26609645	26	6,05557992	26	7,894208	26	9,9572	26	9,44192	26	6,876465	26	4,5072999	26	2,935902	26		
1126336765	8,97E+09	1906584931	8,19E+09	2,35E+09	7,742E+09	3E+09	7,3E+09	2,7E+09	7,404E+09	2,11E+09	7,982E+09	1,491E+09	8,6E+09	1,02E+09	9,07E+09		
t-статистики																	
-1,80723448	3,866052	-2,4608088	4,00408	-2,80966	3,771692	3,1555	-1,89563	3,07277	-0,601174	2,622301	1,2362977	2,1230402	2,07935	1,713447	2,467532		

Рис.5.10. Результаты расчетов с использованием теста Глейзера

Таблица 5.3 – Результаты построения регрессий модулей остатков $|e_i|$ от X в разных степенях

Гамма	a_1	t_1	R^2	F
-1	-3,51301148	-1,80723	0,1116	3,2661
-0,5	-4213242,6	-2,46081	0,1889088	6,0556
-0,25	-562437,9	-2,80966	0,2329073	7,8942
0,25	1640,1	3,155501	0,2769178	9,9572
0	36,1088	3,072771	0,2664055	9,4419
1	0,026129	2,622301	0,2091607	6,8765
1,5	1,98E-05	2,12304	0,147745	4,5073
2	1,49E-08	1,713447	0,1014623	2,9359

Таким образом, проверка нулевой гипотезы о гомоскедастичности остатков регрессионной модели, выполненная с помощью различных тестов, показала не однозначный результат: в большинстве случаев эта гипотеза при уровне значимости 5% была отклонена, а тест Уайта не позволил ее отклонить.

На самом деле противоречия здесь нет: ведь тот факт, что гипотеза не отвергается, не означает, что она на самом деле верна; можно лишь утверждать, что если она на самом деле верна, то наблюдаемый результат возможен с вероятностью, большей, чем уровень значимости.

Пример 2. Применение взвешенного метода наименьших квадратов для оценивания модели с гетероскедастичными остатками

Продолжим решение задачи о построении модели зависимости добавленной стоимости в обрабатывающей промышленности (Y) от валового внутреннего продукта (X) (см. пример 1). Модель регрессии, характеризующая эту зависимость, получена в виде:

$$\hat{Y} = 603,9 + 0,194X.$$

Анализ регрессионных остатков показал наличие гетероскедастичности. Поэтому применение обычного метода наименьших квадратов для оценивания параметров модели привело к получению неэффективных оценок.

Для оценивания параметров модели используем взвешенный метод наименьших квадратов (или обобщенный МНК со вспомогательной

матрицей). Для использования этого метода надо, чтобы остатки не были автокоррелированы¹.

При использовании теста Глейзера выяснили, что наиболее подходящей формой зависимости $|e_i|$ от X оказалась модель вида:

$$|\hat{e}_i| = -2069,79 + 1640,05X^{0.25}$$

Подставляя в это уравнение значения переменной X (величина ВВП), получаем расчетные значения модуля остатков (рис. 5.11), которые будем использовать для преобразования исходных данных:

$$Y_{np} = P^{-1}Y = \begin{pmatrix} \frac{Y_1}{\sqrt{\lambda_1}} \\ \frac{Y_2}{\sqrt{\lambda_2}} \\ \frac{Y_3}{\sqrt{\lambda_3}} \\ \dots \\ \frac{Y_n}{\sqrt{\lambda_n}} \end{pmatrix}; \quad X_{np} = P^{-1}X = \begin{pmatrix} \frac{1}{\sqrt{\lambda_1}} & \frac{X_{11}}{\sqrt{\lambda_1}} & \dots & \frac{X_{1p}}{\sqrt{\lambda_1}} \\ \frac{1}{\sqrt{\lambda_2}} & \frac{X_{21}}{\sqrt{\lambda_2}} & \dots & \frac{X_{2p}}{\sqrt{\lambda_2}} \\ \dots & \dots & \dots & \dots \\ \frac{1}{\sqrt{\lambda_n}} & \frac{X_{n1}}{\sqrt{\lambda_n}} & \dots & \frac{X_{np}}{\sqrt{\lambda_n}} \end{pmatrix}.$$

$$\text{Так, } |\hat{e}_1| = -2069,79 + 1640,05 \cdot 232006^{0.25} = 15296,25 = \sqrt{\lambda_1},$$

$$|\hat{e}_2| = -2069,79 + 1640,05 \cdot 547203^{0.25} = 23908,20 = \sqrt{\lambda_2} \text{ и т. д.}$$

$$44517 / 15296,25 = 2,91,$$

$$112617 / 23908,2 = 4,74 \text{ и т. д. (см. рис. 5.11).}$$

Используя преобразованные таким образом данные, с помощью МНК построили модель:

$$\frac{Y}{\sqrt{\lambda}} = -2333 \frac{1}{\sqrt{\lambda}} + 0,203 \frac{X}{\sqrt{\lambda}}.$$

¹Проверка выполнения данной предпосылки классической регрессии будет выполнена в следующей лабораторной работе. Но в данном случае автокорреляция остатков не имеет места.

E38		fx		{=ЛИНЕЙН(E4:E31;F4:G31;0;1)}			
	A	B	C	D	E	F	G
1	Взвешенный МНК			Преобразованные данные			
2	Исходные данные						
3	ДС (Y)	X0	ВВП (X)	$ \hat{\varepsilon}_i = \sqrt{\lambda_i}$	$Y/\sqrt{\lambda}$	$X_0/\sqrt{\lambda}$	$X/\sqrt{\lambda}$
4	44517	1	232006	15296,25	2,91	6,5376E-05	15,17
5	112617	1	547203	23908,20	4,71	4,1827E-05	22,89
6	13096	1	50919	3938,49	3,33	0,0002539	12,93
7	25927	1	151266	11646,00	2,23	8,5866E-05	12,99
8	21581	1	97624	8291,97	2,60	0,0001206	11,77
9	256316	1	1330998	35008,03	7,32	2,8565E-05	38,02
10	9392	1	98861	8383,37	1,12	0,00011928	11,79
11	11758	1	130823	10493,01	1,12	9,5301E-05	12,47
12	7227	1	41506	2711,20	2,67	0,00036884	15,31
13	17572	1	52662	4146,67	4,24	0,00024116	12,70
14	11349	1	74121	6363,01	1,78	0,00015716	11,65
15	145013	1	1016286	31374,92	4,62	3,1873E-05	32,39
16	161318	1	380820	20043,60	8,05	4,9891E-05	19,00
17	2797	1	24848	-106,78	-26,19	-0,00936496	-232,70
18	18874	1	72505	6214,29	3,04	0,00016092	11,67
19	55073	1	420788	21072,91	2,61	4,7454E-05	19,97
20	48595	1	334286	18737,53	2,59	5,3369E-05	17,84
21	13484	1	122926	10011,27	1,35	9,9887E-05	12,28
22	17025	1	87352	7497,31	2,27	0,00013338	11,65
23	20648	1	71039	6077,21	3,40	0,00016455	11,69
24	2720	1	13746	-2939,61	-0,93	-0,00034018	-4,68
25	4520	1	14386	-2736,42	-1,65	-0,00036544	-5,26
26	80104	1	483652	22552,52	3,55	4,4341E-05	21,45
27	34806	1	198432	13916,76	2,50	7,1856E-05	14,26
28	57503	1	261388	16385,41	3,51	6,103E-05	15,95
29	3317	1	44753	3156,18	1,05	0,00031684	14,18
30	31115	1	135961	10794,86	2,88	9,2637E-05	12,59
31	244397	1	1024609	31481,21	7,76	3,1765E-05	32,55
32	$ \hat{\varepsilon}_i = -20697,9 + 1640,05X^{0,25}$						
33							
34	ЛИНЕЙН						
35	Модель по исходным данным			Модель по преобразованным данным			
36	0,193693	603,8754		0,203162523	-2233,00068	0	
37	0,013343	5699,688		0,01584698	421,882094	#Н/Д	
38	0,890172	23461,99		0,958013548	1,30574994	#Н/Д	
39	210,7339	26		296,623684	26	#Н/Д	
40	1,16E+11	1,43E+10		1011,476618	44,3295554	#Н/Д	
41	t=	14,51668	0,105949	t=	12,82026772	-5,29294965	
42							
43	$Y = 603,875 + 0,194X$			$\frac{Y}{\sqrt{\lambda_i}} = -2233 \frac{1}{\sqrt{\lambda_i}} + 0,203 \frac{X_i}{\sqrt{\lambda_i}}$			
44							
45							
46							
47	$Y = -2233 + 0,203X$						

Рис.5.11. Расчетные значения модуля остатков

Коэффициенты уравнения, полученного на основе преобразованных данных, нужно интерпретировать в рамках исходного уравнения регрессии.

Возвращаясь к исходным обозначениям, запишем модель зависимости добавленной стоимости в обрабатывающей промышленности (Y) от валового внутреннего продукта (X) в виде.

$$Y = -2333 + 0,203X.$$

Проверим, удалось ли построить модель с гомоскедастичными остатками. Применив тест Голдфелда-Квандта к преобразованным данным, убедились в том, что гетероскедастичность удалось устранить (см. рис. 5.12).

	A	B	C	D	E	F	G	H	I
109									
110	Преобразованные данные				Упорядоченные по возрастанию X преобр.				
111									
112	$Y/\sqrt{\lambda}$	$X_0/\sqrt{\lambda}$	$X/\sqrt{\lambda}$		$Y/\sqrt{\lambda}$	$X_0/\sqrt{\lambda}$	$X/\sqrt{\lambda}$		
113	2,91	6,54E-05	15,17	1	-26,19	-0,0093650	-232,70	0,1783	-1620,5388
114	4,71	4,18E-05	22,89	2	-1,65	-0,0003654	-5,26	0,0394	986,8257
115	3,33	0,000254	12,93	3	-0,93	-0,0003402	-4,68	0,9880	0,9864
116	2,23	8,59E-05	12,99	4	1,78	0,0001572	11,65	370,6142	9
117	2,60	0,000121	11,77	5	2,27	0,0001334	11,65	721,1479	8,7562
118	7,32	2,86E-05	38,02	6	3,04	0,0001609	11,67		
119	1,12	0,000119	11,79	7	3,40	0,0001645	11,69		
120	1,12	9,53E-05	12,47	8	2,60	0,0001206	11,77		
121	2,67	0,000369	15,31	9	1,12	0,0001193	11,79		
122	4,24	0,000241	12,70	10	1,35	0,0000999	12,28		
123	1,78	0,000157	11,65	11	1,12	0,0000953	12,47		
124	4,62	3,19E-05	32,39	1				F_{табл} = 3,1789	
125	8,05	4,99E-05	19,00	2					
126	-26,19	-0,00936	-232,70	3	F= 26,7804 / 8,7562 =			3,0585	< 3,1789
127	3,04	0,000161	11,67	4				гомоскедастичность	
128	2,61	4,75E-05	19,97	5					
129	2,59	5,34E-05	17,84	6					
130	1,35	9,99E-05	12,28	1	2,91	0,0000654	15,17	0,1999	-587,1862
131	2,27	0,000133	11,65	2	2,67	0,0003688	15,31	0,0249	4976,2383
132	3,40	0,000165	11,69	3	3,51	0,0000610	15,95	0,9031	1,7250
133	-0,93	-0,00034	-4,68	4	2,59	0,0000534	17,84	41,9216	9
134	-1,65	-0,00037	-5,26	5	8,05	0,0000499	19,00	249,4837	26,7804
135	3,55	4,43E-05	21,45	6	2,61	0,0000475	19,97		
136	2,50	7,19E-05	14,26	7	3,55	0,0000443	21,45		
137	3,51	6,1E-05	15,95	8	4,71	0,0000418	22,89		
138	1,05	0,000317	14,18	9	4,62	0,0000319	32,39		
139	2,88	9,26E-05	12,59	10	7,76	0,0000318	32,55		
140	7,76	3,18E-05	32,55	11	7,32	0,0000286	38,02		
141									

Рис.5.12. Проверка остатков на гетероскедастичность

Сравним результаты оценивания регрессии для моделей, построенных по первоначальным и преобразованным данным (рис.5.11).

Видим, что оценка коэффициента при X отличается незначительно. Этого и следовало ожидать, поскольку гетероскедастичность не приводит к смещению оценок. Оценка коэффициента при X в модели по преобразованным данным должна иметь минимальную дисперсию и поэтому быть более точной. Может показаться странным, что ее стандартная ошибка, равная 0,01585, больше соответствующей величины из модели по исходным данным, равной 0,013343 (см. рис. 5.11). Но это не следует принимать во внимание, так как в условиях гетероскедастичности ее расчет некорректен (ошибка обладает смещением – занижена).

Таким образом, применив при моделировании зависимости добавленной стоимости в обрабатывающей промышленности (Y) от валового внутреннего продукта (X) для межстрановой выборки взвешенный метод наименьших квадратов, получили адекватную модель, которая позволяет сделать вывод, что при изменении ВВП на млн.долл.США добавленная стоимость в среднем изменяется на 0,203 млн.долл.США.

Вопросы и задания для самостоятельной работы

1. В чем суть понятий «гомо- и гетероскедастичности»?
2. Как влияет гетероскедастичность на результаты оценивания эконометрической модели?
3. Охарактеризуйте методы тестирования гетероскедастичности.
4. В чем заключается тест Спирмена?
5. Какова схема теста Голдфелда-Квандта?
6. Опишите тест Глейзера.
7. Каково предположение теста Парка?
8. Как используется тест Уайта?
9. В чем суть метода взвешенных наименьших квадратов?

10. Какие типы преобразований применяются для устранения гетероскедастичности?

Задание 1. В табл.5.4 представлены данные о потребительских расходах C и располагаемом доходе Y_d тридцати семей (ден. ед).

Таблица 5.4 – Исходные данные

Потребление			Доход
10700	10900	11200	12000
11400	11700	12100	13000
12300	12600	13200	14000
13000	13300	13600	15000
13800	14000	14200	16000
14400	14900	15300	17000
15000	15700	16400	18000
15900	16500	16900	19000
16900	17500	18100	20000
17200	17800	18500	21000

- 1) Постройте регрессию C на Y_d и проверьте наличие или отсутствие гетероскедастичности (примените различные тесты).
- 2) Если в п.1) выявлена гетероскедастичность, осуществите коррекцию на гетероскедастичность, предварительно высказав предположение о форме зависимости дисперсии возмущения от независимой переменной.

Задание 2. Выдвигается предположение, что средняя заработная плата наемных рабочих пропорциональна их стажу. Для анализа данного утверждения обследуются по 20 рабочих восьми категорий стажа. Получены следующие статистические данные (табл. 5.5):

Таблица 5.5 – Исходные данные

Стаж	[0,5)	[5,10)	[10,15)	[15,20)	[20,25)	[25,30)	[30,35)	[35,40]
Зар. плата	10000	12500	14300	18700	25400	29000	32000	34300

1) Постройте эмпирическое уравнение регрессии, в котором заработная плата является зависимой переменной, а стаж работы – объясняющей переменной (уравнение строится в предположении, что дисперсии отклонений постоянны).

2) Оцените качество построенной регрессии.

3) Есть ли основания считать, что для данной регрессионной модели весьма вероятна гетероскедастичность? Если да, то почему?

4) Предполагая, что дисперсия отклонений пропорциональна трудовому стажу, постройте на основании тех же данных уравнение с применением взвешенного метода наименьших квадратов.

5) Предполагая, что дисперсия отклонений пропорциональна квадрату трудового стажа, постройте соответствующее уравнение регрессии с применением взвешенного метода наименьших квадратов.

6) Какое из трех предположений относительно дисперсии отклонений наиболее реалистично с вашей точки зрения?

ТЕМА 6. ОБОБЩЕННАЯ МОДЕЛЬ ЛИНЕЙНОЙ РЕГРЕССИИ С АВТОКОРРЕЛИРОВАННЫМИ ОСТАТКАМИ (ТЕСТИРОВАНИЕ И УСТРАНЕНИЕ ПОСЛЕДСТВИЙ АВТОКОРРЕЛЯЦИИ)

Данная тема посвящена изучению одной из разновидностей обобщенной модели линейной регрессии (эконометрической модели), которая отличается от классической тем, что возмущающие переменные взаимозависимы (автокоррелированы).

Цель выполнения лабораторной работы: научиться выявлять нарушения одной из основных предпосылок классической регрессионной модели

(автокоррелированности остатков); описывать последствия этих нарушений с точки зрения качества модели; применять методы оценивания параметров модели, позволяющие уменьшить отрицательные последствия, вызванные этими нарушениями

Для выполнения лабораторной работы № 6 студенты должны

знать:

- содержание основных предпосылок классической регрессии;
- последствия нарушения предпосылок с точки зрения качества построенной методом наименьших квадратов модели;
- основные способы тестирования предпосылок;
- способы устранения негативных последствий нарушения предпосылок;

уметь:

- пользоваться таблицами для проверки статистических гипотез;
- пользоваться пакетом EXCEL:
 - строить таблицы; составлять и копировать формулы;
 - пользоваться МАСТЕРОМ ФУНКЦИЙ: статистическими и математическими функциями.

Задание для выполнения лабораторной работы №6

Для выборочных данных (табл. 6.1) постройте эконометрическую модель, в которой зависимой переменной будет фактор Y , а независимой – X . Для этого необходимо:

- 1) методом наименьших квадратов (МНК) вычислить оценки параметров модели, проверить статистическую надежность модели в целом и оценок параметров;
- 2) проанализировать целесообразность применения МНК с помощью, проверив регрессионные остатки на автокорреляцию с применением различных тестов;
- 3) рассчитать оценки параметров модели с учетом проведенного тестирования на наличие автокорреляции остатков.

Таблица 6.1 – Исходные данные по вариантам [21]

Вариант 1			Вариант 2			Вариант 3		
№ п/п	Темп прироста индекса потребительских цен, Y, %	Темп прироста индекса оптовых цен, X, %	№ п/п	Темп прироста индекса потребительских цен, Y, %	Учетная ставка X, %	№ п/п	Темп прироста индекса оптовых цен, Y, %	Учетная ставка X, %
1	3,2	0,5	1	3,2	3,1	1	0,5	3,1
2	2,2	0,8	2	2,2	2,9	2	0,8	2,9
3	1,1	1,8	3	1,1	2,3	3	1,8	2,3
4	1,8	1,6	4	1,8	2,1	4	1,6	2,1
5	1,8	1,2	5	1,8	2,0	5	1,2	2,0
6	1,1	1,3	6	1,1	1,8	6	1,3	1,8
7	0,9	1,4	7	0,9	1,6	7	1,4	1,6
8	1,0	1,0	8	1,0	1,3	8	1,0	1,3
9	2,2	1,2	9	2,2	1,3	9	1,2	1,3
10	1,0	1,7	10	1,9	1,3	10	1,7	1,3
11	1,9	0,8	11	1,9	2,0	11	0,8	2,0
12	2,4	1,5	12	2,4	2,9	12	1,5	2,9
13	2,3	1,8	13	2,3	2,9	13	1,8	2,9
14	1,2	1,9	14	1,2	3,5	14	1,9	3,5
15	1,2	1,7	15	1,2	3,6	15	1,7	3,6
16	2,3	1,5	16	2,3	3,4	16	1,5	3,4
17	1,4	0,7	17	1,4	3,6	17	0,7	3,6
18	1,0	1,2	18	1,0	4,3	18	1,2	4,3
19	0,1	1,6	19	0,1	6,3	19	1,6	6,3
20	1,2	2,2	20	1,2	6,8	20	2,2	6,8
21	4,8	10,4	21	4,8	6,8	21	10,4	6,8
22	7,2	11,7	22	7,2	6,8	22	11,7	6,8
23	6,8	12,1	23	6,8	6,5	23	12,1	6,5
24	7,0	12,5	24	7,0	6,0	24	12,5	6,0
25	7,3	13,4	25	7,3	6,0	25	13,4	6,0
26	5,1	10,4	26	5,1	6,8	26	10,4	6,8
27	5,4	9,1	27	5,4	7,0	27	9,1	7,0
28	6,1	9,2	28	6,1	7,2	28	9,2	7,2
29	6,1	8,9	29	6,1	7,2	29	8,9	7,2
30	5,9	9,0	30	5,9	7,5	30	9,0	7,5

Вариант 4			Вариант 5			Вариант 6		
№ п/п	Темп прироста индекса оптовых цен, Y, %	Средне-взвешенные процентные ставки по кредитам КБ, X, %	№ п/п	Средне-взвешенные процентные ставки по депозитам КБ, Y, %	Средне-взвешенные процентные ставки по кредитам КБ, X, %	№ п/п	Темп прироста индекса потребительских цен, Y, %	Долларизация экономики X, %
1	0,5	5,3	1	3,0	5,3	1	2,6	32,3
2	0,8	5,2	2	2,9	5,2	2	1,1	34,5
3	1,8	4,4	3	2,4	4,4	3	1,2	34,5
4	1,6	4,4	4	2,3	4,4	4	1,3	34,5
5	1,2	4,5	5	2,4	4,5	5	2,2	30,3
6	1,3	4,2	6	2,3	4,2	6	1,2	30,0
7	1,4	3,9	7	2,4	3,9	7	0,1	28,3
8	1,0	3,5	8	1,9	3,5	8	0,8	29,7

9	1,2	3,5	9	2,0	3,5	9	0,6	28,5
10	1,7	3,2	10	1,9	3,2	10	0,1	28,7
11	0,8	3,5	11	2,0	3,5	11	0,1	28,5
12	1,5	3,6	12	2,4	3,6	12	0,03	28,7
13	1,8	3,7	13	2,5	3,7	13	1,2	28,4
14	1,9	4,1	14	2,4	4,1	14	0,9	27,8
15	1,7	4,1	15	2,5	4,1	15	0,9	27,9
16	1,5	4,0	16	2,5	4,0	16	1,4	26,5
17	0,7	3,9	17	2,5	3,9	17	1,3	28,0
18	1,2	4,0	18	2,6	4,0	18	0,2	30,8
19	1,6	4,4	19	2,6	4,4	19	0,2	28,8
20	2,2	4,8	20	2,8	4,8	20	1,3	30,7
21	10,4	5,3	21	2,1	5,3	21	0,07	33,2
22	11,7	6,0	22	2,4	6,0	22	0,05	28,9
23	12,1	5,5	23	2,6	5,5	23	0,9	31,1
24	12,5	5,7	24	2,5	5,7	24	0,2	34,0
25	13,4	6,0	25	2,5	6,0	25	3,8	42,4
26	10,4	6,0	26	2,8	6,0	26	6,2	41,4
27	9,1	6,3	27	3,2	6,3	27	3,0	40,2
28	9,2	6,2	28	3,2	6,2	28	3,3	39,8
29	8,9	6,4	29	3,3	6,4	29	1,5	39,9
30	9,9	6,5	30	3,2	6,5	30	1,0	40,4

