

ДАЛЬНЕВОСТОЧНЫЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИНСТИТУТ МАТЕМАТИКИ И КОМПЬЮТЕРНЫХ НАУК
КАФЕДРА МАТЕМАТИЧЕСКИХ МЕТОДОВ В ЭКОНОМИКЕ

А. С. ВЕЛИЧКО

**ИЗУЧАЕМ ЭКОНОМЕТРИКУ.
НАЧАЛЬНЫЙ КУРС**

Учебное пособие

Владивосток
Издательство Дальневосточного университета
2007

УДК 330.43
ББК 65
В27

Величко А. С.

В27 Изучаем эконометрику. Начальный курс : учебное пособие / А. С. Величко. – Владивосток : Изд-во Дальневост. ун-та, 2007. – 72 с.

В пособии содержатся материалы к базовому односеместровому курсу эконометрики: рабочая учебная программа курса и его развернутое содержание, рейтинг-план, вопросы к теоретическому экзамену и вопросы для самоконтроля по теоретической части курса, практические задания и лабораторные работы с использованием специализированного программного пакета *Eviews* и рекомендации к их выполнению, лабораторные работы для самостоятельного выполнения, англо-русский глоссарий, список литературы.

Адресовано студентам, аспирантам и преподавателям экономических и экономико-математических специальностей вузов.

В $\frac{0604020105}{180(03) - 2007}$

УДК 330.43
ББК 65

© Величко А.С., 2007
© ИМКН ДВГУ, 2007

Предисловие

Предлагаемое учебное пособие содержит учебные материалы по базовому односеместровому курсу эконометрики, читаемому автором в Дальневосточном государственном университете для студентов специальности 06.18.00 «Математические методы в экономике» в течение трех последних лет.

Цель данного курса – получение студентами теоретических знаний и приобретение навыков эконометрического анализа и моделирования с использованием специализированного программного пакета *Eviews*. Предполагается, что студенты знакомы с курсами линейной алгебры, теории вероятностей и математической статистики, подготовлены к работе с электронными таблицами *Excel* в рамках учебной практики по учебному плану первого курса, имеют навыки программирования на языке Си, что поможет овладеть базовыми функциями пакета *Eviews*.

В результате изучения базового курса эконометрики студент должен знать основные методы и подходы эконометрического анализа для моделирования и прогнозирования экономических процессов и объектов на микро, макро и глобальном уровнях, самостоятельно осуществлять моделирование, прогнозирование, проводить аналитические расчеты в области экономической и управленческой деятельности с использованием эконометрических методов.

По учебному плану базовый курс эконометрики предполагает 34 часа теоретических и 34 часа практических и лабораторных занятий. Методика обучения предполагает чтение лекций для освоения теоретической части курса и проведение лабораторных занятий с использованием специализированного программного пакета *Eviews*. В течение семестра студенты выполняют 3 лабораторных работы и готовятся к экзамену по теоретической части курса.

Данное учебное пособие не содержит систематического изложения теоретического материала по курсу, который подробно излагается на лекционных занятиях. Для самостоятельной подготовки можно использовать рекомендованный в учебной программе курса список литературы.

Настоящее пособие включает в себя рабочую учебную программу курса и его развернутое содержание, рейтинг-план дисциплины для текущего контроля знаний, вопросы к теоретическому экзамену и вопросы для самоконтроля по теоретической части курса, практические задания и лабораторные работы с использованием специализированного программного пакета *Eviews* и рекомендации к их выполнению,

лабораторные работы для самостоятельного выполнения, англо-русский глоссарий, список литературы.

На веб-странице курса в Интернет http://www.iacp.dvo.ru/lab_11/oxho/Velichko/econometr.php размещены рабочая учебная программа курса, вопросы к теоретическому экзамену, часы и аудитории занятий и консультаций, ссылки на дополнительную литературу и ресурсы Интернет.

На структуру курса в целом и содержание данного пособия во многом оказало участие автора в семинарах, летних школах, практикумах, проходивших в разные годы с 1999 г. во Владивостоке и Санкт-Петербурге, организованных Российской экономической школой (Москва) и Европейским университетом в Санкт-Петербурге. Автор выражает искреннюю признательность профессорам РЭШ Катышеву П.К. и Пересецкому А.А. (РЭШ, Москва), профессору Харемзе В. (Лейстерский университет, Великобритания), профессору Макаровой С.Б. (Европейский университет, Санкт-Петербург), профессору Меленбергу Я. (Тилбургский университет, Нидерланды) за существенный опыт, полученный автором на семинарах, летних школах, практикумах по эконометрике.

Учебная программа и содержание курса

Модуль 1. Множественная линейная регрессия (12 часов, 6 недель)

Занятия 1-2. Введение в эконометрику.

Предмет эконометрики. Этапы эконометрического моделирования. Пространственные, временные, панельные статистические данные. Получение, преобразование и предварительная обработка данных. Аномальные наблюдения и погрешности наблюдений. Зависимые и независимые, эндогенные и экзогенные переменные.

Литература:

(Айвазян)¹ Разделы 14.1-14.3. Стр. 597-620.

(Магнус) Разделы 1.1-1.3. Стр. 26-31.

(Берндт) Глава 1.

Занятия 3-4. Постановка задачи множественной линейной регрессии.

Метод наименьших квадратов (МНК).

Задача множественного линейного регрессионного анализа. Основные предположения метода наименьших квадратов (МНК). Вывод формулы для МНК-оценки. Геометрическая интерпретация МНК. Экономическая интерпретация коэффициентов регрессии. Статистические свойства оценок МНК. Теорема Гаусса-Маркова. Частный случай парной регрессии и его геометрическая интерпретация.

Литература:

(Айвазян) Разделы 15.1-15.3. Стр. 621-650.

(Магнус) Разделы 2.1-2.5, 3.1-3.3. Стр. 32-51, 67-74.

(Берндт) Глава 2.

Занятия 5-6. Показатели качества регрессии. Проверка статистических гипотез и построение доверительных интервалов для параметров регрессии.

Стандартизованные коэффициенты регрессии, коэффициент детерминации. Статистика Стьюдента и критерий Фишера. Тест Вальда.

Литература:

(Айвазян) Раздел 15.3. Стр. 650-653.

(Магнус) Разделы 2.6, 3.4, 3.5, 10.6. Стр. 51-55, 74-88, 253-256.

(Берндт) Глава 2.

¹ Полную ссылку краткого обозначения учебников см. в списке литературы.

Модуль 2. Обобщение и модификация задачи линейной регрессии (8 часов, 4 недели)

Занятие 7. Полная и частичная мультиколлинеарность.

Полная и частичная мультиколлинеарность. Методы устранения мультиколлинеарности.

Литература:

(Айвазян) Разделы 15.4, 15.5. Стр. 653-671.

(Магнус) Разделы 4.1, 4.3, 4.4. Стр. 108-112, 124-135.

(Берндт) Глава 3.

Занятие 8-9. Нелинейные регрессионные модели. Искусственные (фиктивные) переменные.

Преобразование нелинейного уравнения регрессии к линейному виду. Коэффициент эластичности. Искусственные (фиктивные) переменные. Экономическая интерпретация коэффициентов при искусственных (фиктивных) переменных.