Вариант 7			Вариант 8			Вариант 9		
№ п/п	Выпуск промышленной продукции Y, млн.руб	Численность ППП X, тыс.чел.	№ п/п	Выпуск промышленной продукции Y, млн.руб	Кредиты пром. предприятиям X, млн.руб	№ п/п	Кредиты пром.предприятиям Y,млн.руб	Средне-взвешенные процентные ставки по кредитамX,%
1	1574	5991	1	1574	838	1	58	17,1
2	1840	5492	2	1840	992	2	101	13,4
3	2296	5063	3	2296	1253	3	303	18,1
4	6320	5022	4	6320	1202	4	391	22,8
5	10849	4872	5	10849	1357	5	538	30,1
6	12456	4812	6	12456	2176	6	792	24,5
7	14853	4614	7	14853	2886	7	1253	13,5
8	17730	4552	8	17730	3029	8	1202	15,2
9	16929	4502	9	16929	2997	9	1357	17,3
10	17002	4383	10	17002	3060	10	2176	10,2
11	15915	4072	11	15915	3332	11	2886	6,5
12	17341	4398	12	17341	4103	12	3029	8,4
13	15942	4134	13	15942	4126	13	2997	9,0
14	16237	4034	14	16237	4487	14	3060	7,0
15	16379	4054	15	16379	5095	15	3332	5,4
16	18624	3842	16	18624	5196	16	4103	5,3
17	17455	3842	17	17455	5001	17	4126	5,0
18	16810	3724	18	16810	5104	18	4487	4,3
19	18000	3677	19	18000	4985	19	5095	3,6

20	22954	3598	20	22954	5102	20	5196	3,4
21	22001	3693	21	22001	5198	21	5001	3,9
22	22982	3646	22	22982	5243	22	5104	4,0
23	25686	3452	23	25686	5968	23	4985	4,8
24	28466	3124	24	28466	5715	24	5102	5,4
25	28224	3087	25	28224	5610	25	5189	5,0
26	29186	2985	26	29186	5845	26	5243	4,7
27	29870	2874	27	29870	6092	27	5968	4,2
28	30045	2830	28	30045	5929	28	5715	4,4
29	31224	2841	29	31224	6101	29	7245	4,1
30	32307	2893	30	32307	6462	30	7331	4,5

Вариант 10			Вариант 11			Вариант 12		
№ п/п	Прибыль предприятия, Y, тыс. руб	Среднемесячная стоимость основных фондов X, тыс. руб	№ п/п	Прибыль предприятия, Y, тыс. руб	Среднемесячная стоимость оборотных средств X, тыс. руб	№ п/п	Прибыль предприятия, Y, тыс. руб	Фондоёмкость продукции X, руб
1	1658	951	1	6658	151	1	1622	0,07
2	2169	996	2	7555	232	2	3298	0,10
3	1934	1024	3	6989	326	3	3054	0,13
4	1862	1062	4	7321	416	4	3197	0,13
5	2345	1127	5	7543	442	5	3292	0,11
6	2095	1003	6	7678	384	6	3351	0,13
7	2246	1068	7	7716	357	7	3367	0,11
8	2465	1052	8	7556	382	8	3298	0,11
9	2318	924	9	7723	406	9	3370	0,12
10	2708	1194	10	7736	452	10	3375	0,10
11	3807	1367	11	7885	473	11	3440	0,12
12	2667	1432	12	7980	504	12	3480	0,13
13	2856	1628	13	7994	555	13	3486	0,13
14	2881	1934	14	8214	583	14	3771	0,15
15	2903	2064	15	8375	498	15	3844	0,14
16	2945	2153	16	9742	548	16	4459	0,12
17	4105	1935	17	9425	713	17	4316	0,17
18	3362	2697	18	9337	679	18	4277	0,17
19	3856	3526	19	9177	952	19	4205	0,20
20	4126	3982	20	9119	991	20	4179	0,22
21	4923	4627	21	9520	1006	21	4359	0,25
22	7268	3062	22	9964	1108	22	4559	0,26
23	6086	5826	23	10212	1227	23	4670	0,32
24	6913	5113	24	9968	1027	24	4561	0,25
25	7025	5324	25	9392	1268	25	4301	0,25
26	6111	5948	26	9785	1452	26	4478	0,27
27	6095	6257	27	10162	1565	27	4648	0,29
28	6532	6391	28	10706	1584	28	4893	0,32
29	7268	6024	29	10794	1690	29	4932	0,28
30	6325	5483	30	10805	1705	30	4937	0,33

Вариант 13			Вариант 14			Вариант 15		
№ п/п	Личные доходы населения региона, Y, млн.руб	Потребительские расходы, X, млн.руб	№ п/п	Личные доходы населения региона, Y, млн.руб	Сбережения в банках X, млн.руб	№ п/п	Индекс потребительских цен, Y,%	Индекс официального курса рубля, X,%
1	15,90	5,80	1	15,90	0,50	1	102,2	110,4
2	18,60	6,30	2	18,60	0,90	2	101,2	97,5
3	19,40	8,70	3	19,40	1,02	3	100,1	99,6
4	21,47	8,90	4	21,47	2,40	4	100,8	100,6
5	23,32	10,20	5	23,32	3,05	5	100,8	99,7
6	24,97	10,90	6	24,97	5,90	6	100,1	100,8
7	26,72	11,07	7	26,72	8,25	7	100,1	99,9
8	28,47	12,87	8	28,47	9,76	8	100,0	99,9
9	30,22	14,67	9	30,22	10,23	9	101,2	100,3
10	37,97	16,47	10	37,97	11,62	10	100,9	100,5
11	33,72	18,27	11	33,72	13,02	11	100,9	100,4
12	35,47	20,00	12	35,47	14,41	12	101,4	100,9
13	37,22	22,00	13	37,22	15,80	13	101,3	100,7
14	38,97	23,60	14	38,97	17,19	14	100,2	102,5
15	40,72	25,47	15	40,72	18,58	15	100,2	103,9
16	42,47	27,27	16	42,47	19,97	16	101,3	100,3
17	44,22	29,07	17	44,22	21,36	17	100,0	100,5
18	45,97	30,87	18	45,97	22,75	18	100,0	100,4
19	47,72	32,67	19	47,72	24,14	19	99,1	102,1
20	49,47	34,47	20	49,47	25,54	20	100,2	103,8
21	51,22	36,27	21	51,22	26,93	21	103,8	127,6
22	52,97	38,07	22	52,97	28,32	22	106,2	122,9
23	54,72	39,87	23	54,72	29,71	23	103,0	100,1
24	56,47	41,67	24	56,47	31,10	24	103,3	100,0
25	58,22	43,47	25	58,22	32,49	25	101,5	100,0
26	59,97	45,27	26	59,97	33,80	26	101,0	101,3
27	61,72	47,07	27	61,72	35,27	27	101,0	108,7
28	63,47	48,87	28	63,47	36,67	28	102,3	104,3
29	65,22	50,67	29	65,22	38,06	29	102,4	99,7
30	66,00	51,00	30	66,00	39,45	30	100,1	100,7

Вариант 16			Вариант 17			Вариант 18		
№ п/п	Темп прироста индекса потребительских цен, Y,%	Темп прироста индекса оптовых цен, X,%	№ п/п	Темп прироста индекса потребительских цен, Y,%	Учетная ставка X,%	№ п/п	Темп прироста индекса оптовых цен, Y,%	Учетная ставка X,%
1	3,3	0,7	1	4,4	3,1	1	0,9	3,1
2	2,4	1,0	2	3,1	2,9	2	1,2	2,9
3	1,3	2,1	3	1,6	2,3	3	2,5	2,3
4	2,0	1,9	4	2,5	2,1	4	2,3	2,1

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

Вариант 19			Вариант 20			Вариант 21		
№ п/п	Темп прироста индекса оптовых цен, Y, %	Средневзвешенные процентные ставки по кредитам КБ, X, %	№ п/п	Средневзвешенные процентные ставки по депозитам КБ, Y, %	Средневзвешенные процентные ставки по кредитам КБ, X, %	№ п/п	Темп прироста индекса потребительских цен, Y, %	Доля ризация экономики X, %
1	1,1	5,3	1	2,0	5,3	1	2,6	31,1
2	1,5	5,2	2	1,9	5,2	2	1,2	33,2
3	2,8	4,4	3	1,5	4,4	3	1,3	33,2
4	2,5	4,4	4	1,4	4,4	4	1,4	33,2
5	2,0	4,5	5	1,5	4,5	5	2,2	29,2
6	2,1	4,2	6	1,4	4,2	6	1,3	28,4
7	2,3	3,9	7	1,5	3,7	7	0,3	26,8
8	1,8	3,5	8	1,1	3,3	8	1,0	28,1
9	2,0	3,5	9	1,2	3,3	9	0,8	27,0
10	2,6	3,2	10	1,1	3,0	10	0,3	27,2
11	1,5	3,5	11	1,2	3,3	11	0,3	27,0
12	2,4	3,6	12	1,5	3,4	12	0,2	2,72
13	2,8	3,7	13	1,6	3,5	13	1,3	26,9
14	2,9	4,1	14	1,5	4,1	14	1,1	26,3

15	2,6	4,1	15	1,6	4,1	15	1,1	26,4
16	2,4	4,0	16	1,6	3,8	16	1,5	25,1
17	1,4	3,9	17	1,6	3,7	17	1,4	26,5
18	2,0	4,0	18	1,7	3,8	18	0,4	29,7
19	2,5	4,4	19	1,7	4,4	19	0,4	27,3
20	2,1	4,8	20	1,8	4,8	20	1,4	29,6
21	13,6	5,3	21	2,2	5,3	21	0,2	32,0
22	15,4	6,0	22	2,4	6,0	22	0,2	27,4
23	15,9	5,5	23	2,6	5,5	23	0,1	30,0
24	16,5	5,7	24	2,5	5,7	24	0,4	32,7
25	17,8	6,0	25	2,5	6,0	25	3,8	40,8
26	13,6	6,0	26	2,7	6,0	26	6,1	39,8
27	11,7	6,3	27	3,0	6,3	27	3,0	38,7
28	11,9	6,2	28	3,0	6,2	28	3,3	38,3
29	11,5	6,4	29	3,1	6,4	29	1,6	38,4
30	11,6	6,5	30	3,0	6,5	30	1,2	38,9

Вариант 22			Вариант 23			Вариант 24		
№ п/п	Выпуск промышленной продукции Y, млн.руб	Численность ППП X, тыс.чел	№ п/п	Выпуск промышленной продукции Y, млн.руб	Кредиты пром. предприятиям X, млн.руб	№ п/п	Кредиты пром. предприятиям Y, млн.р	Средневзвешенные процентные ставки по кредитам КБ, X,%
1	2023	5380	1	1565	729	1	221	17,7
2	2255	4965	2	1786	858	2	268	13,8
3	2652	4609	3	2165	1078	3	490	18,7
4	6152	4575	4	5505	1035	4	587	23,6
5	10093	4451	5	9264	1165	5	749	31,2
6	11491	4401	6	10597	1853	6	1028	25,4
7	13576	4237	7	12587	2449	7	1535	13,9
8	16488	4185	8	16556	2569	8	1479	15,7
9	15807	4144	9	15875	2542	9	1650	17,9
10	15869	4045	10	15937	2595	10	2551	10,5
11	14500	3787	11	13468	2824	11	3332	6,7
12	16157	4057	12	16225	3472	12	3489	8,6
13	14524	3838	13	13491	3491	13	3454	9,3
14	15218	3755	14	13736	3794	14	3523	7,2
15	15339	3772	15	13854	4530	15	3822	5,6
16	17247	3452	16	17315	4617	16	4670	5,5
17	16254	3452	17	16322	4449	17	4696	5,2
18	15706	3351	18	15774	4537	18	5093	4,5
19	16717	3311	19	16785	4212	19	5911	3,8
20	20928	3244	20	20996	4536	20	6032	3,6
21	20118	3325	21	20186	4611	21	5798	4,1
22	20952	3285	22	21020	4657	22	5922	4,2
23	23250	3120	23	23318	5280	23	5641	5,0
24	25613	2841	24	25681	5063	24	5919	5,6
25	25407	2810	25	25475	4973	25	6024	5,2

26	26225	2723	26	26293	5175	26	6089	4,9
27	26807	2629	27	26875	5387	27	6959	4,4
28	26955	2592	28	27023	5247	28	6655	4,6
29	27957	2601	29	28025	5395	29	8491	4,3
30	28878	2645	30	28946	5705	30	8594	4,7

Вариант 25			Вариант 26			Вариант 27		
№ п/п	Прибыль предприятия, Y, тыс.руб	Среднемесячная стоимость основных фондов X, тыс.руб	№ п/п	Прибыль предприятия, Y, тыс.руб	Среднемесячная стоимость оборотных средств X, тыс.руб	№ п/п	Прибыль предприятия, Y, тыс.руб	Фондоемкость продукции и X, руб
1	1902	1142	1	1021	237	1	1796	0,10
2	2464	1192	2	4146	288	2	3639	0,13
3	2205	1222	3	3846	347	3	3372	0,16
4	2126	1264	4	4022	404	4	3529	0,16
5	2658	1336	5	4140	420	5	3634	0,14
6	2383	1199	6	4211	384	6	3698	0,16
7	2549	1271	7	4231	367	7	3716	0,14
8	2790	1253	8	4147	383	8	3640	0,14
9	2628	1112	9	4235	398	9	3719	0,15
10	3057	1409	10	4242	427	10	3725	0,13
11	4266	1600	11	4321	440	11	3796	0,15
12	3012	1671	12	4371	453	12	3840	0,16
13	3220	1887	13	4379	487	13	3847	0,16
14	3247	2223	14	4797	506	14	4160	0,19
15	3271	2366	15	4889	456	15	4240	0,17
16	3318	2464	16	5668	482	16	5339	0,15
17	5041	2225	17	5487	593	17	5168	0,18
18	3776	3451	18	5437	570	18	5120	0,18
19	4320	4446	19	5346	753	19	5034	0,22
20	5066	4993	20	5313	779	20	5002	0,24
21	6023	5767	21	5541	789	21	5219	0,28
22	8837	3889	22	5794	857	22	2459	0,29
23	7418	7206	23	5936	937	23	5592	0,36
24	8409	6351	24	5797	803	24	5461	0,28
25	8545	6604	25	5468	965	25	5150	0,28
26	7448	7353	26	5692	1088	26	5362	0,30
27	7429	7723	27	5907	1164	27	5565	0,33
28	7953	7884	28	6217	1176	28	5859	0,36
29	8837	7444	29	6268	1247	29	5907	0,32
30	7705	6795	30	6274	1257	30	5913	0,38

Вариант 28			Вариант 29			Вариант 30		
№ п/п	Личные доходы населения региона, Y, млн.руб	Потребительские расходы, X, млн.руб	№ п/п	Личные доходы населения региона, Y, млн.руб	Сбережения в банках X, млн.руб	№ п/п	Индекс потребительских цен, Y, %	Индекс официального курса рубля, X, %

1	14,31	6,55	1	9,25	1,98	1	103,2	112,6
2	16,91	7,03	2	10,60	2,20	2	102,2	98,3
3	17,67	9,36	3	11,00	2,26	3	100,6	100,4
4	19,66	9,55	4	12,04	3,02	4	101,3	101,4
5	21,34	10,81	5	12,91	3,38	5	101,3	100,5
6	23,02	11,49	6	13,79	4,95	6	100,6	101,6
7	24,70	11,66	7	14,66	6,24	7	100,6	100,7
8	26,38	13,40	8	15,54	7,07	8	100,5	100,7
9	28,06	15,15	9	16,41	7,33	9	102,2	101,1
10	29,74	16,90	10	17,29	8,09	10	101,4	101,3
11	31,42	18,64	11	18,16	8,86	11	101,4	101,2
12	33,10	20,32	12	19,04	9,63	12	102,4	101,7
13	34,78	22,26	13	19,91	10,39	13	102,3	101,5
14	36,46	23,81	14	20,79	11,15	14	100,7	104,5
15	37,92	25,63	15	27,84	11,92	15	100,7	106,0
16	39,62	27,37	16	28,82	12,68	16	102,3	101,1
17	41,31	29,12	17	29,80	12,69	17	100,5	101,3
18	43,01	32,05	18	30,78	13,50	18	100,5	101,2
19	44,71	33,82	19	31,76	14,30	19	99,6	104,1
20	46,41	35,58	20	32,74	15,11	20	100,7	105,9
21	48,10	37,34	21	33,72	15,92	21	104,9	130,1
22	49,80	39,11	22	34,70	16,73	22	107,3	125,3
23	51,50	40,87	23	35,68	17,53	23	104,1	100,9
24	53,20	42,64	24	36,66	18,34	24	104,4	100,8
25	54,89	44,40	25	37,64	19,15	25	102,5	100,8
26	56,59	46,16	26	38,62	19,95	26	101,5	103,3
27	58,29	47,93	27	39,60	20,76	27	101,5	110,9
28	59,99	49,69	28	40,58	21,57	28	103,3	106,4
29	61,68	51,46	29	41,56	22,38	29	103,4	100,5
30	62,44	51,78	30	42,00	23,18	30	100,6	101,5

Для выполнения работы следует изучить теоретический материал, представленный в темах 7, 8 Учебного пособия по эконометрике [18].

РЕШЕНИЕ ТИПОВЫХ ЗАДАЧ

Пример 1. Тестирование автокорреляции остатков

1) Метод Дарбина-Уотсона

В табл. 5.2 (см. пример1 темы 5) представлены данные о добавленной стоимости в обрабатывающей промышленности и ВВП (млн. долл. США) для межстрановой выборки. Построена модель зависимости добавленной стоимости (Y) от ВВП (X), на основе которой найдены расчетные значения рассматриваемого показателя и остатки (см. рис. 5.2).

Проверим наличие автокорреляции остатков (табл. 6.2), в столбцах 5 и 6 которой приведен расчет составляющих d-статистики Дарбина-Уотсона.

Расчетные значения зависимой переменной, представленные в 3-м столбце таблицы, получены с использованием функции Excel ТЕНДЕНЦИЯ.

$$d = \frac{\sum_{t=2}^{18} (e_t - e_{t-1})^2}{\sum_{t=1}^{18} e_t^2} = \frac{34173896128}{14312086755} = 2,388. \quad 4 - d = 4 - 2,388 = 1,612.$$

Таблица 6.2 – Расчет d-статистики

Y	X	Y расчетн.	Остатки (e_t)	$(e_t - e_{t-1})^2$	e_t^2
44517	232006	45541,85	-1024,85		1050320,343
112617	547203	106593,4	6023,645	49681297,9	36284296,28
13096	50919	10466,54	2629,462	11520473,9	6914072,559
25927	151266	29903,07	-3976,07	43632997,7	15809096,11
21581	97624	19512,98	2068,023	36531007,5	4276720,108
256316	1330998	258409,1	-2093,09	17314845,3	4381016,64
9392	98861	19752,58	-10360,6	68351347,6	107341518,6
11758	130823	25943,4	-14185,4	14629254,7	201225461,9
7227	41506	8643,304	-1416,3	163049716	2005916,593
17572	52662	10804,14	6767,855	66980459,8	45803864,36
11349	74121	14960,61	-3611,61	107733223	13043700,47
145013	1016286	197451,5	-52438,5	2384067999	2749798920
161318	380820	74366,11	86951,89	1,943E+10	7560631920
2797	24848	5416,763	-2619,76	8023081814	6863158,921
18874	72505	14647,6	4226,402	46869974,3	17862472,13
55073	420788	82107,63	-27034,6	977252362	730871438,1
48595	334286	65352,79	-16757,8	105613563	280823461,1
13484	122926	24413,8	-10929,8	33965431,4	119460553,6
			Сумма	34173896128	14312086755

Из таблицы Дарбина-Уотсона при уровне значимости 0,05, задавшись количеством наблюдений $n = 28$ и числом независимых переменных $k = 1$, выбираем граничные значения: $d_L = 1,33$, $d_u = 1,48$.

Поскольку $4 - d = 1,612 > d_u = 1,48$, то с вероятностью 0,95 можно утверждать, что автокорреляции остатков, полученных по модели, оцененной методом наименьших квадратов, не существует.

2) Применение теста серий (Бреуша – Годфри).

В соответствии с тестом Бреуша-Годфри оценим авторегрессионную зависимость остатков от трех предыдущих значений (см. рис. 6.1).

Выполненные расчеты показывают, что коэффициенты при всех лаговых переменных статистически не значимы, поскольку величина всех t -статистик оказывается меньше критического $t_{5\%, v=22} = 2,074$, то есть, существенного влияния на результат наблюдения e_t не оказывает ни одно из предыдущих значение e_{t-i} .

={ЛИНЕЙН(D102:D126;E102:G126;0;1)}					
	D	E	F	G	H
	e_t	e_{t-1}	e_{t-2}	e_{t-3}	
	-1024,851				
	6023,645	-1024,851			
	2629,462	6023,645	-1024,851		
	-3976,065	2629,462	6023,645	-1024,851	
	2068,023	-3976,065	2629,462	6023,645	
	-2093,088	2068,023	-3976,065	2629,462	
	-10360,575	-2093,088	2068,023	-3976,065	
	-14185,396	-10360,575	-2093,088	2068,023	
	-1416,304	-14185,396	-10360,575	-2093,088	
	6767,855	-1416,304	-14185,396	-10360,575	
	-3611,606	6767,855	-1416,304	-14185,396	
	-52438,525	-3611,606	6767,855	-1416,304	
	86951,894	-52438,525	-3611,606	6767,855	
	-2619,763	86951,894	-52438,525	-3611,606	
	4226,402	-2619,763	86951,894	-52438,525	
	-27034,634	4226,402	-2619,763	86951,894	
	-16757,788	-27034,634	4226,402	-2619,763	
	-10929,801	-16757,788	-27034,634	4226,402	
	-498,361	-10929,801	-16757,788	-27034,634	
	6284,356	-498,361	-10929,801	-16757,788	
	-546,382	6284,356	-498,361	-10929,801	
	1129,655	-546,382	6284,356	-498,361	
	-14179,961	1129,655	-546,382	6284,356	
	-4232,797	-14179,961	1129,655	-546,382	
	6270,056	-4232,797	-14179,961	1129,655	
	-5955,226	6270,056	-4232,797	-14179,961	
	4176,408	-5955,226	6270,056	-4232,797	
	45333,367	4176,408	-5955,226	6270,056	
	ЛИНЕЙН				
		-0,21246016	-0,2188084	-0,38402507	
		0,22867015	0,241538797	0,228634328	
		0,13275545	23715,84125	#Н/Д	
		1,12256685	22	#Н/Д	
		1894133289	12373704776	#Н/Д	
	t=	-0,92911191	-0,90589338	-1,67964749	
		$e_t = -0,384e_{t-1} - 0,219e_{t-2} - 0,212e_{t-3}$			

Рисунок 6.1 – Результаты расчетов

Таким образом, тест Бреуша-Годфри подтверждает вывод, сделанный с помощью теста Дарбина-Уотсона об отсутствии автокорреляции остатков в модели.

Пример 2. Устранение автокорреляции остатков

Используя данные табл. 6.3 об экономических показателях, построить эконометрическую модель и выполнить проверку остатков на автокорреляцию. Если остатки окажутся автокоррелированными, построить модель с эффективными оценками.