Литература:

(Айвазян) Разделы 15.9.4, 15.11, 15.12. Стр. 715-717, 735-766.

(Магнус) Раздел 4.2. Стр. 112-118.

(Берндт) Главы 3, 5.

Занятие 10. Линейные регрессионные модели с гетероскедастичными остатками.

Понятие гетеро- и гомоскедастичности. Тест Уайта. Оценивание в условиях гетероскедастичности. Состоятельное оценивание матрицы ковариации ошибок в форме Уайта и Навье-Веста. Обобщенный метод наименьших квадратов. Теорема Айткена.

Литература:

(Айвазян) Разделы 15.6, 15.7, 15.9.1. Стр. 672-690, 699-702.

(Магнус) Разделы 5.2, 5.3, 6.1. Стр. 154-163, 167-183.

(Берндт) Глава 4.

Модуль 3. Анализ временных данных в модели линейной регрессии (10 часов, 5 недель)

Занятия 11-12. Структурная изменчивость коэффициентов. Причинность и одновременность.

Тест Чоу на структурную изменчивость коэффициентов регрессии. Тест Гранжера на причинно-следственную связь между временными рядами.

Литература:

(Айвазян) Раздел 15.11.3. Стр. 675-676.

(Магнус) Разделы 3.5, 11.2. Стр. 85-86, 275-276.

(Берндт) Глава 8.

Занятие 13. Автокорреляция в регрессионном уравнении.

Понятие автокорреляции остатков регрессии. Тесты на автокорреляцию остатков (критерий Дарбина-Уотсона, LM-тест Бреуша-Годфри). Оценивание при наличии автокорреляции остатков (процедуры Кохрейна-Орката и Хилдрета-Лу).

Литература:

(Айвазян) Разделы 15.6, 15.8. Стр. 675-676, 690-699.

(Магнус) Раздел 6.2. Стр. 184-192.

(Берндт) Глава 6.

Занятие 14. Модели с распределенным лагом.

Регрессионная модель с распределенными лагами. Оценивание в моделях полиномиальных (Алмон) и геометрических (Койка) лагов.

Литература:

(Айвазян) Раздел 16.5.3. Стр. 872-888.

(Магнус) Разделы 11.1, 11.2. Стр. 264-268, 272-276.

(Берндт) Глава 8.

Занятие 15. Прогнозирование в регрессионных моделях.

Построение точечных и интервальных прогнозов в линейных регрессионных моделях. Прогнозирование в условиях автокорреляции остатков. Оценивание ошибки прогноза.

Литература:

(Айвазян) Разделы 15.9.2, 15.9.3, 16.6.3. Стр. 702-712, 888-906.

(Магнус) Разделы 7.1-7.3. Стр. 204-210.

Модуль 4. Системы одновременных уравнений

(4 часа, 2 недели)

Занятие 16. Система линейных одновременных уравнений и ее идентификация.

Общий вид системы одновременных уравнений, примеры моделей спроса и предложения. Структурная и приведенная форма системы. Идентификация систем, ранговое и порядковое условие идентифицируемости уравнений системы.

Литература:

(Айвазян) Разделы 17.1-17.3.1. Стр. 907-929.

(Магнус) Разделы 9.1, 9.2. Стр. 220-236.

Занятие 17. Методы оценивания параметров систем одновременных уравнений.

Косвенный метод, метод инструментальных переменных, метод внешне не связанных уравнений (SUR, multivariate regression), двушаговый и трехшаговый методы оценивания параметров систем одновременных уравнений.

Литература:

(Айвазян) Разделы 15.10, 17.3.2-17.3.4. Стр. 717-734, 929-942.

(Магнус) Раздел 5.1, 8.1, 8.2, 9.2. Стр. 148-153, 212-215, 237-241.

Рейтинг-план дисциплины

Эконометрика

(Название дисциплины согласно рабочему учебному плану)

Весовой коэффициент дисциплины в совокупной рейтинговой оценке, рассчитываемой по всем дисциплинам, равен 136.

1 Соотношение видов учебной деятельности студента, учитываемых в рейтинговой оценке по данной дисциплине

№	Виды учебной деятельности студентов	Весовые коэффициенты, %
1	<i>Контрольное мероприятие</i>	<i>100%</i>
	Сумма	100%

2 Максимально возможные баллы за виды контролируемой учебной деятельности студента (оценки «идеального студента» за одну единицу учебной работы)

№	Содержание вида контролируемой учебной деятельности	Единица измерения работы	Максимальное количество баллов за единицу выполненной работы
1.1	<i>Практическое задание</i>	<i>1 работа</i>	<i>10</i>
1.2	<i>Итоговое контрольное мероприятие</i>	<i>1 работа</i>	<i>30</i>

3 Календарный план контрольных мероприятий по дисциплине с указанием максимального количества баллов, потенциально доступных студенту

№	Дата	Название модуля	Форма контроля	Максимальное количество баллов
1	<i>7 неделя</i>	<i>Множественная линейная регрессия</i>	<i>Практическое задание</i>	<i>10</i>
2	<i>11 неделя</i>	<i>Обобщение и модификация задачи линейной регрессии</i>	<i>Практическое задание</i>	<i>10</i>
3	<i>14 неделя</i>	<i>Анализ временных данных в модели линейной регрессии</i>	<i>Практическое задание</i>	<i>10</i>
4	<i>16 неделя</i>	<i>Итоговое контрольное мероприятие</i>	<i>Итоговое контрольное мероприятие</i>	<i>30</i>

Вопросы к теоретическому экзамену

1. Множественная линейная регрессия: задача и основные предположения.
2. Метод наименьших квадратов для множественной линейной регрессии.
3. Геометрическая интерпретация метода наименьших квадратов.
4. Статистические свойства оценок параметров, теорема Гаусса-Маркова.
5. Использование t -статистики для проверки статистических гипотез о параметрах регрессии.
6. Использование коэффициента детерминации R^2 и F -критерия для проверки статистических гипотез о параметрах регрессии.
7. Тестирование гипотез общего линейного вида о параметрах регрессии.
8. Мультиколлинеарность (2 случая).
9. Искусственные (фиктивные) переменные.
10. Гетеро- и гомоскедастичность. Обобщенный метод наименьших квадратов и теорема Айткена.
11. Гетеро- и гомоскедастичность. Тест Уайта на гетероскедастичность.
12. Тест Чоу на структурную изменчивость.
13. Автокорреляция. Тесты на автокорреляцию остатков (критерий Дарбина-Уотсона, тест Бреуша-Годфри).
14. Оценивание при наличии автокорреляции остатков (процедуры Кохрейна-Орката и Хилдрета-Лу).
15. Модели с распределенным лагом. Оценивание в моделях полиномиальных и геометрических лагов.
16. Прогнозирование в регрессионных моделях.
17. Система линейных одновременных уравнений и ее идентификация.
18. Косвенный метод наименьших квадратов и метод инструментальных переменных оценки параметров систем одновременных уравнений.
19. Двухшаговый и трехшаговый методы оценки параметров систем одновременных уравнений.