Таблица 6.3- Исходные данные

№ п/п	Инвестиции, ден.ед	ВВП, ден.ед	Процентная ставка, %			Расчеты по критерию Дарбина-Уотсона		Данные для расчетов по критерию Бреуша- Годфри		
	У	X1	X2	У расч	е	$(e_t - e_{t-1})^2$	e^2	e_{t-1}	e_{t-2}	e_{t-3}
1	2	3	4	5	6	7	8	9	10	11
1	11,55	8,58	18,12	9,49	2,06		4,26			
2	13,25	10,45	11,05	12,24	1,01	1,10	1,03	2,06		
3	10,9	8,35	9	11,00	-0,10	1,24	0,01	1,01	2,06	
4	10,45	10,65	17	11,29	-0,84	0,54	0,70	-0,10	1,01	2,06
5	15,1	9,7	16,25	10,69	4,41	27,48	19,41	-0,84	-0,10	1,01
6	17,5	12	13,8	12,92	4,58	0,03	20,99	4,41	-0,84	-0,10
7	17,77	13,45	19,95	12,89	4,88	0,09	23,77	4,58	4,41	-0,84
8	16,1	14,2	18,74	13,70	2,40	6,11	5,78	4,88	4,58	4,41
9	10,59	14,45	13,8	14,80	-4,21	43,79	17,75	2,40	4,88	4,58
10	10,65	13,85	9,55	15,13	-4,48	0,07	20,06	-4,21	2,40	4,88
11	9,32	16,55	19,3	15,40	-6,08	2,56	36,96	-4,48	-4,21	2,40
12	11	18	15,2	17,27	-6,27	0,04	39,37	-6,08	-4,48	-4,21
13	15,05	18,4	12,4	18,10	-3,05	10,39	9,31	-6,27	-6,08	-4,48
14	15,1	20,4	16,5	18,88	-3,78	0,53	14,29	-3,05	-6,27	-6,08
15	22,7	21	5,95	21,30	1,40	26,87	1,97	-3,78	-3,05	-6,27
16	21,65	23,75	17,5	21,27	0,38	1,05	0,14	1,40	-3,78	-3,05
17	23,1	25,75	16,43	23,01	0,09	0,08	0,01	0,38	1,40	-3,78
18	25,65	24,2	7,4	23,49	2,16	4,28	4,67	0,09	0,38	1,40
19	26,15	25,2	15,45	22,77	3,38	1,49	11,44	2,16	0,09	0,38
20	25,55	26,2	19,15	22,85	2,70	0,47	7,28	3,38	2,16	0,09
21	28,1	28,6	5,45	27,24	0,86	3,37	0,75	2,70	3,38	2,16
22	24,2	30,6	9,52	28,02	-3,82	21,94	14,60	0,86	2,70	3,38
23	32,3	31,32	7,95	28,87	3,43	52,64	11,80	-3,82	0,86	2,70
24	21,5	26	7,45	24,87	-3,37	46,24	11,33	3,43	-3,82	0,86
25	22,95	26,85	19,9	23,21	-0,26	9,63	0,07	-3,37	3,43	-3,82
26	30,45	32,1	8,65	29,34	1,11	1,89	1,24	-0,26	-3,37	3,43

27	24,6	32,95	21,35	27,64	-3,04	17,23	9,22	1,11	-0,26	-3,37
28	32,5	33,3	11,11	29,80	2,70	32,87	7,27	-3,04	1,11	-0,26
29	31,2	33,85	15,82	29,35	1,85	0,72	3,41	2,70	-3,04	1,11
30	29,5	35,6	21,67	29,62	-0,12	3,85	0,01	1,85	2,70	-3,04
				Сумма		318,6	298,9			

Решение

- 1) Методом наименьших квадратов построили линейную регрессионную модель зависимости инвестиций от ВВП и процентной ставки (исходные данные – столбцах 2 – 4 табл. 6.3):

$$\hat{Y} = 6,2437 + 0,7693X_1 - 0,1853X_2$$

ЛИНЕЙН

-0,1853	0,7693	6,2437
0,1264	0,0717	2,5110
0,8164	3,3272	#Н/Д
60,0100	27	#Н/Д
1328,6552	298,8975	#Н/Д
t= -1,4661	10,7233	2,4866

Подставив в модель X_1 и X_2 , получили расчетные значения \hat{Y} (столбец 5) и остатки (столбец 6) как разность между фактическими и расчетными Y .

- 2) Проверим остатки на автокорреляцию с помощью критерия Дарбина-Уотсона. Необходимые для вычисления d-статистики расчеты приведены в столбцах 7 и 8 табл. 6.3.

$$d = \frac{\sum_{t=2}^{30} (e_t - e_{t-1})^2}{\sum_{t=1}^{30} e_t^2} = \frac{318,6}{298,9} = 1,066$$

Из таблицы Дарбина-Уотсона при уровне значимости 0,05, задавшись количеством наблюдений $n = 30$ и числом независимых переменных $k = 2$, выбираем граничные значения: $d_L = 1,284$, $d_u = 1,567$.

Поскольку $d = 1,066 < d_L = 1,284$, то с вероятностью 0,95 можно утверждать, что имеется положительная автокорреляция остатков.

- 3) Выполним проверку автокорреляции остатков, применяя критерий Бреуша-Годфри в более общем виде, чем в примере 5.1 (рис. 6.2).

Для тестирования по этому тесту было выбрано 2 лаговых значения для остатков, т. е. тестировалась автокорреляция до 2-го порядка (из-за малого числа наблюдений больший порядок выбирать не целесообразно).

A36		fx {=ЛИНЕЙН(В4:В31;С4:Ф31;1;1)}					G	H	I
	A	B	C	D	E	F			
1		e	X1	X2	et-1	et-2			
2	1	2,06	8,58	18,12					
3	2	1,01	10,45	11,05	2,06				
4	3	-0,10	8,35	9	1,01	2,06			
5	4	-0,84	10,65	17	-0,10	1,01			
6	5	4,41	9,7	16,25	-0,84	-0,10			
7	6	4,58	12	13,8	4,41	-0,84			
8	7	4,88	13,45	19,95	4,58	4,41			
9	8	2,40	14,2	18,74	4,88	4,58			
10	9	-4,21	14,45	13,8	2,40	4,88			
11	10	-4,48	13,85	9,55	-4,21	2,40			
12	11	-6,08	16,55	19,3	-4,48	-4,21			
13	12	-6,27	18	15,2	-6,08	-4,48			
14	13	-3,05	18,4	12,4	-6,27	-6,08			
15	14	-3,78	20,4	16,5	-3,05	-6,27			
16	15	1,40	21	5,95	-3,78	-3,05			
17	16	0,38	23,75	17,5	1,40	-3,78			
18	17	0,09	25,75	16,43	0,38	1,40			
19	18	2,16	24,2	7,4	0,09	0,38			
20	19	3,38	25,2	15,45	2,16	0,09			
21	20	2,70	26,2	19,15	3,38	2,16			
22	21	0,86	28,6	5,45	2,70	3,38			
23	22	-3,82	30,6	9,52	0,86	2,70			
24	23	3,43	31,32	7,95	-3,82	0,86			
25	24	-3,37	26	7,45	3,43	-3,82			
26	25	-0,26	26,85	19,9	-3,37	3,43			
27	26	1,11	32,1	8,65	-0,26	-3,37			
28	27	-3,04	32,95	21,35	1,11	-0,26			
29	28	2,70	33,3	11,11	-3,04	1,11			
30	29	1,85	33,85	15,82	2,70	-3,04			
31	30	-0,12	35,6	21,67	1,85	2,70			
32									
33	et-2	et-1	X2	X1	X0				
34	0,143236	0,408226	-0,06433	0,019019	0,402367		Chi2 табл =	5,991464547	
35	0,205678	0,20808	0,122577	0,072855	2,516001		nR^2=28*0,238 =	6,670115735 > 5,99	
36	0,238218	3,116643	#Н/Д	#Н/Д	#Н/Д				
37	1,798095	23	#Н/Д	#Н/Д	#Н/Д		t табл =	2,06865761	
38	69,86294	223,4097	#Н/Д	#Н/Д	#Н/Д		p =	0,061990909	
39	0,69641	1,961864	-0,52479	0,261059	0,159923	t	1 - p =	0,938009091	
40									

Рисунок 6.2 – Результаты расчетов

На рис. 6.5 показано построение регрессии остатков e_t (см. табл. 6.2) на X_1 , X_2 , e_{t-1} и e_{t-2} методом наименьших квадратов, реализованного с помощью функции ЛИНЕЙН. Здесь цветом выделены наблюдения, которые оказались потерянными.

Коэффициент детерминации $R^2 = 0,238$.

Умножив коэффициент детерминации на $n = 28$, получим 7,1466 (после сдвигов ряда остатков число активных наблюдений составляет 28). Критическое (табличное) значение статистики хи-квадрат при уровне значимости 5% и числе степеней свободы $k = 2$ составляет 5,99. Поскольку вычисленное значение 6,67 превышает табличное, то нулевую гипотезу об отсутствии автокорреляции с вероятностью 0,95 отклоняем.

Кроме того, выполненные расчеты (рис. 6.2) показывают, что коэффициент при лаговой переменной e_{t-2} статистически не значим, поскольку величина t-статистики оказывается меньше критического $t_{5\%, v=23} = 2,069$. t-статистика, рассчитанная для коэффициента при лаговой переменной e_{t-1} , равная 1,962, тоже меньше $t_{5\%, v=23} = 2,069$. С помощью функции СТЬЮДЕНТ.РАСП.2Х для значения 1,962 найденное распределение вероятностей, равное $p = 0,062$, характеризует вероятность незначимости коэффициента, тогда с вероятностью $1 - p = 0,938$ можем считать коэффициент при e_{t-1} статистически значимым. То есть, существенное влияние на результат наблюдения e_t оказывает предыдущее значение e_{t-1} , или, другими словами, имеется автокорреляция первого порядка.

Убедившись, что регрессионные остатки автокоррелированы, построим модель с применением **обобщенного метода наименьших квадратов** для получения эффективных оценок.

Вначале проверим, остатки на гетероскедастичность, используя графический метод. На рис. 6.3 представлены графики остатков по каждой из объясняющих переменных модели, которые не дают основания подозревать существование гетероскедастичности. Поэтому для устранения автокоррелированности остатков можно воспользоваться обобщенным методом наименьших квадратов.

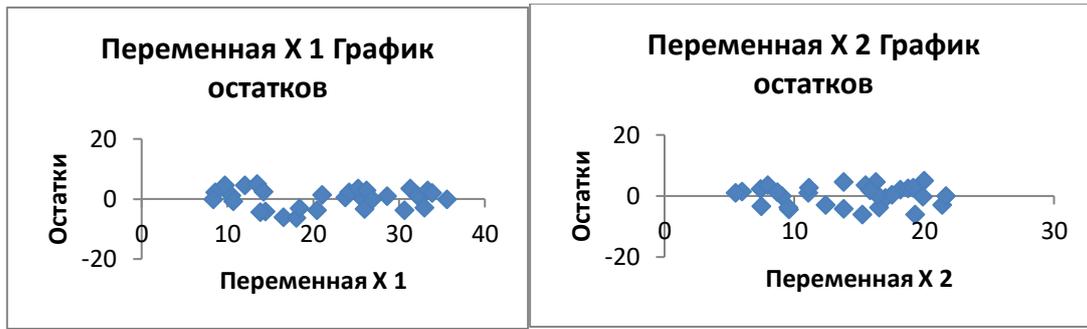


Рис.6.3. Графическое представление регрессионных остатков

При использовании теста Дарбина-Уотсона (см. пример 5.3.2) была рассчитана статистика $d = 1,066$, на основе которой рассчитаем оценку ρ :

$$\rho \approx 1 - \frac{d}{2} = 1 - \frac{1,066}{2} = 0,467$$

Выполним преобразование исходных данных:

$$Y_{np} = P^{-1}Y = \begin{pmatrix} \sqrt{1-\rho^2} Y_1 \\ Y_2 - \rho Y_1 \\ Y_3 - \rho Y_2 \\ \dots \\ Y_n - \rho Y_{n-1} \end{pmatrix}; \quad X_{np} = P^{-1}X = \begin{pmatrix} \sqrt{1-\rho^2} & \sqrt{1-\rho^2} X_{11} & \dots & \sqrt{1-\rho^2} X_{1p} \\ 1-\rho & X_{21} - \rho X_{11} & \dots & X_{2p} - \rho X_{1p} \\ \dots & \dots & \dots & \dots \\ 1-\rho & X_{n1} - \rho X_{n-1,1} & \dots & X_{np} - \rho X_{n-1,p} \end{pmatrix}$$

Фрагмент таблицы исходных данных, приведенных в табл. 6.3, представлен в табл. 6.4.

Таблица 6.4– Исходные данные

№	Y	X1	X2	№	Y	X1	X2
1	11,55	8,58	18,12	3	10,9	8,35	9
2	13,25	10,45	11,05	4	10,45	10,65	17

Расчеты, направленные на получение преобразованных данных, выполним для данных табл. 6.4.

$$Y_{np} = \begin{pmatrix} \sqrt{1-0,467^2} \cdot 11,55 \\ 13,25 - 0,467 \cdot 11,55 \\ 10,9 - 0,467 \cdot 13,25 \\ 10,45 - 0,467 \cdot 10,9 \\ \dots \end{pmatrix} = \begin{pmatrix} 10,213 \\ 7,856 \\ 4,712 \\ 5,360 \\ \dots \end{pmatrix}$$

$$X_{np} = \begin{pmatrix} \sqrt{1-0,467^2} & \sqrt{1-0,467^2} \cdot 8,58 & \sqrt{1-0,467^2} \cdot 18,12 \\ 1-0,467 & 10,45-0,467 \cdot 8,58 & 11,05-0,467 \cdot 18,12 \\ 1-0,467 & 8,35-0,467 \cdot 10,45 & 9-0,467 \cdot 11,05 \\ 1-0,467 & 10,65-0,467 \cdot 8,35 & 17-0,467 \cdot 9 \end{pmatrix} = \begin{pmatrix} 0,884 & 7,587 & 16,022 \\ 0,533 & 6,443 & 2,587 \\ 0,533 & 3,469 & 3,840 \\ 0,533 & 6,750 & 12,797 \end{pmatrix}$$

Результаты полного преобразования и построения новой модели на основе преобразованных данных – на рис. 6.4.

B59		fx {=ЛИНЕЙН(A26:A55;C26:D55;1;1)}						
	A	B	C	D	E	F	G	H
23					Проверка остатков новой модели			
24	Преобразование данных				на автокорреляцию			
25	У	X0	X1	X2	У расч	e(t)	(e(t)-e(t-1))^2	e^2
26	10,213	0,884	7,587	16,022	5,801	4,412		19,468
27	7,856	0,533	6,443	2,587	8,089	-0,233	21,582	0,054
28	4,712	0,533	3,469	3,839	5,482	-0,771	0,289	0,594
29	5,359	0,533	6,750	12,797	5,913	-0,554	0,047	0,307
30	10,219	0,533	4,726	8,310	5,401	4,818	28,861	23,213
31	10,448	0,533	7,470	6,211	8,030	2,417	5,763	5,844
32	9,597	0,533	7,845	13,505	6,597	3,000	0,339	8,998
33	7,801	0,533	7,918	9,422	7,619	0,181	7,942	0,033
34	3,071	0,533	7,818	5,048	8,576	-5,505	32,339	30,308
35	5,704	0,533	7,101	3,105	8,478	-2,774	7,458	7,697
36	4,346	0,533	10,081	14,840	8,019	-3,673	0,808	13,492
37	6,647	0,533	10,270	6,186	10,213	-3,565	0,012	12,712
38	9,913	0,533	9,993	5,301	10,206	-0,294	10,703	0,086
39	8,071	0,533	11,806	10,709	10,337	-2,266	3,888	5,133
40	15,648	0,533	11,472	-1,756	13,025	2,623	23,896	6,879
41	11,048	0,533	13,942	14,721	11,047	0,001	6,875	0,000
42	12,989	0,533	14,658	8,257	13,132	-0,144	0,021	0,021
43	14,861	0,533	12,174	-0,274	13,219	1,642	3,189	2,696
44	14,170	0,533	13,898	11,994	11,658	2,513	0,758	6,313
45	13,337	0,533	14,430	11,934	12,086	1,251	1,592	1,564
46	16,167	0,533	16,363	-3,494	17,237	-1,070	5,386	1,145
47	11,076	0,533	17,243	6,975	15,444	-4,368	10,879	19,082
48	20,998	0,533	17,028	3,504	16,099	4,899	85,877	23,997
49	6,414	0,533	11,372	3,737	11,648	-5,233	102,661	27,389
50	12,909	0,533	14,707	16,421	11,240	1,669	47,640	2,785
51	19,731	0,533	19,560	-0,644	19,047	0,684	0,969	0,468
52	10,379	0,533	17,958	17,310	13,556	-3,177	14,913	10,096
53	21,011	0,533	17,911	1,139	17,344	3,667	46,843	13,445
54	16,021	0,533	18,297	10,631	15,399	0,622	9,272	0,386
55	14,928	0,533	19,791	14,281	15,697	-0,768	1,932	0,590
56		ЛИНЕЙН				Сумма	482,732	244,796
57		-0,236499431	0,77714321	3,693961104		d=	1,972	> du=1,567
58		0,095094691	0,11918283	1,692593602			482,737/244,796=1,972	
59		0,641385207	3,01106879	#Н/Д				
60		24,14484974	27	#Н/Д				
61		437,8202624	244,796452	#Н/Д				
62	t =	-2,486988797	6,52059679	2,182426485				

Рис. 6.4. Результаты расчетов

Из рис. 6.4 видно, что при построении модели по преобразованным данным с помощью обычного МНК не использованы данные преобразованного столбца X_0 . Это связано с тем, что для расчетов применялась функция ЛИНЕЙН.

В примере 2 (Лабораторная работа №5) методом наименьших квадратов построили линейную регрессионную модель зависимости инвестиций от ВВП и процентной ставки по исходным данным вида:

$$\hat{Y} = 6,2437 + 0,7693X_1 - 0,1853X_2.$$

Модель по преобразованным данным (см. рис. 6.4):

$$\hat{Y} = 3,694 + 0,7771X_1 - 0,2365X_2.$$

Проверка остатков этой модели по критерию Дарбина–Уотсона показала, что автокорреляция отсутствует (см. рис. 6.4), поскольку вычисленная d -статистика, равная 1,972, больше верхней границы табличного значения $d_u=1,567$. То есть, коэффициенты при переменных X_1 и X_2 являются эффективными оценками соответствующих истинных значений параметров, и точность прогнозирования инвестиций в зависимости от ВВП и процентной ставки выше, чем при использовании модели, построенной на основе исходных данных.

Пример 3. Применение процедуры Кохрейна – Оркатта для оценивания модели с автокоррелированными остатками

Изучается зависимость среднедушевых расходов на конечное потребление y от среднедушевого дохода x по данным некоторой страны за 16 лет. Исходные (y_t, x_t) и расчетные представлены в табл. 6.5:

Таблица 6.5 – Среднедушевые расходы на конечное потребление (y) и среднедушевые доходы (x)

t	y_t	x_t	t	y_t	x_t
1	70	73	9	103	113
2	73	76	10	109	119
3	78	83	11	112	121
4	83	89	12	114	122
5	86	95	13	115	131
6	89	100	14	118	135
7	96	107	15	122	139
8	96	108	16	123	140

Пусть исходная модель имеет вид $y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$.

1. Применяя МНК к исходным данным (табл. 6.5), получим уравнение регрессии (см. рис. 6.5):

$$\hat{y}_t = 10,987 + 0,806x_t, \quad R^2 = 0,986$$

(2,77) (0,025)

Получили первоначальные оценки параметров β_0, β_1 : $b_0 = 10,987$, $b_1 = 0,806$, в скобках указаны стандартные ошибки оценок.

2. Найдем регрессионные остатки (как разность между фактическими и расчетными значениями Y) и с помощью критерия Дарбина-Уотсона проверим нулевую гипотезу об отсутствии автокорреляции (рис. 6.5).

E88		fx		=КОРРЕЛ(E71:E85;E72:E86)						
	A	B	C	D	E	F	G	H	I	J
67										
68										
69										
70	t	y_t	x_t	y_{rach}	e_t	$(e_t - e_{t-1})^2$	e_t^2	e_{t-1}	$e_t e_{t-1}$	e_{t-1}^2
71	1	70	73	69,821	0,179		0,032			
72	2	73	76	72,239	0,761	0,339	0,580	0,179	0,136	0,032
73	3	78	83	77,880	0,120	0,412	0,014	0,761	0,091	0,580
74	4	83	89	82,716	0,284	0,027	0,081	0,120	0,034	0,014
75	5	86	95	87,552	-1,552	3,370	2,408	0,284	-0,441	0,081
76	6	89	100	91,581	-2,581	1,060	6,664	-1,552	4,006	2,408
77	7	96	107	97,223	-1,223	1,845	1,496	-2,581	3,157	6,664
78	8	96	108	98,029	-2,029	0,650	4,117	-1,223	2,481	1,496
79	9	103	113	102,059	0,941	8,823	0,886	-2,029	-1,910	4,117
80	10	109	119	106,894	2,106	1,356	4,434	0,941	1,982	0,886
81	11	112	121	108,506	3,494	1,927	12,206	2,106	7,357	4,434
82	12	114	122	109,312	4,688	1,426	21,976	3,494	16,378	12,206
83	13	115	131	116,566	-1,566	39,106	2,451	4,688	-7,340	21,976
84	14	118	135	119,789	-1,789	0,050	3,202	-1,566	2,802	2,451
85	15	122	139	123,013	-1,013	0,603	1,027	-1,789	1,813	3,202
86	16	123	140	123,819	-0,819	0,038	0,671	-1,013	0,830	1,027
87					Сумма	61,030	62,243		31,377	61,572
88	ЛИНЕЙН			r =	0,507	d =	0,980	r =	0,509596	
89		0,806	10,987							
90		0,025	2,772							
91		0,987	2,109							
92		1050,435	14,000							
93		4670,194	62,243							
94										

$$r = \frac{\sum_{i=2}^n e_i e_{i-1}}{\sum_{i=2}^n e_{i-1}^2} = \frac{31,377}{61,572} = 0,5096$$

Рис. 6.5. Результаты расчетов

Вычисленное значение статистики Дарбина-Уотсона $d = 0,980$. При уровне значимости 5% для $n = 16$ табличное значение $d_L = 1,106$ и $d_U = 1,371$. Поскольку $d < d_L$, то имеется положительная автокорреляция остатков.

d -статистику можно найти через коэффициент автокорреляции остатков, воспользовавшись приближительной формулой $d \approx 2(1 - r)$. Коэффициент автокорреляции остатков первого порядка составляет $r = 0,507$ (рис. 6.5).

Следовательно, $d = 2(1 - 0,507) = 0,986$. Видим, что отличие от ранее полученного значения невелико.

Итак, имеется положительная автокорреляция остатков

3. Для преобразования модели найдем значения оценки коэффициента ρ по формуле:

$$r = \frac{\sum_{i=2}^n e_i e_{i-1}}{\sum_{i=2}^n e_{i-1}^2},$$

которое равно 0,5096 (см. рис. 6.5).

4. С использованием найденного $r = 0,5096$ найдем оценки параметров уравнения

$$y_t - r y_{t-1} = \alpha_0(1 - r) + \alpha_1(x_t - r x_{t-1}) + (\varepsilon_t - r \varepsilon_{t-1}),$$

методом наименьших квадратов, который будем применять к преобразованным данным (рис. 6.6).

Выполним преобразования исходных данных по формулам:

$$y_t^* = y_t - \rho y_{t-1} = y_t - 0,5096 y_{t-1}$$

$$x_t^* = x_t - \rho x_{t-1} = x_t - 0,5096 x_{t-1} \quad (t \geq 2),$$

(результат – в столбцах F и G рис. 6.6)

Применяя МНК к преобразованным данным, получим оценку преобразованного уравнения (рис. 6.6):

$$\hat{y}_t^* = 6,209 + 0,791 x_t^*, \quad R^2 = 0,948$$

(2,986) (0,052)

(в скобках указаны стандартные ошибки коэффициентов).

F124		fx {=ЛИНЕЙН(F102:F116;G102:G116;1;1)}									
	A	B	C	D	E	F	G	H	I	J	K
100	t	y _t	x _t	y _{t-1}	x _{t-1}	y _t *	x _t *	y _t *rach	e _t *	(e _t *-e _t *(t-1))^2	e _t *t^2
101	1	70	73								
102	2	73	76	70	73	37,328	38,799	36,894	0,434		0,188
103	3	78	83	73	76	40,799	44,270	41,221	-0,422	0,733	0,178
104	4	83	89	78	83	43,251	46,703	43,146	0,106	0,279	0,011
105	5	86	95	83	89	43,703	49,646	45,473	-1,769	3,516	3,131
106	6	89	100	86	95	45,174	51,588	47,009	-1,834	0,004	3,365
107	7	96	107	89	100	50,646	56,040	50,530	0,116	3,803	0,013
108	8	96	108	96	107	47,078	53,473	48,500	-1,421	2,362	2,020
109	9	103	113	96	108	54,078	57,963	52,051	2,028	11,893	4,111
110	10	109	119	103	113	56,511	61,415	54,781	1,730	0,088	2,994
111	11	112	121	109	119	56,454	60,358	53,945	2,509	0,607	6,295
112	12	114	122	112	121	56,925	60,338	53,929	2,995	0,237	8,972
113	13	115	131	114	122	56,906	68,829	60,644	-3,739	45,348	13,978
114	14	118	135	115	131	59,396	68,242	60,181	-0,785	8,727	0,615
115	15	122	139	118	135	61,867	70,204	61,732	0,135	0,846	0,018
116	16	123	140	122	139	60,829	69,166	60,911	-0,082	0,047	0,007
117									сумма	78,490	45,897
118						$y_t^* = y_t - \rho y_{t-1} = y_t - 0,5096 y_{t-1}$					d = 1,710
119						$x_t^* = x_t - \rho x_{t-1} = x_t - 0,5096 x_{t-1} \quad (t \geq 2),$					
120						ЛИНЕЙН					
121											
122						0,791	6,209				
123						0,052	2,986				
124						0,948	1,879				
125						235,154	13,000				
126						830,221	45,897				
127					t =	15,33474	2,079001				
128											

Рис. 6.6. Результаты расчетов

На основе этого уравнения нашли остатки (столбец I рис. 6.6):

$$y_t^* - y_{rach}^* = e_t^*$$

и проверили их на автокорреляцию: $d = 1,71$. При уровне значимости 5% для $n = 15$ табличное значение $d_L = 1,077$ и $d_U = 1,361$. Поскольку $d > d_U$, то автокорреляция остатков отсутствует. (В преобразованной модели наблюдений на одно меньше, чем в исходной, так как при преобразовании используются запаздывающие значения переменных).

Таким образом, процесс преобразований завершен. В полученной модели коэффициент при $x_t^* = x_t - rx_{t-1}$ является оценкой параметра β_1 , а свободный член, деленный на $(1 - r)$, оценивает параметр β_0 .

Пересчитывая оценки параметров модели в соответствии с используемым методом.

$$b_0 = \frac{a}{1 - r} = \frac{6,209}{1 - 0,5096} = 12,66.$$

Окончательно имеем:

$$\hat{y}_t = 12,66 + 0,791x_t.$$

Это уравнение отличается от полученного ранее уравнения, оцененного обычным МНК.

Хотя в преобразованной модели коэффициент детерминации ($R^2 = 0,948$) ниже, чем в непреобразованной ($R^2 = 0,987$), значение d-статистики теперь свидетельствует об отсутствии автокорреляции остатков.

Следует отметить, что стандартные ошибки оценок преобразованной модели ($s_0 = 2,99$; $s_1 = 0,052$) выше, чем в исходной ($s_0 = 2,77$; $s_1 = 0,025$), что подтверждает сделанное ранее замечание о занижении стандартных ошибок в моделях с автокоррелированными остатками и необоснованных выводах о статистической значимости результатов оценивания.

Возрастание стандартных ошибок приводит к увеличению ширины доверительного интервала для коэффициентов модели. Так, при оценивании исходной линейной модели 95%-й доверительный интервал для параметров β_0 и β_1 имеет вид:

$$10,987 - 2,772 \cdot 2,145 < \beta_0 < 10,987 + 2,772 \cdot 2,145$$

$$5,042 < \beta_0 < 16,932;$$

$$0,806 - 0,025 \cdot 2,145 < \beta_1 < 0,806 + 0,025 \cdot 2,145$$

$$0,753 < \beta_1 < 0,859$$

где $t_{0.05, v=14} = 2,145$.

Для преобразованной модели

$$6,209 - 2,986 \cdot 2,16 < \beta_0 < 6,209 + 2,986 \cdot 2,16,$$

$$-0,243 < \beta_0 < 12,66;$$

$$0,791 - 0,0525 \cdot 2,16 < \beta_1 < 0,791 + 0,052 \cdot 2,16,$$

$$0,679 < \beta_1 < 0,902,$$

где $t_{0,05, v=13} = 2,16$.

Таким образом, видим, что вывод о статистической надежности свободного члена в исходной модели не верен, так как после корректировки модели и устранения автокорреляции границы доверительного интервала имеют разные знаки, что не позволяет отклонить гипотезу о равенстве нулю истинного значения параметра β_0 .

Вопросы и задания для самостоятельной работы

- 1) В чем суть проблемы автокорреляции остатков?
- 2) Что означает положительная и отрицательная автокорреляция остатков?
- 3) Какое влияние автокорреляции остатков оказывает на оценки параметров модели?
- 4) Каковы основные методы тестирования автокорреляции?
- 5) Как выполняется проверка с помощью критерия Дарбина-Уотсона?
- 6) Как можно найти оценки регрессионных коэффициентов в случае линейной модели с коррелированными остатками?
- 7) Как реализуется обобщенный метод наименьших квадратов (метод Эйткена) для эконометрической модели с автокоррелированными остатками.

Задание 1.[19] Рассмотрите модель, связывающую количество вакансий w_t и уровень безработицы u_t :

$$\ln w_t = \beta_1 + \beta_2 \ln u_t + \varepsilon_t.$$

Возмущения ε_t независимы и нормально распределены: $N(0, \sigma_\varepsilon^2)$.

Задание .

- 1) Используя (искусственные) данные табл. 6.6, найдите МНК-оценки параметров β_1 и β_2 , а также 95%-доверительный интервал для β_2 .
- 2) Вычислите статистику Дарбина-Уотсона. Что ее значение говорит об исходном предположении об ε_t ? Что можно сказать о доверительном интервале, найденном в п.1)?
- 3) Оцените модель заново (используй метод Эйткена или авторегрессионные преобразования). Найдите 95%-й доверительный интервал для β_2 . Сравните результат с интервалом, полученным в п.1).

Таблица 6. 6 – Исходные данные

t	w_t	u_t	t	w_t	u_t	t	w_t	u_t
1	1,73	8,65	9	5,06	2,87	17	3,15	4,72
2	1,94	4,82	10	2,81	5,29	18	1,92	7,45
3	3,05	2,67	11	4,43	3,31	19	2,26	6,21
4	4,17	2,67	12	3,19	5,44	20	6,18	2,64
5	2,52	2,58	13	2,23	6,80	21	2,07	8,55
6	1,71	8,07	14	2,06	8,25	22	8,39	2,60
7	1,95	8,83	15	3,33	3,44	23	2,75	6,25
8	2,57	5,54	16	2,12	7,80	24	6,10	2,70

Задание 2. [19] Для объяснения изменения ВВП за 10 лет строится регрессионная модель с объясняющими переменными – потреблением (С) и инвестициями (I). Получены следующие статистические данные (табл. 6.7):

Таблица 6.7 – Исходные данные

С, млрд. \$	8	9,5	11	12	13	14	15	16,5	17	18
I, млрд. \$	1,65	1,8	2,0	2,1	2,2	2,4	2,65	2,85	3,2	3,55
ВВП, млрд. \$	14	16	18	20	23	23,5	25	26,5	28,5	30,5

a) Оцените коэффициенты линейной регрессионной модели

$$\text{ВНП} = \beta_0 + \beta_1 I + \beta_2 C + \varepsilon.$$

b) Оцените стандартную ошибку регрессии и стандартные ошибки коэффициентов.

c) Вычислите коэффициент детерминации R^2 и скорректированный коэффициент детерминации \bar{R}^2 ; сравните их значения. Оцените статистическую значимость R^2 .

d) Определите значение статистики Дарбина–Уотсона. Имеет ли место автокорреляция остатков?

e) Через три года предполагаются следующие уровни потребления и инвестиций: $C = 22$, $I = 3,8$. Какой уровень ВНП ожидается при этом?

Задание 3. При оценивании модели $Y_t = X_t' \beta + \varepsilon_t$ обнаружена автокорреляция и оцененная регрессия остатков показала, что $e_t = -0.6e_{t-1}$. Чтобы провести корректное оценивание, необходимо применить метод наименьших квадратов к преобразованным данным, причем для первого наблюдения будет использовано преобразование (выберите правильный ответ):

- 1) $Y_1^* = 0.8Y_1$; $X_1^* = 0.8X_1$; 2) $Y_1^* = 0.6Y_1$; $X_1^* = 0.6X_1$; 3) $Y_1^* = 0.4Y_1$; $X_1^* = 0.4X_1$; 4) $Y_1^* = -0.6Y_1$; $X_1^* = -0.6X_1$; 5) $Y_1^* = Y_1\sqrt{0.84}$; $X_1^* = X_1\sqrt{0.84}$

ТЕМА 7. СИСТЕМЫ ОДНОВРЕМЕННЫХ УРАВНЕНИЙ

Данная тема посвящена изучению подходов к моделированию экономических процессов не одним, а несколькими уравнениями (системой уравнений), т.е. набором взаимосвязанных регрессионных моделей, в которых

одни и те же переменные в различных уравнениях системы могут играть роль результирующих показателей и объясняющих переменных.

Цель проведения лабораторной работы: научиться построению и исследованию систем одновременных уравнений путём оценивания их параметров.

Для выполнения лабораторной работы № 7 студенты должны

знать:

- разновидности систем эконометрических уравнений;
- смысл эндогенных, экзогенных и предопределенных переменных;
- особенности структурной и приведенной форм системы одновременных уравнений;
- содержание, проблемы и методы проверки идентифицируемости структурных уравнений;
- алгоритм оценивания параметров структурной модели с использованием косвенного и двухшагового МНК;
- направления использования систем одновременных уравнений в практике эконометрических исследований и прогнозирования экономических показателей.

уметь:

- пользоваться таблицами для проверки статистических гипотез;
- пользоваться пакетом EXCEL:
 - строить таблицы; составлять и копировать формулы;
 - пользоваться статистическими и математическими функциями.

Задание для выполнения лабораторной работы №7

1. На основе исходных данных (табл.7.1) в соответствии со своим вариантом (1 – 20):

1) постройте эконометрическую модель в виде системы одновременных уравнений, включающей уравнения

$$\text{производительности труда } Y_1 = f(Y_2, X_1, X_2, X_4)$$

и заработной платы

$$Y_2 = f(Y_1, X_2, X_3, X_4)$$

в структурной и приведенной формах;

2) выполнив проверку идентифицируемости модели (по правилу порядка и ранга), выберите метод оценивания структурных параметров;

3) с применением выбранного метода найдите оценки параметров и проверьте статистическую надежность полученных результатов;

4) вычислите эластичность производительности труда и заработной платы в зависимости от влияющих факторов для средних значений этих переменных;

5) сделайте прогноз производительности труда и заработной платы на следующий период.

Таблица 7.1 – Исходные данные

Месяц	Производительность труда, ден. ед. Y_1	Средняя зарплата, ден. ед. Y_2	Фондовооруженность труда, ден. ед. X_1	Коэффициент текучести кадров, % X_2	Потери рабочего времени, % X_3	Стаж работы, лет X_4
Вариант 1						
t	Y_1	Y_2	X_1	X_2	X_3	X_4
1	54	222	74	15	4,7	7
2	55	230	76	14,5	4,8	7,5
3	52	212	74	14	5	7
4	53	222	75	13	5,2	8
5	56	247	72	12,1	5,2	9
6	57	252	69	11	5,3	10
7	59	262	69	10,5	5,4	12
8	54	234	64	10,2	5,6	12
9	62	272	74	10	5,7	12,5
10	62	267	74	7,5	5,7	13
11	64	277	76	7	5,4	15
12	66	282	77	6,7	6	12
13	67	282	78	6,6	6,2	14
14	69	292	82	6	6,3	15
15	69	287	84	6,1	6,7	16
16	64	275	86	6,2	6,8	16,5
17	65	280	86	6,5	6,8	17,5
18	68	285	89	6	6,9	19
19	70	281	92	6	7	18,5
20	72	287	94	5	6,7	19,5
21	71	282	94	6	7,2	19,6
22	74	293	95	7	7,3	19,7
23	76	292	95	7	7,4	19,8
24	78	297	96	8	7,4	19,9
Вариант 2						
t	Y_1	Y_2	X_1	X_2	X_3	X_4
1	62	32	15	17	10	252
2	63	37	14,5	16,3	10,5	234
3	60	35	14	14	9	247

4	61	36	13	14,8	11	251
5	64	38	12	15	12	262
6	65	40	11	14,5	13	262
7	67	42	10,5	13	11,5	267
8	62	43	10,2	13,5	14	260
9	70	47	10	12	12	265
10	71	47	7,5	11	16	267
11	72	48	7	10	14,5	272
12	74	50	6,7	9,5	14	275
13	75	49	6,6	8,5	12	274
14	80	52	6	8	13	282
15	77	51	6,1	8,2	14	267
16	82	53	6,2	7,8	17	287
17	83	52	6,5	7,5	17,5	288
18	85	7,3	6	7	16,5	288
19	83	57	6	6,5	16	285
20	87	58	5	6,7	17	287
21	89	7,8	6	7	17	292
22	90	60	7	7,1	18,5	292
23	92	61	7	6,8	17,7	295
24	94	62	8	7,2	17,8	297

Вариант 3

t	Y ₁	Y ₂	X ₁	X ₂	X ₃	X ₄
1	67	254	32	15	17	10
2	68	256	37	14,5	16,3	10,5
3	65	249	35	14	14	9
4	66	253	36	13	14,8	11
5	69	264	38	12	15	12
6	70	264	40	11	14,5	13
7	72	269	42	10,5	13	11,5
8	67	262	43	10,2	13,5	14
9	75	267	47	10	12	12
10	76	269	47	7,5	11	16
11	77	274	48	7	10	14,5
12	79	277	50	6,7	9,5	14
13	80	277	49	6,6	8,5	12
14	85	284	52	6	8	13
15	82	269	51	6,1	8,2	14
16	87	289	53	6,2	7,8	17
17	88	290	52	6,5	7,5	17,5
18	90	290	55	6	7	16,5
19	88	287	57	6	6,5	16
20	92	289	58	5	6,7	17
21	94	294	60	6	7	17
22	92	294	60	7	7,1	18,5
23	96	297	61	7	6,8	17,7
24	97	299	62	8	7,2	17,8

Вариант 4

t	Y ₁	Y ₂	X ₁	X ₂	X ₃	X ₄
1	65	242	33	15	17	10
2	66	244	38	14,5	16,3	10,5
3	63	237	36	14	14	9

4	64	241	37	13	14,8	11
5	65	252	39	12	15	12
6	68	252	41	11	14,5	13
7	70	257	43	10,5	13	11,5
8	65	250	44	10,2	13,5	14
9	73	255	48	10	12	12
10	74	257	48	7,5	11	16
11	75	262	49	7	10	14,5
12	77	265	51	6,7	9,5	14
13	78	264	50	6,6	8,3	12
14	83	272	53	6	8	13
15	80	257	52	6,1	8,2	14
16	85	277	54	6,2	7,8	17
17	86	278	53	6,5	7,5	17,5
18	88	278	56	6	7	16,5
19	86	275	58	6	6,5	16
20	90	277	59	5	6,7	17
21	90	282	58	6	7	17
22	89	282	56	7	7,1	18,5
23	90	285	58	7	6,8	17,7
24	89	285	58	6	5,2	15,8

Вариант 5

t	Y ₁	Y ₂	X ₁	X ₂	X ₃	X ₄
1	59	232	34	16	17	10
2	60	234	39	15,8	16,3	10,5
3	57	227	37	15	14	9
4	58	231	38	14	14,8	11
5	61	242	40	13	15	12
6	62	242	42	12	14,5	13
7	64	247	44	11,5	13	11,5
8	59	240	45	11,2	13,5	14
9	67	245	49	11	12	12
10	68	245	49	8,5	11	16
11	69	252	50	8	10	14,5
12	71	255	52	7,7	9,5	14
13	72	254	51	7,6	8,5	12
14	77	262	54	7	8	13
15	74	247	53	7,1	8,2	14
16	79	267	55	7,2	7,8	17
17	80	268	54	4,5	7,5	17,5
18	82	268	57	7	7	16,5
19	80	265	59	7	6,5	16
20	84	267	60	6	6,7	17
21	85	272	62	7	7	17
22	86	272	62	8	7,1	18,5
23	87	277	63	8	6,8	17,7
24	88	277	65	9	7,2	17,8

Вариант 6

t	Y ₁	Y ₂	X ₁	X ₂	X ₃	X ₄
1	57	222	35	15	16	10
2	58	224	40	14,5	15,3	10,5
3	55	220	38	14	13	9

4	56	227	39	13	13,8	11
5	59	232	41	12	14	12
6	60	237	43	11	13,5	13
7	62	242	45	10,5	14	11,5
8	57	244	46	10,2	12,5	14
9	65	252	50	9	11	12
10	66	254	50	7,5	10	16
11	67	256	51	7	9	14,5
12	69	257	53	6,7	8,5	14
13	70	262	52	6,6	7,5	12
14	75	267	55	6	7	13
15	72	252	54	6,1	7,2	14
16	77	267	56	6,2	6,8	17
17	78	269	55	6,5	6,5	17,5
18	80	272	58	6	6	16,5
19	78	267	60	6	5,5	16
20	82	277	61	5	5,7	17
21	84	280	62	6	6	17
22	84	280	63	7	6,1	18,5
23	86	282	64	7	5,8	17,7
24	87	284	63	8	6,2	17,8

Вариант 7

t	Y ₁	Y ₂	X ₁	X ₂	X ₃	X ₄
1	62	252	34	14	17	10
2	63	254	39	13,5	16,3	10,5
3	60	247	37	13	14	9
4	61	248	38	12	14,8	11
5	64	254	40	11	15	12
6	65	255	42	10	14,5	13
7	67	257	44	9,5	13	11,5
8	62	252	45	9,3	13,5	14
9	70	258	49	9	12	12
10	71	259	49	6,5	11	16
11	72	262	50	6	10	15,5
12	9,2	264	52	5,7	9,5	14
13	75	267	51	5,6	8,5	12
14	80	272	54	5	8	13
15	77	270	53	5,1	8,2	14
16	82	272	55	5,2	7,8	17
17	83	277	54	5,5	7,5	17,5
18	85	280	57	5	7	16,5
19	83	277	59	5	6,5	16
20	87	280	60	4	6,7	17
21	89	278	62	5	7	17
22	90	282	62	6	5,3	18,5
23	92	282	63	6	6,8	17,7
24	94	287	64	7	7,2	17,8

Вариант 8

t	Y ₁	Y ₂	X ₁	X ₂	X ₃	X ₄
1	54	232	27	15	17	10
2	55	232	32	14,5	16,3	10,5
3	52	227	30	14	14	9

4	53	230	31	13	14,8	11
5	56	237	33	12	15	12
6	57	237	35	11	14,5	13
7	59	242	37	10,5	13	11,5
8	54	232	38	10,2	13,5	14
9	62	242	42	10	12	12
10	63	242	42	7,5	11	16
11	64	244	43	7	10	14,5
12	66	247	45	6,7	9,5	14
13	67	247	44	6,6	8,5	12
14	72	252	47	6	8	13
15	69	245	46	6,1	8,2	14
16	74	257	48	6,2	7,8	17
17	75	257	47	6,5	7,5	17,5
18	77	262	50	6	7	16,5
19	75	256	52	6	6,5	16
20	79	272	53	5	6,7	17
21	82	277	57	6	7	17
22	80	272	54	7	7,1	18,5
23	82	277	56	7	6,8	17,7
24	84	282	58	8	7,2	17,8

Вариант 9

t	Y ₁	Y ₂	X ₁	X ₂	X ₃	X ₄
1	63	242	32	15	18	9
2	64	242	37	14,5	17,3	9,5
3	61	237	35	14	15	8
4	62	242	36	13	15,8	10
5	65	252	38	12	16	11
6	66	257	40	11	15,5	12
7	68	258	42	10,5	14	10,5
8	63	242	43	10,2	14,5	13
9	71	262	47	10	13	11
10	72	262	47	7,5	12	15
11	73	267	48	7	11	13,5
12	75	272	50	6,7	10,5	13
13	76	276	49	6,6	9,5	11
14	81	278	52	6	9	12
15	78	277	51	6,1	9,2	13
16	83	282	53	6,2	8,8	16
17	84	272	52	6,5	8,5	16,5
18	86	277	55	6	8	15,5
19	84	272	57	6	7,5	15
20	88	277	58	5	7,7	16
21	90	277	58	6	8	16
22	91	282	57	7	8,1	17,5
23	92	287	58	7	7,8	16,7
24	97	292	59	8	8,2	16,8

Вариант 10

t	Y ₁	Y ₂	X ₁	X ₂	X ₃	X ₄
1	66	227	32	14	17	10
2	67	232	37	13,5	16,3	10,5
3	64	212	35	13	14	9

4	65	232	36	12	14,8	11
5	68	252	38	11	15	12
6	69	252	40	10	14,5	13
7	71	257	42	9,5	13	11,5
8	66	237	43	9,2	13,5	14
9	74	242	47	9	12	12
10	75	244	47	6,5	11	16
11	76	247	48	6	10	14,5
12	78	248	50	5,7	9,5	14
13	79	252	49	5,6	8,5	12
14	84	257	52	5	8	13
15	81	253	51	5,1	8,2	14
16	86	262	53	5,2	7,8	17
17	87	267	52	5,5	7,5	17,5
18	89	268	55	5	7	16,5
19	87	262	57	5	6,5	16
20	91	267	58	4	6,7	17
21	92	272	59	5	7	17
22	93	272	60	6	7,1	18,5
23	92	273	60	6	6,8	17,7
24	94	277	62	7	7,2	17,8

Вариант 11

t	Y ₁	Y ₂	X ₁	X ₂	X ₃	X ₄
1	50	218	70	11	0,7	3
2	51	226	72	10,5	0,8	3,5
3	48	208	70	10	1	3
4	49	218	71	9	1,2	4
5	52	243	68	8,1	1,2	5
6	53	248	65	7	1,3	6
7	55	258	65	6,5	1,4	8
8	50	230	60	6,2	1,6	8
9	58	268	70	6	1,7	8,5
10	58	263	70	3,5	1,7	9
11	60	273	72	3	1,4	11
12	62	278	73	2,7	2	8
13	63	278	74	2,6	2,2	10
14	65	288	78	2	2,3	11
15	65	283	80	2,1	2,7	12
16	60	271	82	2,2	2,8	12,5
17	61	276	82	2,5	2,8	13,5
18	64	281	85	2	2,9	15
19	66	277	88	2	3	14,5
20	68	283	90	1	2,7	15,5
21	67	278	90	2	3,2	15,6
22	70	289	91	3	3,3	15,7
23	72	288	91	3	3,4	15,8
24	74	293	92	4	3,4	15,9