Вопросы для самоконтроля к теоретической части курса

1. Перечислите этапы построения эконометрических моделей.
2. На основании каких исходных данных могут быть построены эконометрические модели?
3. Какова экономическая интерпретация коэффициентов регрессионного уравнения в линейной спецификации и в модели «в логарифмах»?
4. Какие гипотезы проверяются с помощью критерия Стьюдента?
5. Какие гипотезы проверяются с помощью критерия Дарбина-Уотсона?
6. Что показывают коэффициенты множественной корреляции и детерминации?
7. Какие гипотезы проверяются с помощью критерия Фишера?
8. Что такое «асимптотическая несмещенность», «асимптотическая состоятельность» и «эффективность»?
9. Каковы основные предположения метода наименьших квадратов (МНК)?
10. В чем суть МНК?
11. Приведите формулы расчета оценок коэффициентов линейной модели по МНК?
12. Какими свойствами обладают МНК-оценки классической линейной эконометрической модели?
13. Перечислите свойства вектора ошибок эконометрической модели.
14. Каким образом проверяется наличие автокорреляции ошибок модели?
15. Как оценивается дисперсия истинной ошибки модели?
16. Каковы последствия мультиколлинеарности факторов?
17. Как проверяется обратимость матрицы $X^T X$?
18. Каковы последствия неправильного выбора состава независимых переменных модели?
19. Каковы особенности оценивания параметров с учетом наложенных ограничений?
20. Опишите процедуру получения оценок параметров эконометрической модели с помощью метода максимального правдоподобия (ММП).
21. Какими свойствами обладают ММП-оценки параметров?
22. Как выглядит ковариационная матрица ошибок при наличии автокорреляционных связей в векторе ошибок?

23. Как выглядит ковариационная матрица ошибок модели при наличии гетероскедастичности ошибок?
24. Каковы последствия автокорреляции и гетероскедастичности ошибок?
25. В чем суть обобщенного МНК (ОМНК)?
26. Как определяется ковариационная матрица ОМНК-оценок параметров?
27. В чем суть взвешенного МНК?
28. В чем суть метода инструментальных переменных?
29. Какие проблемы возникают при построении моделей с лаговыми переменными?
30. Что представляет собой модель Койка?
31. Перечислите основные подходы к оценке коэффициентов эконометрической модели, содержащей лаговые зависимые переменные?
32. Охарактеризуйте причины изменчивости структуры модели и способы ее отображения в уравнении регрессии.
33. В чем состоят особенности оценки коэффициентов моделей с переменной структурой?
34. Как оценивается точность прогноза?
35. Что представляет собой «доверительный интервал прогноза»?
36. Охарактеризуйте методы оценки доверительного интервала прогноза в моделях с детерминированными и случайными параметрами.
37. Охарактеризуйте особенности прогнозирования на основе моделей авторегрессионных временных рядов.
38. Чем обусловлена смещенность оценок коэффициентов уравнений, полученных с использованием МНК?
39. Что представляют собой структурная и приведенная формы модели одновременных уравнений?
40. Что представляют собой порядковое и ранговое условия идентифицируемости уравнений структурной формы модели одновременных уравнений?
41. В чем состоит суть двухшагового и трехшагового МНК, используемых для оценки коэффициентов системы одновременных уравнений?

Лабораторные работы и рекомендации к их выполнению

В течение семестра студенты выполняют 3 лабораторные работы по модулям учебной программы. Лабораторные занятия в течение семестра в объеме 34 часов проходят с использованием специализированного программного пакета *Eviews* версии 3.1. Дальнейшее изложение ориентировано именно на эту версию пакета. Предполагается, что студенты подготовлены к работе с электронными таблицами *Excel* в рамках учебной практики по учебному плану первого курса, имеют навыки программирования на языке Си, что поможет овладеть базовыми функциями пакета *Eviews*.

Лабораторная работа 1. Множественная линейная регрессия.

Занятия 1-2. Основы статистической обработки данных в Eviews.

Задачи занятий: ознакомление с программой *Eviews* и инструментами *Excel* для статистической обработки данных, создание рабочего файла, импорт данных из *Excel* в *Eviews*, создание новых рядов данных, создание объекта группы переменных¹.

1. Основные понятия оконного интерфейса пакета Eviews.

- рабочий файл, рабочее окно (*workfile*),
- меню рабочего окна и пункт меню,
- основное меню *Eviews*,
- диапазон данных (*range*),
- поддиапазон данных (*sample*),
- количество наблюдений (*observations*),
- окно ввода команд (командная строка),
- объекты *Series* (ряд данных, выборка, переменная), *Group* (группа переменных), *Equation* (регрессия), *Graph* (график), *Table* (таблица), *Text* (текст),
- окно объекта,
- имя объекта,
- пиктограмма объекта,
- меню объекта и пункт меню,

¹ Далее понятия «переменная» и «ряд данных» будут использоваться как синонимы, поскольку ряд данных есть набор реализаций некоторой наблюдаемой физической, экономической и т.п. переменной.

- диалоговое окно.

2. Подготовка данных в Excel для импорта в Eviews.

Создайте новую рабочую книгу *Excel*. Убедитесь, что все импортируемые данные расположены на одном листе с названием без символов кириллицы или со стандартным названием (Лист1 для русифицированной версии *Excel*). Данные должны быть размещены в столбцах (или строках) без пропусков ячеек. Не допускайте также объединения ячеек импортируемого столбца (строки). В первой ячейке столбца (строки) укажите имя переменной (без символов кириллицы), которое понадобится для импорта данного столбца (строки) в *Eviews*. Желательно сразу указывать такие имена рядов данных, по которым можно будет легко идентифицировать импортируемую переменную в дальнейшем при работе с ней в *Eviews*. Установите формат импортируемых ячеек на «Общий», проверьте отсутствие нечисловых значений в ячейках, например, текста, значений со знаком процента «%». Сохраните рабочую книгу *Excel* в формате *Excel 5.0/95* с именем файла без символов кириллицы.

3. Создание рабочего файла Eviews.

Чтобы начать работу в пакете *Eviews*, после его запуска необходимо создать новый рабочий файл (workfile). Для этого выберите в основном меню *Eviews* пункт *File, New, Workfile*. В появившемся диалоговом окне *Workfile range* необходимо указать тип данных и ввести идентификатор начального (Start date) и последнего (End date) наблюдения. *Eviews* позволяет использовать следующие типы данных:

- годовые (Annual), формат идентификатора первого и последнего наблюдения имеет вид 'год', например, 2001 (2001 год),
- полугодовые (Semi-annual), формат идентификатора первого и последнего наблюдения имеет вид 'год:номер_полугодия', например, 2001:2 (второе полугодие 2001 года),
- поквартальные (Quarterly), формат идентификатора первого и последнего наблюдения имеет вид 'год:номер_квартала', например, 2001:3 (третий квартал 2001 года),
- ежемесячные (Monthly), формат идентификатора первого и последнего наблюдения имеет вид 'год:номер_месяца', например, 2001:11 (ноябрь 2001 года),
- недельные (Weekly), формат идентификатора первого и

последнего наблюдения имеет вид 'год:номер_недели', например, 2001:11 (ноябрь 2001 года),

- подневные (Daily), с пятидневной, то есть без суббот и воскресений, недель (5 day weeks) или семидневной недель (7 day weeks), формат идентификатора первого и последнего наблюдений имеет вид 'месяц:день:год', например, 9:13:1998 (13 сентября 1998 года),
- недатированные или нерегулярные (Undated or irregular), этот тип предназначен для работы с пространственными данными. Необходимо ввести только номер последнего наблюдения, равный количеству наблюдений. Диапазон данных будет по умолчанию начинаться с наблюдения с номером 1.