Вариант 12

t	Y ₁	Y ₂	X ₁	X ₂	X ₃	X ₄
1	58	28	11	13	6	248
2	59	33	10,5	12,3	6,5	230
3	56	31	10	10	5	243

4	57	32	9	10,8	7	247
5	60	34	8	11	8	258
6	61	36	7	10,5	9	258
7	63	38	6,5	9	7,5	263
8	58	39	6,2	9,5	10	256
9	66	43	6	8	8	261
10	67	43	3,5	7	12	263
11	68	44	3	6	10,5	268
12	70	46	2,7	5,5	10	271
13	71	45	2,6	4,5	8	270
14	76	48	2	4	9	278
15	73	47	2,1	4,2	10	263
16	78	49	2,2	3,8	13	283
17	79	48	2,5	3,5	13,5	284
18	81	3,3	2	3	12,5	284
19	79	53	2	2,5	12	281
20	83	54	1	2,7	13	283
21	85	3,8	2	3	13	288
22	86	56	3	3,1	14,5	288
23	88	57	3	2,8	13,7	291
24	90	58	4	3,2	13,8	293

Вариант 13

t	Y ₁	Y ₂	X ₁	X ₂	X ₃	X ₄
1	63	250	28	11	13	6
2	64	252	33	10,5	12,3	6,5
3	61	245	31	10	10	5
4	62	249	32	9	10,8	7
5	65	260	34	8	11	8
6	66	260	36	7	10,5	9
7	68	265	38	6,5	9	7,5
8	63	258	39	6,2	9,5	10
9	71	263	43	6	8	8
10	72	265	43	3,5	7	12
11	73	270	44	3	6	10,5
12	75	273	46	2,7	5,5	10
13	76	273	45	2,6	4,5	8
14	81	280	48	2	4	9
15	78	265	47	2,1	4,2	10
16	83	285	49	2,2	3,8	13
17	84	286	48	2,5	3,5	13,5
18	86	286	51	2	3	12,5
19	84	283	53	2	2,5	12
20	88	285	54	1	2,7	13
21	90	290	56	2	3	13
22	88	290	56	3	3,1	14,5
23	92	293	57	3	2,8	13,7
24	93	295	58	4	3,2	13,8

Вариант 14

t	Y ₁	Y ₂	X ₁	X ₂	X ₃	X ₄
1	61	238	29	11	13	6
2	62	240	34	10,5	12,3	6,5
3	59	233	32	10	10	5

4	60	237	33	9	10,8	7
5	61	248	35	8	11	8
6	64	248	37	7	10,5	9
7	66	253	39	6,5	9	7,5
8	61	246	40	6,2	9,5	10
9	69	251	44	6	8	8
10	70	253	44	3,5	7	12
11	71	258	45	3	6	10,5
12	73	261	47	2,7	5,5	10
13	74	260	46	2,6	4,3	8
14	79	268	49	2	4	9
15	76	253	48	2,1	4,2	10
16	81	273	50	2,2	3,8	13
17	82	274	49	2,5	3,5	13,5
18	84	274	52	2	3	12,5
19	82	271	54	2	2,5	12
20	86	273	55	1	2,7	13
21	86	278	54	2	3	13
22	85	278	52	3	3,1	14,5
23	86	281	54	3	2,8	13,7
24	87	283	56	4	3,2	13,8

Вариант 15

t	Y ₁	Y ₂	X ₁	X ₂	X ₃	X ₄
1	55	228	30	12	13	6
2	56	230	35	11,8	12,3	6,5
3	53	223	33	11	10	5
4	54	227	34	10	10,8	7
5	57	238	36	9	11	8
6	58	238	38	8	10,5	9
7	60	243	40	7,5	9	7,5
8	55	236	41	7,2	9,5	10
9	63	241	45	7	8	8
10	64	241	45	4,5	7	12
11	65	248	46	4	6	10,5
12	67	251	48	3,7	5,5	10
13	68	250	47	3,6	4,5	8
14	73	258	50	3	4	9
15	70	243	49	3,1	4,2	10
16	75	263	51	3,2	3,8	13
17	76	264	50	0,5	3,5	13,5
18	78	264	53	3	3	12,5
19	76	261	55	3	2,5	12
20	80	263	56	2	2,7	13
21	81	268	58	3	3	13
22	82	268	58	4	3,1	14,5
23	83	273	59	4	2,8	13,7
24	84	273	61	5	3,2	13,8

Вариант 16

t	Y ₁	Y ₂	X ₁	X ₂	X ₃	X ₄
1	53	218	31	11	12	6
2	54	220	36	10,5	11,3	6,5
3	51	216	34	10	9	5

4	52	223	35	9	9,8	7
5	55	228	37	8	10	8
6	56	233	39	7	9,5	9
7	58	238	41	6,5	10	7,5
8	53	240	42	6,2	8,5	10
9	61	248	46	5	7	8
10	62	250	46	3,5	6	12
11	63	252	47	3	5	10,5
12	65	253	49	2,7	4,5	10
13	66	258	48	2,6	3,5	8
14	71	263	51	2	3	9
15	68	248	50	2,1	3,2	10
16	73	263	52	2,2	2,8	13
17	74	265	51	2,5	2,5	13,5
18	76	268	54	2	2	12,5
19	74	263	56	2	1,5	12
20	78	273	57	1	1,7	13
21	80	276	58	2	2	13
22	80	276	59	3	2,1	14,5
23	82	278	60	3	1,8	13,7
24	83	280	59	4	2,2	13,8

Вариант 17

t	Y ₁	Y ₂	X ₁	X ₂	X ₃	X ₄
1	58	248	30	10	13	6
2	59	250	35	9,5	12,3	6,5
3	56	243	33	9	10	5
4	57	244	34	8	10,8	7
5	60	250	36	7	11	8
6	61	251	38	6	10,5	9
7	63	253	40	5,5	9	7,5
8	58	248	41	5,3	9,5	10
9	66	254	45	5	8	8
10	67	255	45	2,5	7	12
11	68	258	46	2	6	11,5
12	5,2	260	48	1,7	5,5	10
13	71	263	47	1,6	4,5	8
14	76	268	50	1	4	9
15	73	266	49	1,1	4,2	10
16	78	268	51	1,2	3,8	13
17	79	273	50	1,5	3,5	13,5
18	81	276	53	1	3	12,5
19	79	273	55	1	2,5	12
20	83	276	56	0	2,7	13
21	85	274	58	1	3	13
22	86	278	58	2	4,9	14,5
23	88	278	59	2	2,8	13,7
24	90	283	60	3	3,2	13,8

Вариант 18

t	Y ₁	Y ₂	X ₁	X ₂	X ₃	X ₄
1	50	228	23	11	13	6
2	51	228	28	10,5	12,3	6,5
3	48	223	26	10	10	5

4	49	226	27	9	10,8	7
5	52	233	29	8	11	8
6	53	233	31	7	10,5	9
7	55	238	33	6,5	9	7,5
8	50	228	34	6,2	9,5	10
9	58	238	38	6	8	8
10	59	238	38	3,5	7	12
11	60	240	39	3	6	10,5
12	62	243	41	2,7	5,5	10
13	63	243	40	2,6	4,5	8
14	68	248	43	2	4	9
15	65	241	42	2,1	4,2	10
16	70	253	44	2,2	3,8	13
17	71	253	43	2,5	3,5	13,5
18	73	258	46	2	3	12,5
19	71	252	48	2	2,5	12
20	75	268	49	1	2,7	13
21	78	273	53	2	3	13
22	76	268	50	3	3,1	14,5
23	78	273	52	3	2,8	13,7
24	80	278	54	4	3,2	13,8

Вариант 19

t	Y ₁	Y ₂	X ₁	X ₂	X ₃	X ₄
1	59	238	28	11	14	5
2	60	238	33	10,5	13,3	5,5
3	57	233	31	10	11	4
4	58	238	32	9	11,8	6
5	61	248	34	8	12	7
6	62	253	36	7	11,5	8
7	64	254	38	6,5	10	6,5
8	59	238	39	6,2	10,5	9
9	67	258	43	6	9	7
10	68	258	43	3,5	8	11
11	69	263	44	3	7	9,5
12	71	268	46	2,7	6,5	9
13	72	272	45	2,6	5,5	7
14	77	274	48	2	5	8
15	74	273	47	2,1	5,2	9
16	79	278	49	2,2	4,8	12
17	80	268	48	2,5	4,5	12,5
18	82	273	51	2	4	11,5
19	80	268	53	2	3,5	11
20	84	273	54	1	3,7	12
21	86	273	54	2	4	12
22	87	278	53	3	4,1	13,5
23	88	283	54	3	3,8	12,7
24	93	288	55	4	4,2	12,8

Вариант 20

t	Y ₁	Y ₂	X ₁	X ₂	X ₃	X ₄
1	62	223	28	10	13	6
2	63	228	33	9,5	12,3	6,5
3	60	208	31	9	10	5

4	61	228	32	8	10,8	7
5	64	248	34	7	11	8
6	65	248	36	6	10,5	9
7	67	253	38	5,5	9	7,5
8	62	233	39	5,2	9,5	10
9	70	238	43	5	8	8
10	71	240	43	2,5	7	12
11	72	243	44	2	6	10,5
12	74	244	46	1,7	5,5	10
13	75	248	45	1,6	4,5	8
14	80	253	48	1	4	9
15	77	249	47	1,1	4,2	10
16	82	258	49	1,2	3,8	13
17	83	263	48	1,5	3,5	13,5
18	85	264	51	1	3	12,5
19	83	258	53	1	2,5	12
20	87	263	54	0	2,7	13
21	88	268	55	1	3	13
22	89	268	56	2	3,1	14,5
23	88	269	56	2	2,8	13,7
24	90	273	58	3	3,2	13,8

Для выполнения работы следует изучить теоретический материал, представленный в теме 9 Учебного пособия по эконометрике [18].

РЕШЕНИЕ ТИПОВЫХ ЗАДАЧ

Пример 7.1. Проверка идентифицируемости СОУ и выбор метода оценивания структурных параметров

Одна из версий макроэкономической модели имеет вид:

$$C_t = a_1 + b_{11}Y_t + b_{12}C_{t-1} + \varepsilon_{1t}, \text{ (функция потребления);}$$

$$I_t = a_2 + b_{21}r_t + b_{22}I_{t-1} + \varepsilon_{2t}, \text{ (функция инвестиций); (7.1)}$$

$$r_t = a_3 + b_{31}Y_t + b_{32}M_t + \varepsilon_{3t}, \text{ (функция денежного рынка);}$$

$$Y_t = C_t + I_t + G_t, \text{ (тождество дохода),}$$

Здесь C – расходы на потребление; I – инвестиции; r – процентная ставка; M – денежная масса; G – государственные расходы; Y – ВВП; t – текущий период; $t-1$ – предыдущий период.

Задание:

- 1) классифицируйте переменные на эндогенные, экзогенные и предопределенные;
- 2) применив необходимое и достаточное условия идентифицируемости, определите, идентифицировано ли каждое из уравнений модели;
- 3) предложите метод для оценивания параметров модели.

Решение.

1). Исходная модель представляет систему взаимосвязанных (одновременных) уравнений в структурной форме. Она содержит четыре уравнения (из которых первые три являются уравнениями поведения, и одно тождество), четыре эндогенные переменные (C_t, I_t, r_t, Y_t) и четыре предопределенные переменные ($C_{t-1}, I_{t-1}, M_t, G_t$), из которых – две эндогенные лаговые (C_{t-1}, I_{t-1}) и две - экзогенные (M_t, G_t).

Переобозначим переменные модели: Y – эндогенные текущие, X – предопределенные.

$Y_1 = C_t$ – расходы на потребление в текущем периоде;

$Y_2 = I_t$ – инвестиции в текущем периоде;

$Y_3 = r_t$ – процентная ставка денежного рынка;

$Y_4 = Y_t$ – ВВП;

$X_0 \equiv 1$ фиктивная переменная, приписываемая свободному члену;

$X_1 = C_{t-1}$ – расходы на потребление в прошлом периоде;

$X_2 = I_{t-1}$ – инвестиции в прошлом периоде;

$X_3 = M_t$ – денежная масса (денежный агрегат M_2);

$X_4 = G_t$ – государственные расходы.

С учетом введенных обозначений перепишем систему в следующем виде:

$$\begin{aligned}
Y_1 &= \alpha_{14}Y_4 + \beta_{10}X_0 + \beta_{11}X_1 + \varepsilon_1, \text{(функция потребления);} \\
Y_2 &= \alpha_{23}Y_3 + \beta_{20}X_0 + \beta_{22}X_2 + \varepsilon_2, \quad \text{(функция инвестиций);} \quad (7.2) \\
Y_3 &= \alpha_{34}Y_4 + \beta_{30}X_0 + \beta_{33}X_3 + \varepsilon_3, \text{(функция денежного рынка);} \\
Y_4 &= Y_1 + Y_2 + X_4. \text{(тождество дохода).}
\end{aligned}$$

2). Проверим, идентифицировано ли каждое из уравнений модели, применив необходимое и достаточное условия идентифицируемости.

Для удобства рассуждений запишем (7.2) модель в следующем виде (7.3):

$$\begin{aligned}
Y_1 - \alpha_{14}Y_4 - \beta_{10}X_0 - \beta_{11}X_1 &= \varepsilon_1, \\
Y_2 - \alpha_{23}Y_3 - \beta_{20}X_0 - \beta_{22}X_2 &= \varepsilon_2, \\
Y_3 - \alpha_{34}Y_4 - \beta_{30}X_0 - \beta_{33}X_3 &= \varepsilon_3, \\
Y_4 - Y_1 - Y_2 - X_4 &= 0.
\end{aligned} \quad (7.3)$$

Составим таблицу из коэффициентов при переменных модели (7.3) (табл. 7.2).

Таблица 7.2 – Коэффициенты при переменных модели

№ уравнения	$X_0=1$	Y_1	Y_2	Y_3	Y_4	X_1	X_2	X_3	X_4
1	$-\beta_{10}$	1	0	0	$-\alpha_{14}$	$-\beta_{11}$	0	0	0
2	$-\beta_{20}$	0	1	$-\alpha_{23}$	0	0	$-\beta_{22}$	0	0
3	$-\beta_{30}$	0	0	1	$-\alpha_{34}$	0	0	$-\beta_{33}$	0
4	0	-1	-1	0	1	0	0	0	-1

Вначале применим необходимое условие идентифицируемости (условия порядка, или счетное правило)

Используя данные табл. 7.2, проверим условия порядка для каждого отдельного уравнения. Результат проверки запишем в табл. табл. 7.3.

Таблица 7.3 – Проверка идентифицируемости (условия порядка)

№ уравнения i	Количество отсутствующих предопределенных переменных $(k - k_i)$	Количество включенных эндогенных переменных минус один $(m_i - 1)$	Идентифици- руемость
1	$5 - 2 = 3$	$2 - 1 = 1$	Сверх
2	$5 - 2 = 3$	$2 - 1 = 1$	Сверх
3	$5 - 2 = 3$	$2 - 1 = 1$	Сверх
4	-	-	-

В соответствии со счетным правилом все уравнения сверхидентифицируемы (необходимые условия выполнены, хотя как неравенства: $k - k_i > m_i - 1$), значит модель сверхидентифицируема.

Проведем проверку, используя условие ранга (необходимое и достаточное), опираясь на данные табл. 7.2.

В первом уравнении отсутствуют Y_2, Y_3, X_2, X_3, X_4 (нули в первой строке табл. 7.2). Для проверки этого уравнения на идентифицируемость построим соответствующую матрицу коэффициентов при переменных Y_2, Y_3, X_2, X_3, X_4 , включенных в остальные уравнения системы, кроме первого:

$$A_1 = \begin{bmatrix} 1 & -\alpha_{23} & -\beta_{22} & 0 & 0 \\ 0 & 1 & 0 & -\beta_{33} & 0 \\ -1 & 0 & 0 & 0 & -1 \end{bmatrix}$$

Определитель матрицы, составленной, например, из первых трех столбцов матрицы A_1 , не равен нулю ($\Delta = -\beta_{22}$), следовательно, ранг матрицы A_1 равняется трем (равен количеству эндогенных переменных модели без единицы). Условие ранга выполнено и, значит, первое уравнение идентифицируемо.

Второе уравнение модели (7.3) не содержит Y_1, Y_4, X_1, X_3, X_4 . Составим матрицу коэффициентов при этих переменных из других уравнений (первого, третьего и четвертого) (см. табл. 7.2):

$$A_2 = \begin{bmatrix} 1 & -\alpha_{23} & -\beta_{22} & 0 & 0 \\ 0 & 1 & 0 & -\beta_{33} & 0 \\ -1 & 0 & 0 & 0 & -1 \end{bmatrix}$$

Определитель матрицы, составленной из первого, второго и пятого столбцов матрицы A_2 , не равен нулю ($\Delta = -1$), следовательно, ранг матрицы A_2 равняется трем. Условие ранга выполнено и, значит, второе уравнение идентифицируемо.

В третьем уравнении нет переменных Y_1, Y_2, X_1, X_2, X_4 . Соответствующая матрица:

$$A_3 = \begin{bmatrix} 1 & 0 & -\beta_{11} & 0 & 0 \\ 0 & 1 & 0 & -\beta_{22} & 0 \\ -1 & -1 & 0 & 0 & -1 \end{bmatrix}$$

Определитель матрицы, составленной из первого, второго и пятого столбцов матрицы A_3 , не равен нулю ($\Delta = -1$), следовательно, ранг матрицы A_3 равняется трем. Условие ранга выполнено и, значит, третье уравнение идентифицируемо.

Казалось бы достаточно иметь только правило ранга. Однако правило порядка обладает дополнительным свойством: *выполнение правила ранга обеспечивает идентифицируемость уравнения модели, а с помощью правила порядка относят это уравнение к тому или другому виду.*

Таким образом, условие ранга выполнено для всех уравнений, что подтверждает сделанный ранее вывод о сверхидентифицируемости модели.

3). Для оценивания структурных параметров точно идентифицируемых уравнений модели можно применить косвенный метод наименьших квадратов (КМНК). Параметры сверхидентифицируемых уравнений оцениваются с помощью двухшагового метода наименьших квадратов (2МНК).

Причем доказано, что если для уравнения выполнено ранговое условие идентифицируемости и порядковое условие выполняется как равенство (точная идентифицируемость), то 2МНК-оценка совпадает с оценкой, полученной косвенным методом наименьших квадратов. Это положение будет проиллюстрировано в примерах 7.3 и 7.4.

Пример 7.2. Оценивание параметров структурной формы модели косвенным методом наименьших квадратов (КМНК).

Рассмотрим следующую линейную взаимозависимую модель в структурной форме, описывающую денежное обращение:

$$Y_{1t} = -\alpha_{12}Y_{2t} + \beta_{11}X_{1t} + \beta_{12}X_{2t} + \varepsilon_{1t}, (7.4)$$
$$Y_{2t} = -\alpha_{21}Y_{1t} + \beta_{21}X_{1t} + \beta_{23}X_{3t} + \varepsilon_{2t},$$

где Y_{1t} - денежное обращение;

Y_{2t} - оборачиваемость денег;

X_{1t} - фиктивная переменная при свободном члене уравнения, тождественно равная 1;

X_{2t} - денежные доходы населения;

X_{3t} - размер вклада в сберегательную кассу,

ε_{it} - остатки уравнений, являющиеся реализациями случайных составляющих модели, или возмущений.

Отрицательные коэффициенты в правых частях уравнений при переменных Y_{1t} и Y_{2t} указывают на их обратную зависимость: денежное обращение (количество денег в обороте) должно быть тем больше, чем меньше величина оборачиваемости денег.

Исходные данные для расчетов приведены в табл.7.4.

Таблица 7.4 – Исходные данные

t	Y_1	Y_2	X_1	X_2	X_3
1	25	10	1	1	22
2	28	11	1	4	19
3	29	9	1	3	13
4	31	7	1	5	16
5	35	8	1	6	13
6	38	6	1	6	9
7	40	4	1	8	6
8	39	2	1	8	8
9	42	1	1	9	3
10	43	2	1	10	1

Необходимо найти оценки параметров модели и сделать выводы.

Решение.

Выполним проверку идентифицируемости модели (7.4), используя счетное правило.

В системе (7.4) три экзогенных переменных (X_1, X_2 и X_3) и две эндогенных (Y_1, Y_2); оба уравнения включают по две экзогенные переменные и по 2 эндогенные, так что

$$k = 3, \quad k_1 = 2, \quad k_2 = 2;$$

$$m = 2, \quad m_1 = 2, \quad m_2 = 2.$$

Проверим выполнение условия для каждого из уравнений:

$$k - k_i \geq m_i - 1$$

Для первого уравнения имеем:

$$3 - 2 \geq 2 - 1, \quad \text{получили } 1 = 1;$$

для второго уравнения:

$$3 - 2 \geq 2 - 1, \quad \text{получили } 1 = 1;$$

Таким образом, для обоих уравнений необходимое условие выполнено как равенство, следовательно они (значит, и модель в целом) точно идентифицируемы, и можно выполнить оценку параметров модели с помощью *косвенного метода наименьших квадратов*. (Здесь не приводим проверку с помощью рангового условия, поскольку его выполнение очевидно).

Косвенный метод наименьших квадратов будем применять в такой последовательности:

- 1) преобразование структурной формы модели в приведенную;
- 2) оценивание параметров каждого уравнения приведенной формы с помощью обычного МНК;
- 3) определение оценок структурных параметров по найденным на втором этапе коэффициентам приведенных уравнений.

Используя матричную форму записи, представим систему (7.4) в виде:

$$AY = BX + e, \quad (7.5)$$

где

$$A = \begin{pmatrix} 1 & a_{12} \\ a_{21} & 1 \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} & 0 \\ b_{21} & 0 & b_{23} \end{pmatrix}, \quad Y = \begin{pmatrix} Y_{1t} \\ Y_{2t} \end{pmatrix} = \begin{pmatrix} Y_{11} & Y_{12} \dots & Y_{1n} \\ Y_{21} & Y_{22} \dots & Y_{2n} \end{pmatrix},$$

$$X = \begin{pmatrix} X_{1t} \\ X_{2t} \\ X_{3t} \end{pmatrix} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ X_{21} & X_{22} & \dots & X_{2n} \\ X_{31} & X_{32} & \dots & X_{3n} \end{pmatrix}, \quad e = \begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix} = \begin{pmatrix} e_{11} & e_{12} & \dots & e_{1n} \\ e_{21} & e_{22} & \dots & e_{2n} \end{pmatrix}.$$

Умножив (7.5) слева на матрицу A^{-1} (на обратную матрицу), получим:

$$A^{-1}AY = A^{-1}BX + A^{-1}e,$$

или
$$Y = CX + U, \quad (7.6)$$

где $U = A^{-1}e, \quad C = A^{-1}B. \quad (7.7)$

Выражение $C = A^{-1}B$ определяет соотношение между структурными и приведенными параметрами. При выполнении расчетов это выражение удобно представлять в виде:

$$AC = B. \quad (7.7')$$

Форма эконометрической модели, задаваемой в виде (7.6), называется **приведенной**.

Представим матричное уравнение (7.6) в развернутом виде:

$$Y_{1t} = c_{11}X_{1t} + c_{12}X_{2t} + c_{13}X_{3t} + u_{1t}, \quad (7.8)$$

$$Y_{2t} = c_{21}X_{1t} + c_{22}X_{2t} + c_{23}X_{3t} + u_{2t},$$

Каждое уравнение приведенной формы (7.8) оценим **методом наименьших квадратов**, используя функцию ЛИНЕЙНЕХСЕЛ (см. рис. 7.1 и 7.2)

Итак, имеем оценку уравнений в приведенной форме:

$$\hat{Y}_{1t} = 32,401X_{1t} + 1,218X_{2t} - 0,428X_{3t},$$

$$\hat{Y}_{2t} = 6,625X_{1t} - 0,575X_{2t} + 0,257X_{3t}.$$

ЛИНЕЙН =ЛИНЕЙН(B2:B11;E2:F11;1;1)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	t	Y1	Y2	X1	X2	X3									
2	1	25	10	1	1	22									
3	2	28	11	1	4	19									
4	3	29	9	1	3	13									
5	4	31	7	1	5	16									
6	5	35	8	1	6	13									
7	6	38	6	1	6	9									
8	7	40	4	1	8	6									
9	8	39	2	1	8	8									
10	9	42	1	1	9	3									
11	10	43	2	1	10	1									
12															
13															
14															
15															
16															
17															
18															
19															
20															
21															

Аргументы функции

ЛИНЕЙН

Известные_значения_y: B2:B11 = {25;28;29;31;35;38}

Известные_значения_x: E2:F11 = {1;22;4;19;3;13;5;1}

Конст: 1 = ИСТИНА

Статистика: 1 = ИСТИНА

= {-0,427927927927912;1}

Возвращает параметры линейного приближения по методу наименьших квадратов.

Статистика логическое значение, которое указывает, требуется ли вернуть дополнительную статистику по регрессии (ИСТИНА) или только коэффициенты m и константу b (ЛОЖЬ или отсутствие значения).

[Справка по этой функции](#) Значение: -0,427927928

Получили:

-0,42793	1,217718	32,4009
0,171971	0,415349	4,328579
0,968435	1,281154	#Н/Д
107,384	7	#Н/Д
352,5105	11,48949	#Н/Д

Рис.7.1. Оценка приведенных параметров первого уравнения с помощью МНК

ЛИНЕЙН =ЛИНЕЙН(C2:C11;E2:F11;1;1)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	t	Y1	Y2	X1	X2	X3									
2	1	25	10	1	1	22									
3	2	28	11	1	4	19									
4	3	29	9	1	3	13									
5	4	31	7	1	5	16									
6	5	35	8	1	6	13									
7	6	38	6	1	6	9									
8	7	40	4	1	8	6									
9	8	39	2	1	8	8									
10	9	42	1	1	9	3									
11	10	43	2	1	10	1									
12															
13															
14															
15															
16															
17															
18															
19															
20															
21															
22															
23															
24															
25															

Аргументы функции

ЛИНЕЙН

Известные_значения_y: C2:C11 = {10;11;9;7;8;6;4;2}

Известные_значения_x: E2:F11 = {1;22;4;19;3;13;5;1}

Конст: 1 = ИСТИНА

Статистика: 1 = ИСТИНА

= {0,256756756756759;-0}

Возвращает параметры линейного приближения по методу наименьших квадратов.

Статистика логическое значение, которое указывает, требуется ли вернуть дополнительную статистику по регрессии (ИСТИНА) или только коэффициенты m и константу b (ЛОЖЬ или отсутствие значения).

[Справка по этой функции](#) Значение: 0,256756757

Получили

0,256757	-0,57508	6,626126
0,207086	0,50016	5,21244
0,856374	1,542755	#Н/Д
20,86878	7	#Н/Д
99,33934	16,66066	#Н/Д

Рис.7.2. Оценка приведенных параметров второго уравнения с помощью МНК

Так как оба структурных уравнения точно идентифицируемы, параметры структурной формы однозначно определяются по параметрам приведенной формы на основе системы уравнений (см. (7.7')):

$$B = AC,$$

$$\begin{pmatrix} b_{11} & b_{12} & 0 \\ b_{21} & 0 & b_{23} \end{pmatrix} = \begin{pmatrix} 1 & a_{12} \\ a_{21} & 1 \end{pmatrix} \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \end{pmatrix} = \begin{pmatrix} c_{11} + c_{21}a_{12} & c_{12} + c_{22}a_{12} & c_{13} + c_{23}a_{12} \\ c_{11}a_{21} + c_{21} & c_{12}a_{21} + c_{22} & c_{13}a_{21} + c_{23} \end{pmatrix}.$$

Из равенства матриц получим оценки:

$$\begin{aligned} b_{11} &= c_{11} + c_{21}a_{12}, & b_{12} &= c_{12} + c_{22}a_{12}, & 0 &= c_{13} + c_{23}a_{12}, \\ b_{21} &= c_{11}a_{21} + c_{21}, & 0 &= c_{12}a_{21} + c_{22}, & b_{23} &= c_{13}a_{21} + c_{23}. \end{aligned}$$

$$a_{12} = \frac{-c_{13}}{c_{23}} = 1,667; \quad a_{21} = \frac{-c_{22}}{c_{12}} = 0,472;$$

$$\begin{aligned} b_{11} &= 32,401 + 6,626 \cdot 1,667 = 43,446; & b_{21} &= 32,401 \cdot 0,472 + 6,626 = 21,919; \\ b_{12} &= 1,218 - 0,575 \cdot 1,667 = 0,258; & b_{23} &= -0,428 \cdot 0,472 + 0,257 = 0,055. \end{aligned}$$

Таким образом, структурные уравнения имеют вид:

$$\hat{Y}_{1t} + 1,667 \hat{Y}_{2t} = 43,446 + 0,258 X_{2t},$$

$$\text{или } \hat{Y}_{1t} = -1,667 \hat{Y}_{2t} + 43,446 + 0,258 X_{2t},$$

$$0,472 \hat{Y}_{1t} + \hat{Y}_{2t} = 21,919 + 0,055 X_{3t},$$

$$\text{или } \hat{Y}_{2t} = -0,472 \hat{Y}_{1t} + 21,919 + 0,055 X_{3t}.$$

Итак, получили структурную модель:

$$\begin{cases} \hat{Y}_{1t} = -1,667 \hat{Y}_{2t} + 43,446 + 0,258 X_{2t}, \\ \hat{Y}_{2t} = -0,472 \hat{Y}_{1t} + 21,919 + 0,055 X_{3t}. \end{cases}$$

По этим уравнениям мы можем сделать следующие выводы:

1. Параметры приведенной формы отражают общее (прямое и косвенное) воздействие предопределенных переменных на совместно зависимые переменные Y_1 и Y_2 (см. рис. 7.3). Так, $c_{12} = 1,218$ указывает на прямое и косвенное (возникающее на основе одновременных соотношений между Y_1 и Y_2) воздействие денежных доходов населения на денежное обращение; $c_{23} = 0,257$ указывает на прямое и косвенное влияние размера вклада в сберегательную кассу на оборачиваемость денег:



Рис. 7.3. Схема взаимодействия эндогенных и предопределенных переменных

2. Параметры структурной формы отражают только непосредственное воздействие факторов, отличие которого от общего воздействия становится очевидным при сравнении $b_{12} = 0,259$ с $c_{12} = 1,218$ или $b_{23} = 0,055$ с $c_{23} = 0,257$.

Кроме того, взаимное влияние между денежным обращением и оборачиваемостью денег можно определить по параметрам a_{12} и a_{21} .

Пример 7.3. Оценивание параметров структурной формы модели двухшаговым методом наименьших квадратов (2МНК)

Применим 2МНК для оценивания конъюнктурной макроэкономической модели, проверка идентифицируемости которой показала, что модель сверхидентифицируема (см. пример 7.1).

1. На первом шаге с помощью МНК построим регрессионное уравнение зависимости объясняющих эндогенных переменных от всех predetermined переменных и вычислим ее теоретические (расчетные) значения.
2. На втором шаге используем теоретические значения эндогенных объясняющих переменных вместо их действительных значений для оценки регрессии с помощью МНК.

1 шаг. Запишем модель в разных формах:

структурная
$Y_1 = \alpha_{14}Y_4 + \beta_{10}X_0 + \beta_{11}X_1 + \varepsilon_1$
$Y_2 = \alpha_{23}Y_3 + \beta_{20}X_0 + \beta_{22}X_2 + \varepsilon_2$
$Y_3 = \alpha_{34}Y_4 + \beta_{30}X_0 + \beta_{33}X_3 + \varepsilon_3$
$Y_4 = Y_1 + Y_2 + X_4$
приведенная
$Y_{1t} = c_{10}X_{0t} + c_{11}X_{1t} + c_{12}X_{2t} + c_{13}X_{3t} + c_{14}X_{4t} + u_{1t}$
$Y_{2t} = c_{20}X_{0t} + c_{21}X_{1t} + c_{22}X_{2t} + c_{23}X_{3t} + c_{24}X_{4t} + u_{2t}$
$Y_{3t} = c_{30}X_{0t} + c_{31}X_{1t} + c_{32}X_{2t} + c_{33}X_{3t} + c_{34}X_{4t} + u_{3t}$
$Y_{4t} = c_{40}X_{0t} + c_{41}X_{1t} + c_{42}X_{2t} + c_{43}X_{3t} + c_{44}X_{4t} + u_{4t}$

Используя данные табл. 7.5, с помощью МНК построим регрессионные уравнения зависимости объясняющих эндогенных переменных от всех predetermined и вычислим теоретические (расчетные) значения эндогенных переменных (рис.7.4).

Таблица 7.5 – Макроэкономические показатели Швеции (млрд. долл. США в ценах 1995 г.) [22], [23]

Год	Потребление(Y ₁)	Инвестиции(Y ₂)	Процентная ставка, %(Y ₃)	ВВП (Y ₄)	Денежная масса (X ₃)	Государственные расходы (X ₄)
1980	61,9	25,8	14,12	125	28,05	37,4
1981	63	24,1	13,71	124,8	35,65	37,7
1982	64,8	23,6	13,91	126,3	38,51	38
1983	64	26,9	15,37	128,6	41,29	37,3
1984	65,1	31	14,36	134,2	44,79	38
1985	67,4	31,2	12,97	137,1	46,27	38,4
1986	69,7	32,3	9,98	140,8	52,58	38,8
1987	73,6	32,9	9,34	145,6	55,71	39
1988	75,2	35	10,62	149,3	59,94	39,1
1989	75,3	37,6	12,62	153,3	66,51	40,4
1990	75,6	36,7	14,85	155	67,02	42,7
1991	78,7	32,5	13,75	153,3	69,81	42,14
1992	78,5	29,5	11,76	150,6	72,10	42,6
1993	78,4	29	7,58	152,3	75,10	44,8
1994	80,2	33,4	8,2	158,7	75,62	45,1
1995	80,6	39,5	8,72	165,1	77,96	45
1996	82	38,7	4,17	167,2	84,87	46,6
1997	84,4	40,1	4,83	171,3	85,82	46,7
1998	87	41,7	3,6	177,5	85,45	48,8
1999	90,7	43,9	3,63	185,7	92,72	51
2000	95,1	46,7	4,14	193,8	94,50	52
2001	102,7	50,3	3,89	210,6	115,70	57,5

С учетом выполненной оценки (рис. 7.4) имеем систему приведенных уравнений:

$$\hat{Y}_{1t} = -4,77 + 0,68X_{1t} + 0,24X_{2t} - 0,01X_{3t} + 0,54X_{4t}, \quad R^2 = 0,99, F = 309,6$$

$$\hat{Y}_{2t} = -7,76 + 0,11X_{1t} + 0,75X_{2t} - 0,01X_{3t} + 0,23X_{4t}, \quad R^2 = 0,89, F = 33,2$$

$$\hat{Y}_{3t} = 24,3 + 0,17X_{1t} + 0,03X_{2t} - 0,14X_{3t} - 0,43X_{4t}, \quad R^2 = 0,72, F = 10,5$$

$$\hat{Y}_{4t} = -12,66 + 0,79X_{1t} + 0,98X_{2t} - 0,02X_{3t} + 1,78X_{4t}, \quad R^2 = 0,98, F = 260,1.$$

Коэффициенты детерминации R^2 и F-критерии Фишера свидетельствуют о достоверности уравнений.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	МНК Оценка приведенных моделей												
2	Исходные данные					Расчетные значения							
3	год	Y1	Y2	Y3	Y4	X1	X2	X3	X4	Y1	Y2	Y3	Y4
4	1980	61,9	25,8	14,12	125			28,05	37,4				
5	1981	63	24,1	13,71	124,8	61,9	25,8	35,65	37,7	63,61	26,54	14,13	127,93
6	1982	64,8	23,6	13,91	126,3	63	24,1	38,51	38	64,08	25,43	13,73	127,60
7	1983	64	26,9	15,37	128,6	64,8	23,6	41,29	37,3	64,78	25,06	13,93	127,23
8	1984	65,1	31	14,36	134,2	64	26,9	44,79	38	65,36	27,56	13,10	130,99
9	1985	67,4	31,2	12,97	137,1	65,1	31	46,27	38,4	67,28	30,83	13,04	136,55
10	1986	69,7	32,3	9,98	140,8	67,4	31,2	52,58	38,8	69,04	31,25	12,37	139,13
11	1987	73,6	32,9	9,34	145,6	69,7	32,3	55,71	39	70,95	32,34	12,26	142,31
12	1988	75,2	35	10,62	149,3	73,6	32,9	59,94	39,1	73,76	33,19	12,29	146,06
13	1989	75,3	37,6	12,62	153,3	75,2	35	66,51	40,4	75,98	35,16	11,14	151,54
14	1990	75,6	36,7	14,85	155	75,3	37,6	67,02	42,7	77,90	37,64	10,17	158,25
15	1991	78,7	32,5	13,75	153,3	75,6	36,7	69,81	42,14	77,56	36,84	10,04	156,54
16	1992	78,5	29,5	11,76	150,6	78,7	32,5	72,10	42,6	78,90	34,11	9,90	155,65
17	1993	78,4	29	7,58	152,3	78,5	29,5	75,10	44,8	79,20	32,32	8,40	156,40
18	1994	80,2	33,4	8,2	158,7	78,4	29	75,62	45,1	79,17	32,00	8,16	156,36
19	1995	80,6	39,5	8,72	165,1	80,2	33,4	77,96	45	81,37	35,44	8,32	161,85
20	1996	82	38,7	4,17	167,2	80,6	39,5	84,87	46,6	83,87	40,34	6,92	170,82
21	1997	84,4	40,1	4,83	171,3	82	38,7	85,82	46,7	84,68	39,91	6,95	171,31
22	1998	87	41,7	3,6	177,5	84,4	40,1	85,45	48,8	87,79	41,70	6,54	178,33
23	1999	90,7	43,9	3,63	185,7	87	41,7	92,72	51	91,06	43,61	5,05	185,70
24	2000	95,1	46,7	4,14	193,8	90,7	43,9	94,50	52	94,63	45,87	5,05	192,52
25	2001	102,7	50,3	3,89	210,6	95,1	46,7	115,70	57,5	101,02	49,48	0,52	208,03
26		Y1(Xj)	ЛИНЕЙН(B5:B25;F5:I25;1;1)					Y2(Xj)	ЛИНЕЙН(C5:C25;F5:I25;1;1)				
27		0,54	-0,01	0,24	0,68	-4,77		0,23	-0,01	0,75	0,11	-7,76	
28		0,20	0,11	0,10	0,23	10,80		0,39	0,21	0,20	0,45	21,40	
29		0,99	1,32	#Н/Д	#Н/Д	#Н/Д		0,89	2,62	#Н/Д	#Н/Д	#Н/Д	
30		309,58	16,00	#Н/Д	#Н/Д	#Н/Д		32,23	16,00	#Н/Д	#Н/Д	#Н/Д	
31		2167,68	28,01	#Н/Д	#Н/Д	#Н/Д		886,16	109,98	#Н/Д	#Н/Д	#Н/Д	
32	t	2,72	-0,11	2,34	3,00	-0,44	t	0,58	-0,05	3,71	0,24	-0,36	
33		Y3(Xj)	ЛИНЕЙН(D5:D25;F5:I25;1;1)					Y4(Xj)	ЛИНЕЙН(E5:T25;F5:I25;1;1)				
34		-0,43	-0,14	0,03	0,17	24,30		1,78	-0,02	0,98	0,79	-12,66	
35		0,37	0,20	0,19	0,43	20,19		0,46	0,25	0,24	0,53	25,22	
36		0,72	2,47	#Н/Д	#Н/Д	#Н/Д		0,98	3,09	#Н/Д	#Н/Д	#Н/Д	
37		10,53	16,00	#Н/Д	#Н/Д	#Н/Д		260,06	16,00	#Н/Д	#Н/Д	#Н/Д	
38		257,85	97,92	#Н/Д	#Н/Д	#Н/Д		9931,24	152,75	#Н/Д	#Н/Д	#Н/Д	
39	t	-1,16	-0,70	0,17	0,39	1,20	t	3,83	-0,10	4,12	1,49	-0,50	
40													

Рис. 7.4. Оценка приведенной формы модели

2 шаг. Найдем оценки структурных параметров методом наименьших квадратов, заменяя фактические значения эндогенных объясняющих переменных их теоретическими значениями (рис. 7.5).

	A	B	C	D	E	F	G	H	I	J	K
48	Оценка структурных параметров										
49		Фактические значения				Для 1 уравнения		Для 2 уравнения		Для 3 уравнения	
50	год	Y1	Y2	Y3	Y4	Y4 расч	X1	Y3 расч	X2	Y4 расч	X3
51	1980	61,9	25,8	14,12	125						28,05
52	1981	63	24,1	13,71	124,8	127,93	61,9	14,13	25,8	127,93	35,65
53	1982	64,8	23,6	13,91	126,3	127,60	63	13,73	24,1	127,60	38,51
54	1983	64	26,9	15,37	128,6	127,23	64,8	13,93	23,6	127,23	41,29
55	1984	65,1	31	14,36	134,2	130,99	64	13,10	26,9	130,99	44,79
56	1985	67,4	31,2	12,97	137,1	136,55	65,1	13,04	31	136,55	46,27
57	1986	69,7	32,3	9,98	140,8	139,13	67,4	12,37	31,2	139,13	52,58
58	1987	73,6	32,9	9,34	145,6	142,31	69,7	12,26	32,3	142,31	55,71
59	1988	75,2	35	10,62	149,3	146,06	73,6	12,29	32,9	146,06	59,94
60	1989	75,3	37,6	12,62	153,3	151,54	75,2	11,14	35	151,54	66,51
61	1990	75,6	36,7	14,85	155	158,25	75,3	10,17	37,6	158,25	67,02
62	1991	78,7	32,5	13,75	153,3	156,54	75,6	10,04	36,7	156,54	69,81
63	1992	78,5	29,5	11,76	150,6	155,65	78,7	9,90	32,5	155,65	72,10
64	1993	78,4	29	7,58	152,3	156,40	78,5	8,40	29,5	156,40	75,10
65	1994	80,2	33,4	8,2	158,7	156,36	78,4	8,16	29	156,36	75,62
66	1995	80,6	39,5	8,72	165,1	161,85	80,2	8,32	33,4	161,85	77,96
67	1996	82	38,7	4,17	167,2	170,82	80,6	6,92	39,5	170,82	84,87
68	1997	84,4	40,1	4,83	171,3	171,31	82	6,95	38,7	171,31	85,82
69	1998	87	41,7	3,6	177,5	178,33	84,4	6,54	40,1	178,33	85,45
70	1999	90,7	43,9	3,63	185,7	185,70	87	5,05	41,7	185,70	92,72
71	2000	95,1	46,7	4,14	193,8	192,52	90,7	5,05	43,9	192,52	94,50
72	2001	102,7	50,3	3,89	210,6	208,03	95,1	0,52	46,7	208,03	115,70
73											
74	Y1(Y4,X1)		Y2(Y3,X2)			Y3(Y4,X3)					
75	ЛИНЕЙН(B52:B72;F52:G72;1;1)			ЛИНЕЙН(C52:C72;H52:I72;1;1)			ЛИНЕЙН(D52:D72;J52:K72;1;1)				
76	0,47	0,27	-0,48		0,77	-0,52	13,89		-0,10	-0,07	27,06
77	0,16	0,07	2,72		0,17	0,31	8,50		0,14	0,13	11,01
78	0,99	1,25	#Н/Д		0,89	2,49	#Н/Д		0,71	2,41	#Н/Д
79	689,47	18,00	#Н/Д		71,53	18,00	#Н/Д		21,54	18,00	#Н/Д
80	2167,39	28,29	#Н/Д		884,81	111,33	#Н/Д		250,92	104,84	#Н/Д
81	2,92	3,98	-0,18	t	4,51	-1,71	1,63	t	-0,70	-0,55	2,46
82											

Рис. 7.5. Оценка структурной формы модели

Получили структурную модель:

$$\hat{Y}_1 = 0,27Y_4 - 0,48 + 0,47X_1$$

$$\hat{Y}_2 = -0,52Y_3 + 13,89 + 0,77X_2$$

$$\hat{Y}_3 = -0,07Y_4 + 27,06 \pm 0,10X_3$$

$$\hat{Y}_4 = Y_1 + Y_2 + X_4$$

Здесь Y_1 —расходы на потребление в текущем периоде; Y_2 — инвестиции в текущем периоде; Y_3 — процентная ставка; Y_4 – ВВП; X_1 —расходы на потребление в прошлом периоде; X_2 —инвестиции в прошлом периоде; X_3 — денежная масса; X_4 — государственные расходы.

И коэффициенты детерминации R^2 , а также F-критерии Фишера свидетельствуют о достоверности уравнений.

Оценки структурной формы показывают, что увеличение государственных расходов (X_4) на 1 млрд. долл. при прочих равных условиях вызывают рост потребительских расходов в среднем на 0,27 млрд. долл.; увеличение процентной ставки (Y_3) на 1% способствует уменьшению инвестиции в текущем периоде (Y_2) на 0,52 млрд. долл. Обратное воздействие на изменение процентной ставки оказывают ВВП и денежная масса: так, при увеличении ВВП (Y_4) на 1 млрд. долл. и неизменной денежной массе (X_3) величина процентной ставки (Y_3) снижается в среднем на 0,07%; если же при неизменном ВВП денежная масса вырастет на 1 млрд. долл., то в среднем уменьшение процентной ставки составит 0,1%.

Как указывалось ранее, 2МНК можно использовать для оценивания точно идентифицируемого уравнения, причем результат (оценки структурных параметров) в точности совпадут с теми, которые позволят получить КМНК. Иллюстрация этого положения выполнена для точно идентифицируемой модели, которая оценивалась косвенным методом наименьших квадратов в примере 7.2 (см. рис. 7.6).