По завершению ввода диапазона данных после нажатия кнопки «ОК» *Eviews* создаст рабочий файл без имени и откроет рабочее окно (Workfile). Рабочее окно всегда *Eviews* содержит служебный вектор коэффициентов C , обозначенный пиктограммой в виде буквы β и служебный ряд данных *resid*. Данные имена переменных являются зарезервированными и не могут быть использованы для сохранения других рядов данных.

4. Импорт данных из Excel в Eviews.

Убедитесь, что файл *Excel*, содержащий импортируемые данные, закрыт или открыт с опцией «только для чтения».

Для начала импорта данных в основном меню *Eviews* (или в меню рабочего окна) выберите пункт *Procs, Import, Read Text-Lotus-Excel*, затем выберите в диалоговом окне *Open* сохраненный вами файл *Excel* с импортируемыми данными. Далее в появившемся диалоговом окне *Excel Spreadsheet Import* необходимо:

- в разделе *Order of data* (порядок данных) выбрать пункт *By Observation - Series in columns*, если импортируемые данные расположены в столбцах листа *Excel*, или выбрать пункт *By Series - Series in rows*, если импортируемые данные расположены в строках листа *Excel*,
- в поле *Upper-left data cell* (адрес левой верхней ячейки диапазона данных) указать имя ячейки, с которой начинаются импортируемые данные (то есть адрес левой верхней ячейки диапазона данных) на листе *Excel*,

- в поле `Excel 5+ sheet name` указать имя листа *Excel* с импортируемыми данными, если это имя является стандартным (Лист 1 для русифицированной версии), то поле можно оставить пустым,
- в окне `Names for Series or Number of Series if names in file` достаточно указать число импортируемых рядов данных, если вы не забыли указать имена каждого диапазона данных на листе *Excel* в первой строке (столбце), иначе необходимо ввести имена переменных (без символов кириллицы), разделяя их пробелом, количество которых равно количеству импортируемых рядов данных,
- в окне `Sample to import` указан диапазон данных, который был введен при создании рабочего файла *Eviews*, его изменять не нужно, если вы еще не изменяли данные, например, добавив новые строки (столбцы) на листе *Excel*,
- можно указать, изменять ли после импорта поддиапазон данных (`Reset sample to`), устанавливая его на поддиапазон импортируемых данных (`Current sample`), диапазон данных, уже установленный в рабочем файле (`Workfile range`).

После нажатия кнопки «ОК», если импорт данных прошел успешно, ряды данных будут сохранены в виде объектов `Series` и обозначены в рабочем окне *Eviews* в виде пиктограмм с галочкой, справа от пиктограммы объекта `Series` указывается имя переменной.

Проверьте правильность импорта данных. Откройте двойным щелчком левой кнопки мыши по изображению объекта типа `Series` в рабочем окне *Eviews*, чтобы отобразить таблицу значений ряда данных. Если объект `Series` уже открыт, вернуться к отображению значений ряда данных можно, выбрав пункт `View SpreadSheet` в меню объекта `Series`. Проверьте правильность импорта значений рядов данных, количество значений в импортированных рядах данных в созданных объектах `Series`.

Сохраните рабочий файл *Eviews* на диск, выбрав в основном меню пункт `File, Save`. Рабочие файлы *Eviews* имеют расширение `wf1`.

5. Создание новых рядов данных

Создать новый ряд можно путем вычислений над уже имеющимися рядами. Пакет *Eviews* содержит внутренний язык программирования, работа с которым осуществляется в окне ввода команд. Чтобы создать

новую переменную, необходимо написать команду *genr* в окне ввода команд, затем через пробел указать имя создаваемой переменной, затем знак равенства и, наконец, произвольное леводопустимое выражение. Например, чтобы создать (или перезаписать) ряд данных x , со значениями, равными 1, для всех наблюдений, необходимо ввести в окне ввода команд *genr x=1* и нажать Enter.

Поскольку работа с командой *genr* часто необходима, можно выбрать в рабочем окне *Eviews* пункт меню *Genr*, затем в диалоговом окне *Generate by Series* указать необходимое выражение.

Создадим переменную y , значения которой будут равны 1 для тех наблюдений, у которых значение некоторой уже имеющейся переменной x больше какого-то числа, например 15, а для всех остальных наблюдений значения переменной y будут равны 0. Для этого необходимо ввести команду *genr y=(x>15)*.

Заметьте, что, обрабатывая выражение, *Eviews* обращается последовательно ко всем наблюдениям. Фактически, выполняя данное присваивание, *Eviews* иницирует цикл по всем наблюдениям. Зачастую необходимо сгенерировать ряд, значения которого представляют собой прогрессию, заданную рекуррентным соотношением. *Eviews* легко справляется с этой задачей. Выражение $x(2)$ означает ряд данных для переменной x , взятый со сдвигом (или лагом) на два наблюдения вперед, $x(-1)$ – со сдвигом назад на одно наблюдение. Введем последовательно следующие две команды в окне ввода команд:

$$\begin{aligned} & \textit{genr} \ x=1 \\ & \textit{genr} \ x(1)=x+1 \end{aligned}$$

Вводим Enter после каждого выражения. Выполняя первую команду, мы создаем ряд наблюдений x со значениями для всех наблюдений, равными 1. При выполнении второй команды, применяя рекуррентное выражение последовательно ко всем наблюдениям, *Eviews* присвоит переменной x значения 2,3,... и так далее до последнего наблюдения в диапазоне данных.

6. Создание объекта группы рядов данных (переменных).

Для создания объекта группы переменных (*Group*) необходимо последовательно выбрать левой кнопкой мыши нужные объекты *Series* в рабочем окне *Eviews*, удерживая клавишу *Ctrl*. Затем, отпустив клавишу *Ctrl* и нажав правую кнопку мыши, необходимо выбрать в появившемся меню *Open as Group*. Чтобы сохранить вновь созданный

объект Group необходимо выбрать в меню объекта Group пункт Name и ввести допустимое имя без символов кириллицы. После нажатия кнопки «ОК» объект Group будет сохранен с введенным вами именем и отображен в рабочем окне пиктограммой в виде буквы G.

Занятия 3-4. Основные статистические показатели и визуализация данных.

Задачи занятий: проанализировать наличие выбросов, получить основные статистические показатели рядов данных, гистограмму, корреляционную матрицу для независимых переменных, построить графики рядов данных.

1. Анализ и исключение «выбросов» (outliers).

Выбросы (outliers) – значения наблюдений ряда, которые значительно отклоняются от среднего. Исключить выбросы при расчете статистических показателей, оценивании, построении графиков можно, сделав двойной щелчок по поддиапазону данных (sample) в рабочем окне *Eviews* (либо выбрав в рабочем окне *Eviews* пункт меню *Sample* или *Proc, Sample*). Затем в появившемся диалоговом окне *Sample* в поле *IF condition (optional)* можно ввести условие отбора наблюдений, например, $x \leq 1000$, что исключит из анализа наблюдения, для которых значения ряда данных x строго больше 1000.