Из рис. 7.6 видно, что полученные с помощью двухшагового метода наименьших квадратов оценки совпадают с соответствующими КМНК-оценками, так как эта модель точно идентифицируема.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	F																	
1	Y1	Y2	X1	X2	X3							2МНК																					
2	25	10	1	1	22				1) этап			Функция ТЕНДЕНЦИЯ																					
3	28	11	1	4	19							для расчета		У1 расч и У2 расч																			
4	29	9	1	3	13				$\hat{Y}_1 = 32,401 + 1,218 X_2 - 0,428 X_3$			У1 расч	X3	У2 расч	X2																		
5	31	7	1	5	16							24,2042	22	11,6997	1																		
6	35	8	1	6	13							29,14114	19	9,204204	4																		
7	38	6	1	6	9							30,49099	13	8,238739	3																		
8	40	4	1	8	6				$\hat{Y}_2 = 6,626 - 0,575 X_2 + 0,258 X_3$			31,64264	16	7,858859	5																		
9	39	2	1	8	8							34,14414	13	6,513514	6																		
10	42	1	1	9	3							35,85586	9	5,486486	6																		
11	43	2	1	10	1							39,57508	6	3,566066	8																		
12												38,71922	8	4,07958	8																		
13	1) этап			КМНК			2) этап			В=АС			42,07658	3	2,220721	9																	
14	ЛИНЕЙН(A2:A11;D2:E11;1;1)												44,15015	1	1,132132	10																	
15	-0,42793	1,217718	32,4009	$a_{12} = \frac{-c_{13}}{c_{23}} = 1,667; a_{21} = \frac{-c_{22}}{c_{12}} = 0,472;$ $b_{11} = 32,401 + 6,626 \cdot 1,667 = 43,446;$ $b_{21} = 32,401 \cdot 0,472 + 6,626 = 21,919;$ $b_{12} = 1,218 - 0,575 \cdot 1,667 = 0,258;$ $b_{23} = -0,428 \cdot 0,472 + 0,257 = 0,055.$						Для У1: ТЕНДЕНЦИЯ(A2:A11;D2:E11;D2:E11;1)																							
16	0,171971	0,415349	4,328579							Для У2: ТЕНДЕНЦИЯ(B2:B11;D2:E11;D2:E11;1)																							
17	0,968435	1,281154	#Н/Д																														
18	107,384	7	#Н/Д																														
19	352,5105	11,48949	#Н/Д																														
20	$\hat{Y}_1 = 32,401 + 1,218 X_2 - 0,428 X_3$												2) этап																				
21										ЛИНЕЙН(A2:A11;L5:M14;1;1)			ЛИНЕЙН(B2:B11;N5:O14;1;1)																				
22	ЛИНЕЙН(B2:B11;D2:E11;1;1)									0,259259			-1,66667			43,44444																	
23	0,256757	-0,57508	6,626126							0,786733			0,669781			8,708072																	
24	0,207086	0,50016	5,21244							0,968435			1,281154			#Н/Д																	
25	0,856374	1,542755	#Н/Д							107,384			7			#Н/Д																	
26	20,86878	7	#Н/Д							352,5105			11,48949			#Н/Д																	
27	99,33934	16,66066	#Н/Д													0,054665			-0,47226			21,92766											
28	$\hat{Y}_2 = 6,626 - 0,575 X_2 + 0,258 X_3$															0,376291			0,410736			18,45631											
29																			0,856374			1,542755			#Н/Д								
30																						20,86878			7			#Н/Д					
31																									99,33934			16,66066			#Н/Д		
32																																	
33																																	
34																																	
35																																	
36																																	
37																																	
38																																	
39																																	
40																																	
41																																	
42																																	
43																																	
44																																	
45																																	
46																																	
47																																	
48																																	
49																																	
50																																	
51																																	
52																																	
53																																	
54																																	
55																																	
56																																	
57																																	
58																																	
59																																	
60																																	
61																																	
62																																	
63																																	
64																																	
65																																	
66																																	
67																																	
68																																	
69																																	
70																																	
71																																	
72																																	
73																																	
74																																	
75																																	
76																																	
77																																	
78																																	
79																																	
80																																	
81																																	
82																																	
83																																	
84																																	
85																																	
86																																	
87																																	
88																																	
89																																	
90																																	
91																																	
92																																	
93																																	
94																																	
95																																	
96																																	
97																																	
98																																	
99																																	
100																																	

Рис. 7.6. Результаты оценивания модели денежного обращения с помощью КМНК и 2МНК

Пример 7.4. Оценивание структурных параметров модели, включающей неидентифицируемые уравнения

Рассмотрим модель, представленную системой одновременных уравнений, имеет вид:

$$\text{функция дохода } Y_{1t} = \alpha_{12}Y_{2t} + \beta_{10} + \beta_{11}X_{1t} + \beta_{12}X_{2t} + \varepsilon_{1t},$$

$$\text{функция предложения денег } Y_{2t} = \alpha_{21}Y_{1t} + \beta_{20} + \varepsilon_{2t},$$

где Y_{1t} – доход; Y_{2t} – запас денег; X_{1t} – инвестиционные затраты; X_{2t} – затраты правительства на товары и услуги.

Переменные Y_{1t} и Y_{2t} – эндогенные, X_{1t} и X_{2t} – экзогенные.

Уравнение дохода показывает, что доход (Y_{1t}) определяется предложением денег (Y_{2t}), инвестиционными затратами (X_{1t}) и затратами правительства (X_{2t}).

Уравнение предложения денег показывает, что запас денег (Y_{2t}) определяется в соответствии с уровнем доходов (Y_{1t}).

Очевидно, рассматриваемая модель представляет систему одновременных уравнений в структурной форме.

Задание. Требуется выбрать метод и найти оценки структурных параметров модели. Сделать выводы.

Решение.

- 1) Проверим идентифицируемость системы, используя условие порядка (необходимое условие):

$$k - k_i \geq m_i - 1.$$

$$k = 3; \quad k_1 = 3; \quad k_2 = 1; \quad m_1 = 2; \quad m_2 = 2.$$

Первое уравнение:

$$3 - 3 \geq 2 - 1, \quad 0 < 1,$$

т.е. необходимое условие идентифицируемости не выполнено, и первое уравнение неидентифицируемо, поэтому *первое уравнение без введения дополнительных условий оценить нельзя*.

Второе уравнение: $3 - 1 \geq 2 - 1$, $2 > 1$ – уравнение
сверхидентифицируемо.

Для оценивания второго уравнения используем 2МНК.

Исходные данные приведены в табл. 7.6.

Таблица 7.6 – Исходные данные (млрд. руб.)

Доход Y_{1t}	Предложение денег Y_{2t}	Инвестиции X_{1t}	Правительственные расходы X_{2t}	\hat{Y}_1
1015,5	216,6	148,8	218,2	1048,09
1102,7	230,8	172,5	232,4	1136,43
1212,8	252	202	250	1246,08
1359,3	265,9	238,8	266,5	1361,25
1472,8	277,5	240,8	299,1	1493,14
1598,4	291,1	219,6	335	1606,72
1782,8	310,3	277,7	356,9	1772,11
1990,5	335,3	344,1	387,3	1982,41
2249,7	363	416,8	425,2	2230,95
2508,2	389	454,8	467,8	2451,17
2732	414,8	437	530,3	2674,76
3052,6	441,8	515	588,1	3009,33
3166	480,8	447,3	641,7	3130,13
3401,6	528	501,9	675,7	3338,73
3774,7	585,5	674	736,8	3813,69
3992,5	624,7	670,4	814,6	4117,12

1 шаг. Учитывая, что во втором уравнении структурной формы модели в качестве объясняющей переменной выступает эндогенная переменная Y_1 (величина дохода), фактические значения которой на втором шаге применения 2МНК придется заменять теоретическими (расчетными), оценим первое уравнение приведенной формы, описывающее зависимость эндогенной переменной Y_1 от всех predetermined. Далее используем полученное уравнение регрессии для нахождения расчетных значений Y_1 .

Построенное регрессионное уравнение зависимости дохода Y_1 (ВВП) от predetermined (экзогенных) переменных с применением МНК (функция ЛИНЕЙН в Excel) имеет вид(см.рис. 7.7).:

$$\hat{Y}_1 = -17,88 + 1,3529X_1 + 3,9627X_2$$

$$(34,181) \quad (0,300) \quad (0,263)$$

$$t: (-0,523) \quad (4,506) \quad (15,090) \quad (7.9)$$

$$R^2 = 0,9976, \quad F = 2681,89.$$

В скобках указаны стандартные ошибки оценок параметров модели и t-статистики Стьюдента, характеризующие статистическую надежность оценок.

	A	B	C	D	E	F	G	H	I
1	ФУНКЦИЯ ДЕНЕЖНОГО ПРЕДЛОЖЕНИЯ. 2МНК.								
2	Y1	Y2	X1	X2	Y1расч				
3	1015,5	216,6	148,8	218,2	1048,09				
4	1102,7	230,8	172,5	232,4	1136,43		Y1 = f (X1, X2)		
5	1212,8	252	202	250	1246,08		ЛИНЕЙН(A3:A18;C3:D18;1;1)		
6	1359,3	265,9	238,8	266,5	1361,25		3,96269	1,35291	-17,88
7	1472,8	277,5	240,8	299,1	1493,14		0,2626	0,30026	34,1809
8	1598,4	291,1	219,6	335	1606,72		0,99758	51,8991	#Н/Д
9	1782,8	310,3	277,7	356,9	1772,11		2681,89	13	#Н/Д
10	1990,5	335,3	344,1	387,3	1982,41		1,4E+07	35015,7	#Н/Д
11	2249,7	363	416,8	425,2	2230,95				
12	2508,2	389	454,8	467,8	2451,17		Y2 = f (Y1расч)		
13	2732	414,8	437	530,3	2674,76		ЛИНЕЙН(B3:B18;E3:E18;1;1)		
14	3052,6	441,8	515	588,1	3009,33		0,12918	81,4689	
15	3166	480,8	447,3	641,7	3130,13		0,00298	7,33777	
16	3401,6	528	501,9	675,7	3338,73		0,99263	11,3093	
17	3774,7	585,5	674	736,8	3813,69		1884,9	14	
18	3992,5	624,7	670,4	814,6	4117,12		241079	1790,6	
19									
20							ТЕНДЕНЦИЯ(A3:A18;C3:D18;C3:D18;1)		
21									
22									
23							2,30333	9,00683	
24									

Рис. 7.7. Результаты расчетов

2 шаг. Оценим функцию денежного предложения методом наименьших квадратов (функция ЛИНЕЙН в Excel), заменив эмпирические (фактические) значения эндогенной переменной Y_1 теоретическими (вычисленными по формуле (7.9) с использованием функции ТЕНДЕНЦИЯ в Excel (рис. 7.7):

$$\hat{Y}_2 = 81,4689 + 0,1292\hat{Y}_1$$

$$(7,3378) \quad (0,0030)$$

$$t: (11,103) \quad (43,416)$$

$$R^2 = 0,9926, \quad F = 1884,9.$$

Величина t-статистик достаточно велика и показывает, что оценки параметров являются статистически надежными. Коэффициент детерминации и F-критерий Фишера свидетельствуют о достоверности уравнения.

Стандартные ошибки по отношению к величине оценки составляют:

$$\frac{7,3378}{81,4689} \cdot 100 = 2,3\%; \quad \frac{0,0030}{0,1292} \cdot 100 = 9,0\%,$$

что говорит о несущественном смещении оценок.

Полученное уравнение позволяет сделать вывод о том, что увеличение дохода на 1 млрд. руб. способствует увеличению денежной массы (предложения денег) в среднем на 0,1292 млрд. руб.

Обратимся теперь к первому уравнению модели, которое оказалось неидентифицируемым и тем самым не позволило выполнить оценивание параметров всей модели.

Известно, что для того чтобы сделать подобное уравнение идентифицированным, **необходимо ввести дополнительную предопределенную переменную в уравнение, смежное с неидентифицируемым**. В качестве таких переменных часто используются лаговые эндогенные переменные.

В рассматриваемом примере, чтобы сделать уравнение спроса идентифицируемым, добавим во второе уравнение модели лаговое значение предложения денег (запаса денег) $Y_{2,t-1}$. Тогда введенная предопределенная переменная $X_{3t} = Y_{2,t-1}$ снимает проблему неидентифицируемости для первого уравнения модели. В результате все уравнения модели становятся идентифицируемыми. Убедимся в этом, выполнив проверку с помощью условия порядка.

Получившаяся после введения лаговой переменной во второе уравнение структурная модель имеет вид:

$$\text{функция дохода } Y_{1t} = \alpha_{12}Y_{2t} + \beta_{10} + \beta_{11}X_{1t} + \beta_{12}X_{2t} + \varepsilon_{1t},$$

$$\text{функция предложения денег } Y_{2t} = \alpha_{21}Y_{1t} + \beta_{20} + \beta_{23}X_{3t} + \varepsilon_{2t}, \quad (7.10)$$

где $X_{3t} = Y_{2,t-1}$ —денежная масса в прошлом периоде (лаговая эндогенная переменная, включаемая в состав predetermined).

$$\text{Тогда } k = 4; \quad k_1 = 3; \quad k_2 = 2; \quad m_1 = 2; \quad m_2 = 2.$$

И условие

$$k - k_i \geq m_i - 1$$

оказывается выполненным :

$$\text{для первого уравнения } 4 - 3 \geq 2 - 1, \quad 1 = 1,$$

т.е. теперь первое уравнение оказалось точно идентифицируемым, и его параметры могут быть оценены КМНК;

$$\text{и для второго уравнения } 4 - 2 \geq 2 - 1, \quad 2 > 1,$$

которое продолжает оставаться сверхидентифицируемым.

Для оценки параметров первого уравнения (которое теперь точно идентифицируемо) применим КМНК, а второго – 2МНК.

Используя КМНК, оценим параметры первого уравнения и, заодно, убедимся в невозможности однозначного оценивания с помощью этого метода параметров сверхидентифицируемого уравнения.

Представим модель (7.10) в приведенной форме (7.11):

$$\text{функция дохода } Y_{1t} = c_{10} + c_{11}X_{1t} + c_{12}X_{2t} + c_{13}X_{3t} + u_{1t},$$

$$\text{функция предложения денег } Y_{2t} = c_{20} + c_{21}X_{1t} + c_{22}X_{2t} + c_{23}X_{3t} + u_{2t} \quad (7.11)$$

и оценим ее параметры методом наименьших квадратов (см. рис.7.8).

Из рис. 7.8 приведенная форма модели представляется в виде:

$$\hat{Y}_{1t} = 296,215 + 1,662X_{1t} + 5,93X_{2t} - 3,765X_{3t}$$

$$\hat{Y}_{2t} = 0,323 + 0,026X_{1t} + 0,058X_{2t} + 0,971X_{3t}$$

87	Фактические значения					Для 1 уравнения			Для 2 уравнения	
88	Y1	Y2	X1	X2	x3=Y(2,t-1)	Y2 ras	X1	X2	Y1 ras	x3=Y(2,t-
89	1015,5	216,6	148,8	218,2						
90	1102,7	230,8	172,5	232,4	216,6	228,665	172,5	232,4	1145,673	216,6
91	1212,8	252	202	250	230,8	244,247	202	250	1245,621	230,8
92	1359,3	265,9	238,8	266,5	252	266,752	238,8	266,5	1324,829	252
93	1472,8	277,5	240,8	299,1	265,9	282,198	240,8	299,1	1469,141	265,9
94	1598,4	291,1	219,6	335	277,5	294,998	219,6	335	1603,115	277,5
95	1782,8	310,3	277,7	356,9	291,1	310,992	277,7	356,9	1778,361	291,1
96	1990,5	335,3	344,1	387,3	310,3	333,135	344,1	387,3	1996,726	310,3
97	2249,7	363	416,8	425,2	335,3	361,511	416,8	425,2	2248,204	335,3
98	2508,2	389	454,8	467,8	363	391,877	454,8	467,8	2459,708	363
99	2732	414,8	437	530,3	389	420,296	437	530,3	2702,859	389
100	3052,6	441,8	515	588,1	414,8	450,744	515	588,1	3078,141	414,8
101	3166	480,8	447,3	641,7	441,8	478,315	447,3	641,7	3181,804	441,8
102	3401,6	528	501,9	675,7	480,8	519,587	501,9	675,7	3327,365	480,8
103	3774,7	585,5	674	736,8	528	573,459	674	736,8	3798,076	528
104	3992,5	624,7	670,4	814,6	585,5	633,726	670,4	814,6	4036,978	585,5
105										
106	Оценка приведенных параметров									
107	ЛИНЕЙН(A90:A104;C90:E104;1;1)					ЛИНЕЙН(B90:B104;C90:E104;1;1)				
108		-3,764	5,930	1,662	296,215		0,971	0,058	0,026	0,323
109	Y1(x1, x2,x3)	1,101	0,604	0,248	93,151	Y2(x1, x2,x3)	0,199	0,109	0,045	16,803
110		0,999	38,550	#Н/Д	#Н/Д		0,998	6,954	#Н/Д	#Н/Д
111		2864,818	11	#Н/Д	#Н/Д		1485,005	11	#Н/Д	#Н/Д
112		12771979,535	16346,794	#Н/Д	#Н/Д		215424,423	531,910	#Н/Д	#Н/Д
113	t	-3,418	9,817	6,706	3,180	t	4,888	0,534	0,583	0,019
114										
115	Оценка структурных параметров									
116	ЛИНЕЙН(A90:A104;F90:H104;1;1)					ЛИНЕЙН(B90:B104;I90:J104;1;1)				
117		6,155	1,763	-3,877	297,469		1,007	0,012	-2,520	
118	Y1(Y2, x1,x2)	0,667	0,261	1,134	93,502	Y2(Y1, x3)	0,129	0,015	11,787	
119		0,999	38,550	#Н/Д	#Н/Д		0,998	6,669	#Н/Д	
120		2864,818	11,000	#Н/Д	#Н/Д		2422,089	12,000	#Н/Д	
121		12771979,535	16346,794	#Н/Д	#Н/Д		215422,688	533,645	#Н/Д	
122	t	9,231	6,747	-3,418	3,181	t	7,780	0,760	-0,214	
123										

Рис. 7.8. Расчет приведенных и структурных коэффициентов с помощью 2МНК

Оценим структурные параметры:

$$A = \begin{bmatrix} 1 & -\alpha_{12} \\ -\alpha_{21} & 1 \end{bmatrix} B = \begin{bmatrix} \beta_{10} & \beta_{11} & \beta_{12} & 0 \\ \beta_{20} & 0 & 0 & \beta_{23} \end{bmatrix}$$

$$C = \begin{bmatrix} c_{10} & c_{11} & c_{12} & c_{13} \\ c_{20} & c_{21} & c_{22} & c_{23} \end{bmatrix} = \begin{bmatrix} 296,215 & 1,662 & 5,93 & -3,765 \\ 0,323 & 0,026 & 0,058 & 0,971 \end{bmatrix}$$

$$AC = \begin{bmatrix} 296,215 - 0,323\alpha_{12} & 1,662 - 0,026\alpha_{12} & 5,93 - 0,058\alpha_{12} & -3,765 - 0,971\alpha_{12} \\ -296,215\alpha_{21} + 0,323 & -1,662\alpha_{21} + 0,026 & -5,93\alpha_{21} + 0,058 & 3,765\alpha_{21} + 0,971 \end{bmatrix}$$

Условие, связывающее параметры структурной и сприведенной форм:

$$AC = B.$$

В табл. 7.7 и 7.8 выполнено сравнение элементов матриц AC и B для первого и второго уравнений соответственно.

Таблица 7.7 –Сравнение элементов матриц AC и B для первого уравнения

Реализация условия AC = B	Расчет оценок параметров для второго уравнения
$296,215 - 0,323\alpha_{12} = \beta_{10}$ $-296,215\alpha_{21} + 0,323 = \beta_{20}$	$\beta_{10} = 296,215 - 0,323 \cdot (-3,877) = 297,469$
$1,662 - 0,026\alpha_{12} = \beta_{11}$ $-1,662\alpha_{21} + 0,026 = 0$	$\beta_{11} = 1,662 - 0,026 \cdot (-3,877) = 1,763$
$5,93 - 0,058\alpha_{12} = \beta_{12}$ $-5,93\alpha_{21} + 0,058 = 0$	$\beta_{12} = 5,93 - 0,058 \cdot (-3,877) = 6,155$
$-3,765 - 0,971\alpha_{12} = 0$ $3,765\alpha_{21} + 0,971 = \beta_{23}$	$\alpha_{12} = \frac{3,765}{-0,971} = -3,877$

Таблица 7.8 –Сравнение элементов матриц AC и B для второго уравнения

Реализация условия AC = B	Расчет оценок параметров для первого уравнения
$296,215 - 0,323\alpha_{12} = \beta_{10}$ $-296,215\alpha_{21} + 0,323 = \beta_{20}$	
$1,662 - 0,026\alpha_{12} = \beta_{11}$ $-1,662\alpha_{21} + 0,026 = 0$	$\alpha_{21} = \frac{0,026}{1,662} = 0,016$
$5,93 - 0,058\alpha_{12} = \beta_{12}$ $-5,93\alpha_{21} + 0,058 = 0$	$\alpha_{21} = \frac{0,058}{5,93} = 0,098$
$-3,765 - 0,971\alpha_{12} = 0$ $3,765\alpha_{21} + 0,971 = \beta_{23}$	

Данные табл. 7.7 позволяют записать оцененное структурное уравнение дохода (которое теперь оказалось точно идентифицируемым) следующим образом:

$$\hat{Y}_{1t} = -3,877Y_{2t} + 291,469 + 1,763X_{1t} + 6,155X_{2t}. \quad (7.12)$$

Из табл. 7.8 видно, что для сверхидентифицируемого второго уравнения не удастся получить однозначный результат: величина коэффициента α_{21} имеет различные оценки.

Из рис. 7.8 видно, что полученные с помощью 2МНК оценки для первого уравнения совпадают с теми, что позволил получить КМК (см. уравнение (7.12)).

Вопросы и задания для самостоятельной работы

- 1) Каково назначение систем эконометрических уравнений?
- 2) Приведите основные типы систем уравнений?
- 3) Охарактеризуйте переменные систем эконометрических уравнений (эндогенные, экзогенные, predetermined).
- 4) В чем отличие системы взаимозависимых уравнений от системы независимых уравнений? В чем особенность системы рекурсивных уравнений?
- 5) Что представляет собой структурная форма модели, приведенная форма модели? Как они связаны между собой?
- 6) Для чего нужна приведенная форма модели?
- 7) В чем суть понятия «идентифицируемость модели»?
- 8) Дайте определение точной идентифицируемости, неидентифицируемости, сверхидентифицируемости отдельных уравнений, модели в целом
- 9) Как идентифицируется отдельное уравнение в системе по счетному правилу?
- 10) В чем состоит достаточное условие идентифицируемости отдельного уравнения?
- 11) В чем суть понятия «идентификации модели»?
- 12) Какие методы идентификации систем эконометрических уравнений вы знаете?

Задача 1. Рассмотрите расширенную кейнсианскую модель определения дохода:

функция потребления $C_t = \beta_1 + \beta_2 Y_t + \beta_3 T_t + \varepsilon_{1t},$

функция инвестиций $I_t = \alpha_0 + \alpha_1 Y_{t-1} + \varepsilon_{2t},$

функция налогообложения $T_t = \gamma_0 + \gamma_1 Y_t + \varepsilon_{3t},$

тождество дохода $Y_t = C_t + I_t + G_t,$

где C_t – совокупное потребление в период t ; I_t – инвестиции в период времени t ; T_t – налоги в период времени t ; Y_t – совокупный доход в период t ; G_t – государственные расходы в период времени t ; Y_{t-1} – совокупный доход в период $t-1$.

- 1) Классифицируйте переменные на эндогенные, экзогенные и предопределенные.
- 2) Проверьте идентифицируемость модели, применив необходимое и достаточное условия. Укажите, каким методом вы будете оценивать структурные параметры каждого уравнения.
- 3) Напишите приведенную форму модели. Для чего она используется?
- 4) Что произойдет, если в правую часть функции инвестиций ввести экзогенную переменную – ставку процента r_t ?

Задача 2. Изучается зависимость потребления (C) от доходов (Y). Каким методом вы будете определять параметры функции потребления, если эконометрическая модель имеет следующий вид:

Модель А: $C = a + bY + \varepsilon$ (функция потребления);

Модель Б: $C = a + bY + \varepsilon$ (функция потребления);

$Y = C + I$ (тождество дохода);

Модель В: $C = a + bY + \varepsilon$ (функция потребления);

$$Y = C + I + G \quad (\text{тождество дохода}),$$

где I – инвестиции; G – государственные расходы.

Переменные C, Y – эндогенные.

Дайте развернутый ответ по каждой из моделей А – В, включающий обоснование выбранного вами метода и краткое описание методики расчетов.

Задача 3. Результаты оценивания параметров некоторой эконометрической модели выглядят следующим образом.

Структурная форма модели:

$$Y_1 = -4 + ??? Y_2 - 9,4X_2 + \varepsilon_1,$$

$$Y_2 = 12,83 - 2,67Y_1 + ??? X_1 + \varepsilon_2,$$

$$Y_3 = 1,36 - 1,76Y_1 + 0,828Y_2 + \varepsilon_3.$$

Приведенная форма модели:

$$Y_1 = 2 + 4X_1 - 3X_2 + u_1,$$

$$Y_2 = 7,5 + 5X_1 + 8X_2 + u_2,$$

$$Y_3 = 4 + ??? X_1 + ??? X_2 + u_3.$$

- 1) Какими методами оценены параметры структурной и приведенной форм модели?
- 2) Обоснуйте возможность применения косвенного метода наименьших квадратов для расчета структурных оценок.
- 3) Восстановите пропущенные характеристики

Задача 4. Модель Менгеса:

$$Y_t = a_1 + b_{11} Y_{t-1} + b_{12} I_t + \varepsilon_1;$$

$$I_t = a_2 + b_{21} Y_t + b_{22} Q_t + \varepsilon_2;$$

$$C_t = a_3 + b_{31} Y_t + b_{32} C_{t-1} + b_{33} P_t + \varepsilon_3;$$

$$Q_t = a_4 + b_{41} Q_{t-1} + b_{42} R_t + \varepsilon_4;$$

где Y – национальный доход; C – расходы на личное потребление; I – инвестиции; Q – валовая прибыль экономики; P – индекс стоимости жизни; R – объем продукции промышленности; t – текущий период; $t-1$ – предыдущий период.

Задание:

- 1) выделить эндогенные и экзогенные переменные;
- 2) применив необходимое и достаточное условие идентификации, определить, идентифицировано ли каждое из уравнений модели;
- 3) определить метод оценки параметров модели;
- 4) записать приведенную форму модели;
- 5) определить коэффициенты приведенной формы модели;
- 6) определить коэффициенты структурной формы модели;
- 7) проверить значимость полученных уравнений и их коэффициентов.

Таблица 7.4 - Исходные данные

Текущий период t	Внутренние инвестиции I (млн руб.)	Национальный доход Y (млн руб.)	Расходы на личное потребление C (млн руб.)	Валовая прибыль экономики Q (млн руб.)	Индекс стоимости жизни P (%)	Объем продукции промышленности R (млн руб.)
1	211	310	450	725,6	200	600
2	2 670	5 328	7500	11 390,5	210	1 300
3	27 125	49 730	40 600	76 961,7	220	8 500
4	108 810	172 380	124 000	251 944,4	238	129 000
5	266 974	437 007	310 000	662 374,4	195	384 000
6	375 998	558 500	260 000	790 819,2	208	1 108 000
7	408 797	711 600	390 000	881 001,1	229	1 469 000
8	407 086	686 000	490 000	1 032 768,6	204	1 626 000
9	970 439	1 213 600	990 000	2 050 276,8	180	1 707 000
10	1 165 181	2 097 700	1 650 000	3 033 247,2	181	3 150 000

СПИСОК ЛИТЕРАТУРЫ

1. Frisch, R. Editorial / R. Frisch // *Econometrica*. – 1933. – № 1. – Р. 2.
2. Айвазян, С. А. Прикладная статистика и основы эконометрики : учеб. пособие для вузов / С. А. Айвазян, В. С. Мхитарян. – Москва : ЮНИТИ, 1998. – 1022 с.
3. Бородич, С. А. Эконометрика : учеб. пособие / С. А. Бородич. – 2-е изд., испр. – Минск : Новое знание, 2006. – 408 с.
4. Доугерти, К. Введение в эконометрику : учеб. для вузов : пер. с англ. / К. Доугерти. - 2-е изд. - Москва : ИНФРА-М, 2004. – 432 с.
5. Методические указания к выполнению контрольных заданий по дисциплине «Эконометрия» / сост.: В. Н. Калинина, В. И. Соловьев. – Москва : [б. и.], 2004. – 33 с.
6. Методические указания к лабораторным и контрольным работам курса «Эконометрика» [Электронный ресурс] / сост.: Ю. Е. Воскобойников, Т. Н. Воскобойникова. - Электрон. дан. - Новосибирск : [б. и.], 2006. – Режим доступа: <http://window.edu.ru/resource/306/63306/files/econometr.pdf>. - Загл. с экрана.
7. Кейн, Э. Экономическая статистика и эконометрия : введение в количественный экономический анализ / Э. Кейн. – Москва : Статистика, 1977. – 255 с.
8. Кремер, Н. Ш. Эконометрика : учеб. для вузов / Н. Ш. Кремер, Б. А. Путко. – Москва : ЮНИТИ-ДАНА, 2002. – 311 с.
9. Лондар, С. Л. Економетрія засобами MS Excel : навч. посіб. / С. Л. Лондар, Р. В. Юринец. – Київ : Вид-во Європ. ун-ту, 2004. – 242 с.
10. Магнус, Я. Р. Эконометрика : начальный курс : учеб. для вузов / Я. Р. Магнус, П. К. Катышев, А. А. Пересецкий. – 6-е изд., перераб. и доп. – Москва : Дело, 2004. – 576 с.
11. Носко, В. П. Эконометрика : учебник. Кн. 1 / В. П. Носко. – Москва : Дело, 2011. – 672 с.
12. Практикум по эконометрике (+CD) : учеб. пособие / И. И. Елисеева [и др.] ; под ред. И. И. Елисейвой. – 2-е изд., перераб. и доп. – Москва : Финансы и статистика, 2007. – 344 с. : ил.

13. Айвазян, С. А. Прикладная статистика. Основы эконометрики : учеб. для вузов : в 2 т. Т. 1. Теория вероятностей и прикладная статистика / С. А. Айвазян, В. С. Мхитарян. – Москва : ЮНИТИ-ДАНА, 2001. – 656 с.
14. Айвазян, С. А. Прикладная статистика. Основы эконометрики : учеб. для вузов : в 2 т. Т. 2. Основы эконометрики / С. А. Айвазян. – Москва : ЮНИТИ-ДАНА, 2001. – 432 с.
15. Ферстер, Э. Методы корреляционного и регрессионного анализа : руководство для экономистов : пер. с нем. / Э. Фёрстер, Б. Рёнц. – Москва : Финансы и статистика, 1983. – 302 с.
16. Эконометрика : учебник / К. В. Балдин [и др.] ; под ред. В. Б. Уткина. – 2-е изд. – Москва : Дашков и К°, 2012. – 564 с.
17. Эконометрика : учебник / И. И. Елисеева [и др.] ; под ред. И. И. Елисеевой. – 2-е изд., перераб. и доп. – Москва : Финансы и статистика, 2007. – 576 с. : ил.
18. Слепнева, Л. Д. Эконометрика [Электронный ресурс] : учеб. пособие / Л. Д. Слепнева, Е. А. Рытикова ; ГОУВПО «ДОННТУ». – Электрон. дан. (1 файл). – Донецк : ДОННТУ, 2019. – Систем. требования: Acrobat Reader.
19. Сборник задач к начальному курсу эконометрики : учеб. пособие / П. К. Катышев [и др.]. – 4-е изд., перераб. и доп. – Москва : Дело, 2007. – 368 с.
20. Наконечний, С. І. Економетрія : навч.-метод. посіб. для самостійного вивч. дисципліни / С. І. Наконечний, Т. О. Терещенко. – Київ : КНЕУ, 2001. – 192 с.
21. Эконометрия : учеб. пособие / В. И. Жлуктенко [и др.]. – Киев : Изд-во Европ. ун-та, 2005. – 552 с.
22. Колемаев, В. А. Эконометрика : учебник / В. А. Колемаев. – Москва : ИНФРА-М, 2007. – 160 с.
23. Евростат : [стат. служба Европейского Союза] [Электронный ресурс] // TrendEconomy.ru : сайт. - Электрон. дан. - [Россия]. - 2019. - Режим доступа: <http://data.trendeconomy.ru/dataviewer>. - Загл. с экрана.

ПРИЛОЖЕНИЕ А

МАТЕМАТИКО-СТАТИСТИЧЕСКИЕ ТАБЛИЦЫ

Таблица А.1 – Значения F -критерия Фишера при уровне значимости $\alpha = 0,05$

$k_1 \backslash k_2$	1	2	3	4	5	6	8	12	24	∞
1	2	3	4	5	6	7	8	9	10	11
1	161,5	199,5	215,7	224,6	230,2	233,9	238,9	243,9	249,0	254,3
2	18,51	19,00	19,16	19,25	19,30	19,33	19,37	19,41	19,45	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,84	8,74	8,64	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,04	5,91	5,77	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,82	4,68	4,53	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4,00	3,84	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,73	3,57	3,41	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,28	3,12	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,23	3,07	2,90	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,07	2,91	2,74	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	2,95	2,79	2,61	2,40
12	4,75	3,88	3,49	3,26	3,11	3,00	2,85	2,69	2,50	2,30
13	4,67	3,80	3,41	3,18	3,02	2,92	2,77	2,60	2,42	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,70	2,53	2,35	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,64	2,48	2,29	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,59	2,42	2,24	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,55	2,38	2,19	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,51	2,34	2,15	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,48	2,31	2,11	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,45	2,28	2,08	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,42	2,25	2,05	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,40	2,23	2,03	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,38	2,20	2,00	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,36	2,18	1,98	1,73
25	4,24	3,38	2,99	2,76	2,60	2,49	2,34	2,16	1,96	1,71
26	4,22	3,37	2,98	2,74	2,59	2,47	2,32	2,15	1,95	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,30	2,13	1,93	1,67
28	4,20	3,34	2,95	2,71	2,56	2,44	2,29	2,12	1,91	1,65
29	4,18	3,33	2,93	2,70	2,54	2,43	2,28	2,10	1,90	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,27	2,09	1,89	1,62
35	4,12	3,26	2,87	2,64	2,48	2,37	2,22	2,04	1,83	1,57
40	4,08	3,23	2,84	2,61	2,45	2,34	2,18	2,00	1,79	1,51
45	4,06	3,21	2,81	2,58	2,42	2,31	2,15	1,97	1,76	1,48
50	4,03	3,18	2,79	2,56	2,40	2,29	2,13	1,95	1,74	1,44
60	4,00	3,15	2,76	2,52	2,37	2,25	2,10	1,92	1,70	1,39
70	3,98	3,13	2,74	2,50	2,35	2,23	2,07	1,89	1,67	1,35
80	3,96	3,11	2,72	2,49	2,33	2,21	2,06	1,88	1,65	1,31
90	3,95	3,10	2,71	2,47	2,32	2,20	2,04	1,86	1,64	1,28
100	3,94	3,09	2,70	2,46	2,30	2,19	2,03	1,85	1,63	1,26
125	3,92	3,07	2,68	2,44	2,29	2,17	2,01	1,83	1,60	1,21
150	3,90	3,06	2,66	2,43	2,27	2,16	2,00	1,82	1,59	1,18
200	3,89	3,04	2,65	2,42	2,26	2,14	1,98	1,80	1,57	1,14
300	3,87	3,03	2,64	2,41	2,25	2,13	1,97	1,79	1,55	1,10
400	3,86	3,02	2,63	2,40	2,24	2,12	1,96	1,78	1,54	1,07
500	3,86	3,01	2,62	2,39	2,23	2,11	1,96	1,77	1,54	1,06
1000	3,85	3,00	2,61	2,38	2,22	2,10	1,95	1,76	1,53	1,03
∞	3,84	2,99	2,60	2,37	2,21	2,09	1,94	1,75	1,52	1

Таблица А.2 – Критические значения t -критерия Стьюдента при уровне значимости 0,10, 0,05, 0,01 (двухсторонний)

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

Таблица А.3 – Критические значения хи-квадрат (χ^2) при уровне значимости 0,01, 0,05, 0,1

Число степеней свободы	α			Число степеней свободы	α			Число степеней свободы	α		
	0,01	0,05	0,1		0,01	0,05	0,1		0,01	0,05	0,1
1	6,6349	3,8415	2,7055	18	34,805	28,869	25,989	35	57,342	49,802	46,059
2	9,2103	5,9915	4,6052	19	36,191	30,144	27,204	36	58,619	50,998	47,212
3	11,345	7,8147	6,2514	20	37,566	31,41	28,412	37	59,893	52,192	48,363
4	13,277	9,4877	7,7794	21	38,932	32,671	29,615	38	61,162	53,384	49,513
5	15,086	11,07	9,2364	22	40,289	33,924	30,813	39	62,428	54,572	50,66
6	16,812	12,592	10,645	23	41,638	35,172	32,007	40	63,691	55,758	51,805
7	18,475	14,067	12,017	24	42,98	36,415	33,196	41	64,95	56,942	52,949
8	20,09	15,507	13,362	25	44,314	37,652	34,382	42	66,206	58,124	54,09
9	21,666	16,919	14,684	26	45,642	38,885	35,563	43	67,459	59,304	55,23
10	23,209	18,307	15,987	27	46,963	40,113	36,741	44	68,71	60,481	56,369
11	24,725	19,675	17,275	28	48,278	41,337	37,916	45	69,957	61,656	57,505
12	26,217	21,026	18,549	29	49,588	42,557	39,087	46	71,201	62,83	58,641
13	27,688	22,362	19,812	30	50,892	43,773	40,256	47	72,443	64,001	59,774
14	29,141	23,685	21,064	31	52,191	44,985	41,422	48	73,683	65,171	60,907
15	30,578	24,996	22,307	32	53,486	46,194	42,585	49	74,919	66,339	62,038
16	32	26,296	23,542	33	54,776	47,4	43,745	50	76,154	67,505	63,167
17	33,409	27,587	24,769	34	56,061	48,602	44,903				

Таблица А.4 – Значения статистик Дарбина – Уотсона при уровне значимости $\alpha = 0,01$

n*	m* = 1		m* = 2		m* = 3		m* = 4		m* = 5	
	DW _L	DW _U								
6	0,390	1,142	-	-	-	-	-	-	-	-
7	0,435	1,036	0,294	1,676	-	-	-	-	-	-
8	0,497	1,003	0,345	1,489	0,229	2,102	-	-	-	-
9	0,554	0,998	0,408	1,389	0,279	1,875	0,183	2,433	-	-
10	0,604	1,001	0,466	1,333	0,340	1,733	0,230	2,193	0,150	2,690
11	0,653	1,010	0,519	1,297	0,396	1,640	0,286	2,030	0,193	2,453
12	0,697	1,023	0,569	1,274	0,449	1,575	0,339	1,913	0,244	2,280
13	0,738	1,038	0,616	1,261	0,499	1,526	0,391	1,826	0,294	2,150
14	0,776	1,054	0,660	1,254	0,547	1,490	0,441	1,757	0,343	2,049
15	0,811	1,070	0,700	1,252	0,591	1,464	0,488	1,704	0,391	1,967
16	0,844	1,086	0,737	1,252	0,633	1,446	0,532	1,663	0,437	1,900
17	0,874	1,102	0,772	1,255	0,672	1,432	0,574	1,630	0,480	1,847
18	0,902	1,118	0,805	1,259	0,708	1,422	0,613	1,604	0,522	1803
19	0,928	1,132	0,835	1,265	0,742	1,415	0,650	1,584	0,561	1,767
20	0,952	1,147	0,863	1,271	0,773	1,411	0,685	1,567	0,598	1,737
21	0,975	1,161	0,890	1,277	0,803	1,408	0,718	1,554	0,633	1,712
22	0,997	1,174	0,914	1,284	0,831	1,407	0,748	1,543	0,667	1,691
23	1,018	1,187	0,938	1,291	0,858	1,407	0,777	1,534	0,698	1,673
24	1,037	1,199	0,960	1,298	0,882	1,407	0,805	1,528	0,728	1,658
25	1,055	1,211	0,981	1,305	0,906	1,409	0,831	1,523	0,756	1,645
26	1,072	1,222	1,001	1,312	0,928	1,411	0,855	1,518	0,783	1,635
27	1,089	1,233	1,019	1,319	0,949	1,413	0,878	1,515	0,808	1,626
28	1,104	1,244	1,037	1,325	0,969	1,415	0,900	1,513	0,832	1,618
29	1,119	1,254	1,054	1,332	0,988	1,418	0,921	1,512	0,855	1,611
30	1,133	1,263	1,070	1,339	1,006	1,421	0,941	1,511	0,877	1,606
31	1,147	1,273	1,085	1,345	1,023	1,425	0,960	1,510	0,897	1,601
32	1,160	1,282	1,100	1,352	1,040	1,428	0,979	1,510	0,917	1,597
33	1,172	1,291	1,114	1,358	1,055	1,432	0,996	1,510	0,936	1,594
34	1,184	1,299	1,128	1,364	1,070	1,435	1,012	1,511	0,954	1,591
35	1,195	1,307	1,140	1,370	1,085	1,439	1,028	1,512	0,971	1,589
36	1,206	1,315	1,153	1,376	1,098	1,442	1,043	1,513	0,988	1,588
37	1,217	1,323	1,165	1,382	1,112	1,446	1,058	1,514	1,004	1,586
38	1,227	1,330	1,176	1,388	1,124	1,449	1,072	1,515	1,019	1,585
39	1,237	1,337	1,187	1,393	1,137	1,453	1,085	1,517	1,034	1,584
40	1,246	1,344	1,198	1,398	1,148	1,457	1,098	1,518	1,048	1,584
45	1,288	1,376	1,245	1,423	1,201	1,474	1,156	1,528	1,111	1,584
50	1,324	1,403	1,285	1,446	1,245	1,491	1,205	1,538	1,164	1,587
55	1,356	1,427	1,320	1,466	1,284	1,506	1,247	1,548	1,209	1,592
60	1,383	1,449	1,350	1,484	1,317	1,520	1,283	1,558	1,249	1,598
65	1,407	1,468	1,377	1,500	1,346	1,534	1,315	1,568	1,283	1,604
70	1,429	1,485	1,400	1,515	1,372	1,546	1,343	1,578	1,313	1,611
75	1,448	1,501	1,423	1,529	1,395	1,557	1,368	1,587	1,340	1,617
80	1,466	1,515	1,441	1,541	1,416	1,568	1,390	1,595	1,364	1,624
85	1,482	1,528	1,458	1,553	1,435	1,578	1,411	1,603	1,386	1,630
90	1,496	1,540	1,474	1,563	1,452	1,587	1,429	1,611	1,406	1,636
95	1,510	1,552	1,489	1,573	1,468	1,596	1,446	1,618	1,425	1,642
100	1,522	1,562	1,503	1,583	1,482	1,604	1,462	1,625	1,441	1,647
150	1,611	1,637	1,598	1,651	1,584	1,665	1,571	1,679	1,557	1,693
200	1,664	1,684	1,653	1,693	1,643	1,704	1,633	1,715	1,623	1,725

*n – число наблюдений, m – количество независимых переменных

Таблица А.5 – Значения статистик Дарбина – Уотсона при уровне значимости $\alpha = 0,05$

N*	$m^* = 1$		$m^* = 2$		$m^* = 3$		$m^* = 4$		$m^* = 5$	
	DW_L	DW_U								
6	0,610	1,400	-	-	-	-	-	-	-	-
7	0,700	1,356	0,467	1,896	-	-	-	-	-	-
8	0,763	1,332	0,559	1,777	0,368	2,287	-	-	-	-
9	0,824	1,320	0,629	1,699	0,455	2,128	0,296	2,588	-	-
10	0,879	1,320	0,697	1,641	0,525	2,016	0,376	2,414	0,243	2,822
11	0,927	1,324	0,658	1,604	0,595	1,928	0,444	2,283	0,32	2,645
12	0,971	1,331	0,812	1,579	0,658	1,864	0,512	2,177	0,379	2,506
13	1,010	1,340	0,861	1,562	0,715	1,816	0,574	2,094	0,445	2,390
14	1,045	1,350	0,905	1,551	0,767	1,779	0,632	2,030	0,505	2,296
15	1,077	1,361	0,946	1,543	0,814	1,750	0,685	1,977	0,562	2,220
16	1,106	1,371	0,982	1,539	0,857	1,728	0,734	1,935	0,615	2,157
17	1,133	1,381	1,015	1,536	0,897	1,710	0,779	1,900	0,664	2,104
18	1,158	1,391	1,046	1,535	0,933	1,696	0,820	1,872	0,710	2,060
19	1,180	1,401	1,074	1,536	0,967	1,685	0,859	1,849	0,752	2,023
20	1,201	1,411	1,100	1,537	0,998	1,676	0,894	1,828	0,792	1,991
21	1,222	1,420	1,125	1,538	1,026	1,669	0,927	1,812	0,829	1,964
22	1,239	1,429	1,147	1,541	1,053	1,664	0,958	1,797	0,863	1,940
23	1,257	1,437	1,168	1,543	1,078	1,660	0,986	1,785	0,895	1,920
24	1,273	1,446	1,188	1,546	1,101	1,656	1,013	1,775	0,925	1,902
25	1,288	1,454	1,206	1,550	1,123	1,654	1,038	1,767	0,953	1,886
26	1,302	1,461	1,224	1,553	1,143	1,652	1,062	1,759	0,979	1,873
27	1,316	1,469	1,240	1,556	1,162	1,651	1,084	1,753	1,004	1,861
28	1,328	1,476	1,255	1,560	1,181	1,650	1,104	1,747	1,028	1,850
29	1,341	1,483	1,270	1,563	1,198	1,650	1,124	1,743	1,050	1,841
30	1,352	1,489	1,284	1,567	1,214	1,650	1,143	1,739	1,071	1,833
31	1,363	1,496	1,297	1,570	1,229	1,650	1,160	1,735	1,090	1,825
32	1,373	1,502	1,309	1,574	1,244	1,650	1,177	1,732	1,109	1,819
33	1,383	1,508	1,321	1,577	1,258	1,651	1,193	1,730	1,127	1,813
34	1,393	1,514	1,333	1,580	1,271	1,652	1,028	1,728	1,144	1,808
35	1,402	1,519	1,343	1,584	1,283	1,653	1,222	1,726	1,160	1,803
36	1,411	1,525	1,354	1,587	1,295	1,654	1,236	1,724	1,175	1,799
37	1,419	1,530	1,364	1,590	1,307	1,655	1,249	1,723	1,190	1,795
38	1,427	1,535	1,373	1,594	1,318	1,656	1,261	1,722	1,204	1,792
39	1,435	1,540	1,382	1,597	1,328	1,658	1,273	1,722	1,218	1,789
40	1,442	1,544	1,391	1,600	1,338	1,659	1,285	1,721	1,230	1,786
45	1,475	1,566	1,430	1,615	1,383	1,666	1,336	1,720	1,287	1,776
50	1,503	1,585	1,462	1,628	1,421	1,674	1,378	1,721	1,335	1,771
55	1,528	1,601	1,490	1,641	1,452	1,681	1,414	1,724	1,374	1,768
60	1,549	1,616	1,514	1,652	1,480	1,689	1,444	1,727	1,408	1,767
65	1,567	1,629	1,536	1,662	1,503	1,696	1,471	1,731	1,438	1,767
70	1,583	1,641	1,554	1,672	1,525	1,703	1,494	1,735	1,464	1,768
75	1,598	1,652	1,571	1,680	1,543	1,709	1,515	1,739	1,487	1,770
80	1,611	1,662	1,586	1,688	1,560	1,715	1,534	1,743	1,507	1,772
85	1,624	1,671	1,600	1,696	1,575	1,721	1,550	1,747	1,525	1,774
90	1,635	1,679	1,612	1,703	1,589	1,726	1,566	1,751	1,542	1,776
95	1,645	1,687	1,623	1,709	1,602	1,732	1,579	1,755	1,557	1,778
100	1,654	1,694	1,634	1,715	1,613	1,736	1,592	1,758	1,571	1,780
150	1,720	1,746	1,706	1,760	1,693	1,774	1,679	1,788	1,665	1,802
200	1,758	1,778	1,748	1,789	1,738	1,799	1,728	1,810	1,718	1,820

* n – число наблюдений, m – количество независимых переменных