Те же действия необходимо проделать для того, чтобы изменить сам поддиапазон данных (sample). В диалоговом окне *Sample* в поле *Sample range pairs* можно указать другой диапазон, границы которого не могут выходить за пределы основного диапазона данных (range).

Увеличить или уменьшить основной диапазон данных (range) можно, сделав двойной щелчок по диапазону данных в рабочем окне *Eviews* (либо выбрав в рабочем окне *Eviews* пункт меню *Proc, Change workfile range*). В случае расширения границ диапазона данных *Eviews* добавит ко всем рядам данным необходимое количество наблюдений. Значения рядов данных рабочего окна *Eviews* для вновь добавленных наблюдений будут не определены и помечены в таблице значений объекта *Series* как *NA (not available)* до тех пор, пока вы их не зададите явно или не присвоите им значения в результате вычислений, например, с помощью команды *genr*. В случае сокращения диапазона данных произойдет потеря значений для наблюдений во ВСЕХ рядах данных рабочего файла.

2. Работа с надстройкой «Анализ данных» в Excel.

Подключите надстройку «Анализ данных» (Analysis ToolPak), выбрав пункт меню *Сервис, Надстройки*. Чтобы вызвать надстройку, выберите пункт меню *Сервис, Анализ данных* (Data Analysis). Доступны следующие инструменты надстройки:

- корреляция (correlation), позволяет создать корреляционную матрицу рядов данных,
- ковариация (covariance), позволяет создать ковариационную матрицу рядов данных,
- описательная статистика (descriptive statistics), позволяет получить основные статистические характеристики рядов данных (среднее значение, среднее квадратическое отклонение, минимальное и максимальное значение и т.п.),
- гистограмма (histogram), позволяет получить таблицу частот для последующего построения гистограммы распределения ряда данных,
- регрессия (regression), позволяет оценить методом наименьших квадратов линейное уравнение множественной регрессии и получить основные статистические показатели качества регрессии.

3. Основные статистические показатели рядов данных.

Необходимо открыть объект Series двойным щелчком мышью на пиктограмме нужного ряда данных (пиктограмма с галочкой) в рабочем окне *Eviews*.

Описательную статистику ряда и гистограмму можно создать, выбрав в меню объекта Series пункт *View Descriptive Statistics, Histogram and Stats* (второй способ: выбрать в основном меню *Eviews* пункт *Quick, Series Statistics, Descriptive Statistics, Histogram and Stats*).

Описательная статистика ряда содержит следующие элементы:

- Mean – среднее значение ряда данных,
- Median – медиана ряда данных,
- Maximum – максимальное значение ряда данных,
- Minimum – минимальное значение ряда данных,
- Std. Dev. – среднее квадратическое отклонение ряда данных,
- Skewness – коэффициент асимметрии ряда данных,
- Kurtosis – эксцесс ряда данных.

Дополнительные возможности в меню объекта Series предоставляет пункт View Descriptive Statistics, Stats by Classification. Можно получить расширенный перечень статистических показателей, сделать расчеты по нескольким подвыборкам (bins) и т.д.

Описательную статистику удобно создавать в объекте Group сразу для нескольких рядов данных. Для этого в меню объекта Group необходимо выбрать пункт View Descriptive Stats, Common Sample.

Построить ковариационную и корреляционную матрицы для нескольких рядов наблюдений можно, выбрав в меню объекта Group пункт View Covariances, Common Sample и View Correlations, Common Sample.

Заметьте, что при расчете статистических показателей результаты не сохраняются в отдельном объекте, создания нового окна с результатами не происходит, и это не совсем удобно. Для сохранения результатов расчетов при работе с объектами Series и Group можно выбрать в меню этих объектов пункт Freeze, в результате будет создан новый объект Table, куда будет скопирован текущий текст результатов расчетов. Необходимо сохранить вновь созданный объект Table, выбрав в меню объекта пункт Name и введя допустимое имя без символов кириллицы. После нажатия кнопки «ОК» объект Table будет сохранен с введенным вами именем и отображен в рабочем окне *Eviews* в виде пиктограммы с изображением таблицы и надписью «Table».

Однако, после выбора пункта Freeze в меню объекта Series с созданной гистограммой и описательной статистикой (Histogram and Stats) создается объект Graph, а не Table. Поэтому, чтобы сохранить описательную статистику в текстовом виде, необходимо ее создать в объекте Group для нескольких рядов данных (выбирая в меню объекта Group пункт View Descriptive Stats, Common Sample), а затем воспользоваться пунктом Freeze в меню объекта Group.

4. Построение графиков в *Eviews*.

Для построения графика ряда наблюдений в *Eviews* необходимо открыть объект Series двойным щелчком мыши на пиктограмме нужного ряда данных (пиктограмма с галочкой) в основном рабочем окне *Eviews*.

В меню View объекта Series доступны следующие пункты:

- Spreadsheet – показать значения ряда данных в виде таблицы значений, которые можно редактировать, копировать,
- Graph – построить различные виды графиков: Line, Area, Bar, Spike, Seasonal.

Если создан объект Group с несколькими рядами данных, то тогда, выбрав в меню этого объекта пункт View Graph и далее необходимый вид графика, можно получить графики рядов данных, совмещенные на одном рисунке. Выбирая в меню этого объекта пункт View Multiple Graph и далее необходимый вид графика, можно получить отдельные графики для каждого ряда данных.

Графики можно сохранить в специальном объекте Graph, поскольку построение графиков происходит в окне объекта Series или Group, и создания нового окна не происходит. Для этого необходимо в меню объекта Series или Group выбрать пункт Freeze, в результате будет создан новый объект Graph. Для его сохранения необходимо в меню этого объекта выбрать пункт Name, ввести допустимое имя без символов кириллицы и нажать «ОК». В результате объект Graph будет отображен в рабочем окне *Eviews* в виде пиктограммы с изображением графика.

Занятие 5. Метод наименьших квадратов.

Задачи занятия: составить уравнение регрессии, основываясь на известных моделях экономической теории, рассчитать параметры линейного уравнения регрессии, воспользовавшись методом наименьших квадратов, дать экономическую интерпретацию коэффициентов регрессии.

1. Оценивание параметров уравнения регрессии.

Определите зависимую переменную и набор независимых переменных (факторов) и составьте какое-либо эконометрическое уравнение на основе известных моделей экономической теории. Например, можно анализировать зависимость доли расходов на потребительские товары в регионе или реальный уровень потребления в зависимости от показателей, характеризующих реальный доход, богатство и реальные денежные остатки в данном регионе.

Согласно экономической теории доля расходов на потребительские товары убывает с ростом реального дохода и возрастает с ростом реальных денежных остатков и богатства. Поэтому ожидаемый знак коэффициента при независимых переменных, характеризующих

реальный доход в регионе, – отрицательный, а при показателях богатства и реальных денежных остатков – положительный.

Для создания регрессионного уравнения и его оценки служит объект *Equation*. Необходимо последовательно выделить левой кнопкой мыши сначала зависимую переменную, а затем последовательно независимые переменные, удерживая клавишу *Ctrl*. Отпустите клавишу *Ctrl*. Нажав правую кнопку мыши и выбрав в появившемся меню *Open as Equation*, появится диалоговое окно *Equation Specification*. В окне будет указана спецификация уравнения для оценивания в *Eviews*. Сначала указано имя зависимой переменной, затем через пробел перечислены имена независимых переменных, и в конце указана служебная переменная *C*, в которой будут сохранены оценки коэффициентов регрессии.

Создать объект *Equation* можно другим способом, выбрав в основном меню *Eviews* пункт *Quick, Estimate Equation*. Затем в появившемся диалоговом окне *Equation Specification* необходимо указать спецификацию уравнения, указав сначала имя зависимой переменной, затем через пробел имена независимых переменных, и в конце указав служебную переменную *C*, в которой будут сохранены оценки коэффициентов регрессии.

Провести оценивание регрессионного уравнения можно в окне ввода команд с помощью команды *ls*, указав имя зависимой переменной, затем через пробел имена независимых переменных, и в конце указав служебную переменную *C*.

Если зависимая и независимые переменные уже сохранены в объекте группа (*Group*), то, открыв этот объект (сделав двойной щелчок по его пиктограмме в рабочем окне *Eviews*), можно создать уравнение регрессии, выбрав пункт *Proc, Make Equation* в меню объекта *Group*. Отредактируйте спецификацию уравнения регрессии по описанным выше правилам, если это необходимо.

В поле *Method* необходимо выбрать метод оценки коэффициентов уравнения. По умолчанию выбран метод наименьших квадратов *LS – Least Squares (NLS and ARMA)*. В поле *Sample* можно указать произвольный поддиапазон данных, по которому будет производиться оценивание.

После нажатия кнопки «ОК» будет создан объект *Equation*, в котором показаны результаты оценивания. Для сохранения нового объекта необходимо выбрать пункт *Name* в меню объекта *Equation* и

ввести допустимое имя без символов кириллицы. После нажатия кнопки «ОК» объект Equation будет сохранен с введенным вами именем и отображен в основном рабочем окне в виде пиктограммы со знаком равенства.

Заметим, что оценки метода наименьших квадратов существуют не всегда. В данном случае объекта Equation создано не будет, и появится предупреждение об ошибке Near singular matrix (матрица вырождена), что говорит о несуществовании МНК-оценок. Подробнее этот случай мы рассмотрим на занятии 6.

Выбрав в меню объекта Equation пункт View Representations, можно получить в текстовом виде спецификацию регрессионного уравнения, оцениваемое уравнение регрессии и, что особенно важно, уравнение регрессионной прямой (регрессионное уравнение с подставленными коэффициентами). Выделив текст, его можно копировать в другое приложение, например, текстовый редактор.

Для сохранения результатов расчетов можно выбрать в меню объекта Equation пункт Freeze, в результате будет создан новый объект Table, куда будет скопирован текущий текст результатов расчетов. Необходимо сохранить вновь созданный объект Table, выбрав в меню объекта пункт Name и введя допустимое имя без символов кириллицы. После нажатия кнопки «ОК» объект Table будет сохранен с введенным вами именем и отображен в рабочем окне *Eviews* в виде пиктограммы с изображением таблицы и надписью «Table».

2. Экономическая интерпретация оценок коэффициентов регрессии.

Результаты оценивания регрессионного уравнения можно отобразить, выбрав в меню объекта Equation пункт View Estimation Output или пункт Stats.

Обозначим за CS ряд данных для реального потребления, за GDP – реальный ВВП Великобритании, используя поквартальные данные за период с 1955 по 2001 г.

В таблице 1 показан вывод результатов, отображаемый в объекте Equation после проведения оценивания коэффициентов регрессионного уравнения $CS = C(1) \times GDP + C(2)$ методом наименьших квадратов. Коэффициент $C(1)$ есть наклон регрессионной прямой, $C(2)$ – константа данной парной регрессии.

Уравнение регрессионной прямой (регрессионное уравнение с подставленными коэффициентами) имеет вид

$$CS = 0.7045856401 \times GDP - 10544.66422,$$

для его отображения в меню объекта Equation выберите пункт View Representations.

Таблица 1. Результаты оценивания уравнения регрессии в *Eviews*.

Dependent Variable: CS

Method: Least Squares

Sample: 1955:1 2001:4

Included observations: 188

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GDP	0.704586	0.005208	135.2764	0.0000
C	-10544.66	701.3370	-15.03509	0.0000
R-squared	0.989938	Mean dependent var		79941.78
Adjusted R-squared	0.989884	S.D. dependent var		28740.23
S.E. of regression	2890.634	Akaike info criterion		18.78692
Sum squared resid	1.55E+09	Schwarz criterion		18.82135
Log likelihood	-1763.971	F-statistic		18299.69
Durbin-Watson stat	0.061335	Prob(F-statistic)		0.000000

Коэффициент при переменной GDP, равный (примерно) 0.7, означает, что по имеющимся наблюдениям при росте (снижении) реального ВВП на единицу реальное потребление увеличивается (уменьшается) в среднем на 0.7. Данный коэффициент в макроэкономике называется предельной склонностью к потреблению, его оценка находится в пределах от нуля до единицы, что согласуется с экономической теорией и полученное значение имеет экономический смысл.

Оценка свободного коэффициента (константы регрессии) равна - 10544.66. В экономической теории данная величина интерпретируется как автономное потребление за вычетом предельной склонности к потреблению, умноженной на величину прямых налогов.

Насколько надежны полученные результаты с точки зрения тестирования гипотез, мы рассмотрим на занятии 5.

3. Графическое представление результатов регрессии.

Для парной регрессии можно графически представить наблюдения в виде точек двумерной декартовой системы координат, поставив в соответствие каждому наблюдению пару значений. Абсциссе точки для каждого наблюдения соответствует значение независимой переменной, а ординате – значение зависимой переменной. Данную «карту наблюдений» можно совместить с регрессионной прямой с помощью *Eviews*, что позволит наглядно отразить качество подгонки кривой.

Для этого необходимо в меню объекта Group, содержащего зависимую и все независимые переменные, выбрать пункт View Graph, Scatter, Scatter with Regression. Если набор переменных не был предварительно сохранен в объекте Group, то необходимо в меню объекта Equation выбрать пункт Procs, Make Regressor Group, после чего будет создан объект Group с зависимой и независимой переменной. Далее полученный график можно скопировать в объект Graph и сохранить в рабочем файле *Eviews*.

Если вы изменяли данные, изменялся поддиапазон наблюдений, повторно выполнить оценивание (или изменить спецификацию уравнения) можно, выбрав в меню сохраненного в рабочем окне объекта Equation пункт Estimate или Procs, Specify/Estimate.

Важно проанализировать статистические показатели остатков регрессии, в частности, получить описательную статистику и построить гистограмму. После оценивания по методу наименьших квадратов в рабочем окне обновляются значения служебного ряда данных resid, который содержит остатки оцененного вами регрессионного уравнения. Выбрав пункт Resid в меню объекта Equation, получим совмещенный график остатков (Resid), фактических (Actual) и прогнозных значений (Fitted) зависимой переменной. Построенный график остатков можно скопировать в новый объект Graph выбором пункта меню Freeze.

Другие типы графиков остатков можно построить, выбрав в меню объекта Equation пункт View Actual, Fitted, Residual и далее один из предлагаемых типов графиков.

Для построения гистограммы и описательной статистики для вектора остатков регрессии можно использовать стандартный подход, который рассматривался ранее на занятии 2 в п.3 «Основные статистические показатели рядов данных». Для этого после оценивания регрессии необходимо открыть служебный ряд данных resid, который является объектом типа Series, дважды щелкнув по его изображению в рабочем окне *Eviews* мышью. Затем в меню объекта Series выбрать пункт View Descriptive Statistics, Histogram and Stats (второй способ: выбрать в основном меню *Eviews* пункт Quick, Series Statistics, Descriptive Statistics, Histogram and Stats).

Выбрав в меню объекта Series пункт View Descriptive Statistics, Stats by Classification, можно получить расширенный перечень статистических показателей, сделать расчеты по

нескольким подвыборкам (bins) и т.д. Существует и другой, иногда более удобный, способ. В меню объекта Equation достаточно выбрать пункт View Residual Tests, Histogram - Normality Test.

Занятие 6. Показатели качества регрессии. Тестирование статистических гипотез.

Задачи занятия: рассчитать остаточную сумму квадратов регрессионного уравнения, коэффициент детерминации R^2 , оценить статистическую значимость параметров регрессионного уравнения с помощью t -статистики Стьюдента, проверить гипотезу о значимости уравнения в целом с помощью статистики Фишера F , построить доверительные интервалы для коэффициентов регрессионного уравнения.

1. Показатели качества регрессии. Тестирование статистических гипотез о значимости отдельных коэффициентов регрессии.

В окне результатов оценивания регрессии объекта Equation рассчитываются следующие показатели качества регрессионного уравнения:

- остаточная сумма квадратов регрессии (Sum squared resid),
- коэффициент детерминации R^2 (R-squared),
- скорректированный (на число регрессоров) коэффициент детерминации R_{adj}^2 (Adjusted R-squared),
- оценка среднеквадратического отклонения ошибки регрессии σ (S.E. of regression),
- среднеквадратические отклонения для коэффициентов регрессии $se(\beta_i)$ (Std. Error),
- фактические значения t -статистик Стьюдента для каждого коэффициента регрессии (t-Statistic),
- p -значения (фактические вероятности принятия нулевой гипотезы) для каждого коэффициента регрессии (Prob),
- значение статистики Фишера F (F-statistic),
- p -значения (фактические вероятности принятия нулевой гипотезы) для статистики Фишера F (Prob(F-statistic)).

Величина коэффициента детерминации R^2 имеет смысл только в том случае, когда константа включена в состав регрессоров, и может быть интерпретирована как доля вариации зависимой переменной, объясненная вариацией независимых переменных (факторов) регрессионного уравнения.

Поскольку проверка гипотез основана на предположении о нормальности распределения ошибок регрессионного уравнения, необходимо проверить его выполнение для полученного вектора остатков регрессии.

Для проверки нулевой гипотезы H_0 о равенстве нулю некоторого коэффициента регрессионного уравнения ($H_0: \beta_i=0$) необходимо сравнить фактическое значение статистики, найденное по формуле $\left| \frac{\hat{\beta}_i}{se(\beta_i)} \right|$, которое указывается в колонке *t-Statistic* в окне результатов оценивания регрессии объекта *Equation*, с критическим значением *t*-статистики Стьюдента для выбранного уровня значимости α , то есть со значением двусторонней $(1-\alpha)$ квантили *t*-статистики Стьюдента с $n-k$ степенями свободы.

Величина α характеризует допустимый уровень вероятности ошибиться, отвергнув нулевую гипотезу, когда она верна. В теории проверки статистических гипотез величину α называют ошибкой первого рода.

Если фактическое значение *t*-статистики Стьюдента больше критического значения статистики, то нулевая гипотеза отвергается для данного уровня значимости α , иначе нулевая гипотеза не может быть отвергнута для данного уровня значимости α .

В случае отвержения нулевой гипотезы для уровня значимости α говорят, что коэффициент β_i регрессионного уравнения *значим* на уровне значимости α (или, говорят, что оценка коэффициента β_i значимо отличается от нуля), и соответствующий ему регрессор объясняет вариацию зависимой переменной. В противном случае говорят, что коэффициент *незначим* на уровне значимости α .

Второй способ проверки гипотезы – сравнить *p*-значение (фактическую вероятность принятия нулевой гипотезы данного коэффициента регрессии) с выбранным уровнем значимости α . Если выполняется условие $p < \alpha$, то нулевая гипотеза отвергается на уровне значимости α , иначе нулевая гипотеза не может быть отвергнута для данного уровня значимости α .

Eviews не возвращает критического значения *t*-статистики Стьюдента в окне результатов оценивания регрессии, но соответствующая двусторонняя квантиль может быть найдена с использованием встроенной функции *@qtdist*. Введем команду

show @qtdist(v,p)

в окне ввода команд, $v=1-\alpha$ для односторонней квантили, $v=1-\alpha/2$ для двусторонней квантили, $p=n-k$ – количество степеней свободы. Команда `show` показывает на экран значение функции `@qtdist`.

Например, в рассмотренном выше выводе результатов оценивания (занятие 5, п.2, таблица 1), количество наблюдений n равно 188, количество регрессоров k равно 2, двусторонняя 0.95 квантиль t -статистики Стьюдента для 186 степеней свободы равна `@qtdist(0.975, 186)=1.973`. Таким образом, критическое значение t -статистики Стьюдента для 186 степеней свободы и выбранного уровня значимости 0.05 (5%) для проверки тестируемой нулевой гипотезы равно 1.973.

В рассматриваемом примере фактическое значение t -статистики Стьюдента для коэффициента $C(1)$ по модулю равно (примерно) 135.3, что больше рассчитанного критического значения t -статистики Стьюдента 1.973, поэтому нулевая гипотеза отвергается для выбранного уровня значимости 0.05 (5%). Другими словами, нулевая гипотеза отвергается с вероятностью ошибиться (то есть отвергнуть нулевую гипотезу, когда она верна), равной 5%.

В этом случае говорят, что коэффициент регрессионного уравнения $C(1)$ значим на уровне значимости 0.05 (5%) (или, говорят, что данный коэффициент значимо отличается от нуля), и соответствующий ему регрессор (то есть, ВВП) объясняет вариацию зависимой переменной CS (потребление).

Воспользовавшись вторым методом тестирования гипотез для рассматриваемого примера, получим, что p -значения для обоих коэффициентов (`Prob`) близки к нулю, поэтому на любом достаточно малом уровне значимости α можно отвергнуть тестируемую нулевую гипотезу.

Для проверки нулевой гипотезы H_0 о равенстве коэффициента β_i регрессионного уравнения некоторой наперед заданной константе b ($H_0: \beta_i=b$) необходимо вычислить фактическое значение статистики по формуле $\left| \frac{\hat{\beta}_i - b}{se(\hat{\beta}_i)} \right|$ и сравнить его с критическим значением t -статистики Стьюдента для выбранного уровня значимости α , то есть с *положительным* значением $(1-\alpha)$ *двусторонней* квантили t -статистики Стьюдента с $n-k$ степенями свободы.

Если фактическое значение t -статистики Стьюдента *больше* критического значения статистики, то нулевая гипотеза отвергается для

данного уровня значимости α , иначе нулевая гипотеза не может быть отвергнута для данного уровня значимости α .

Для проверки нулевой гипотезы для некоторого коэффициента регрессионного уравнения вида $H_0: \beta_i \leq b$, где b – некоторая наперед заданная константа, необходимо вычислить фактическое значение статистики по формуле $\frac{\hat{\beta}_i - b}{se(\beta_i)}$ и сравнить его с критическим значением t -статистики Стьюдента для выбранного уровня значимости α , в данном случае – с *положительным* значением $(1-\alpha)$ *односторонней* квантили t -статистики Стьюдента с $n-k$ степенями свободы.

Если фактическое значение t -статистики Стьюдента *больше* критического значения статистики, то нулевая гипотеза отвергается для данного уровня значимости α , иначе нулевая гипотеза не может быть отвергнута для данного уровня значимости α .

Для проверки нулевой гипотезы для некоторого коэффициента регрессионного уравнения вида $H_0: \beta_i \geq b$, где b – некоторая наперед заданная константа, необходимо вычислить фактическое значение статистики по формуле $\frac{\hat{\beta}_i - b}{se(\beta_i)}$ и сравнить его с критическим значением t -статистики Стьюдента для выбранного уровня значимости α , в данном случае – с *отрицательным* значением $(1-\alpha)$ *односторонней* квантили t -статистики Стьюдента с $n-k$ степенями свободы.

Если фактическое значение t -статистики Стьюдента *меньше* критического значения статистики, то нулевая гипотеза отвергается для данного уровня значимости α , иначе нулевая гипотеза не может быть отвергнута для данного уровня значимости α .

2. Построение доверительного интервала для коэффициента регрессии.

Доверительный интервал для коэффициента β_i для уровня значимости α (или $(1-\alpha)$ доверительный интервал) представляет собой интервал $[\hat{\beta}_i - t_\alpha \cdot se(\beta_i), \hat{\beta}_i + t_\alpha \cdot se(\beta_i)]$, где $\hat{\beta}_i$ – оценка коэффициента β_i по методу наименьших квадратов, t_α – двусторонняя $(1-\alpha)$ квантиль распределения Стьюдента с $n-k$ степенями свободы, n – количество наблюдений (included observations), k – количество регрессоров, включая константу.

В рассмотренном выше примере п.2 занятия 4 получаем для коэффициента β_1 95% доверительный интервал (для 5% уровня значимости) вида $[0.7 - 1.973 \times 0.0052, 0.7 + 1.973 \times 0.0052] = [0.69, 0.71]$.

3. Проверка статистической гипотезы о значимости уравнения в целом.

Для проверки гипотезы о значимости уравнения в целом необходимо воспользоваться статистикой Фишера F . В этом случае нулевая гипотеза имеет вид $H_0: \{\beta_1=\beta_2=\dots=0\}$, то есть тестируется одновременное равенство нулю всех коэффициентов регрессионного уравнения кроме константы регрессии β_0 . Необходимо сравнить фактическое значение статистики Фишера F , возвращаемое *Eviews* (*F-statistic*), и сравнить его с критическим значением статистики Фишера F для выбранного уровня значимости α , то есть со значением $(1-\alpha)$ квантили статистики Фишера с $(k-1, n-k)$ степенями свободы.

Если фактическое значение статистики Фишера F больше критического значения, то нулевая гипотеза отвергается для данного уровня значимости α , иначе нулевая гипотеза не может быть отвергнута для данного уровня значимости α .

В случае отвержения нулевой гипотезы для уровня значимости α говорят, что регрессионное уравнение значимо в целом на уровне значимости α и вариация независимых переменных объясняет вариацию зависимой переменной в регрессионном уравнении. В противном случае говорят, что уравнение в целом незначимо на уровне значимости α и включенные в регрессию факторы не улучшают прогноз для зависимой переменной по сравнению с ее средним значением.

Eviews не возвращает критического значения статистики Фишера F в окне результатов оценивания регрессии, но соответствующая квантиль может быть найдена с использованием встроенной функции *@qfdist*. Введем в окне ввода команд строку `show @qfdist(v,p1,p2)`, где $v=1-\alpha$, $p_1=k-1$, $p_2=n-k$. Команда `show` показывает на экран значение функции *@qfdist*.

Например, в рассмотренном выше примере п.2 занятия 4 в выводе результатов оценивания количество наблюдений n равно 188, количество регрессоров k равно 2, 0.95 квантиль статистики Фишера для 186 степеней свободы равна *@qfdist(0.95,1,186)=253.629*. Критическое значение статистики Фишера для 186 степеней свободы и выбранного уровня значимости 0.05 (5%) для проверки тестируемой нулевой гипотезы равно 253.629.

В рассматриваемом примере фактическое значение статистики Фишера F равно 18299.7, что больше рассчитанного критического значения 253.629 статистики Фишера F , поэтому нулевая гипотеза отвергается для выбранного уровня значимости 0.05 (5%).

В этом случае говорят, что уравнение значимо в целом на уровне значимости 0.05 (5%), и соответствующий ему регрессор (ВВП) объясняет вариацию зависимой переменной CS (потребление).

Воспользовавшись вторым методом тестирования гипотез для рассматриваемого примера, получим, что p -значение для статистики Фишера $\text{Prob}(F\text{-statistic})$ близко к нулю, поэтому на любом достаточно малом уровне значимости α можно отвергнуть тестируемую нулевую гипотезу.

4. Проверка статистической линейной гипотезы общего вида.

Проверим статистическую линейную гипотезу вида $H\beta=q$, H – матрица, в которой количество строк равно количеству тестируемых ограничений, а количество столбцов равно количеству коэффициентов β_i в регрессионном уравнении, q – вектор правых частей ограничений, размерность которого равна количеству тестируемых ограничений. Для проверки гипотезы воспользуемся тестом Вальда (Wald Test).

Вспомним, что коэффициенты регрессионного уравнения в *Eviews* имеют обозначения $C(1)$ (коэффициент β_1 при первом регрессоре), $C(2)$ и так далее, последним коэффициентом будет константа регрессии. Это справедливо, если в спецификации регрессионного уравнения при его оценивании служебная переменная C указана в конце спецификации, как и было сделано в п.1 занятия 4.

Заметим, что пользователь может использовать другой порядок именованной констант, поставив служебную переменную C в начало спецификации регрессионного уравнения, что, впрочем, может быть не всегда удобным, поскольку в этом случае $C(1)$ будет именем константы регрессии, $C(2)$ – именем *первой* независимой переменной и т.д. Чтобы не ошибиться в записи линейного ограничения, откройте окно спецификации уравнения, выбрав в меню объекта Equation пункт View Representations. В разделе Estimation Equation записана спецификация уравнения через коэффициенты $C(1)$, $C(2)$ и т.д.

Для работы с тестом Вальда выберем в меню объекта Equation пункт View Coefficient Tests, Wald – Coefficient Restrictions. В диалоговом окне Wald Test необходимо ввести линейные ограничения на коэффициенты уравнения регрессии, разделяя их запятой. Помните, что ограничения должны быть линейно независимыми, а количество вводимых ограничений должно быть строго меньше количества оцениваемых переменных. Для парной регрессии из рассмотренного выше примера п.2 занятия 4 можно ввести только одно

ограничение, например, $C(1)+2\cdot C(2)=7$, что соответствует записи $\beta_1+2\cdot\beta_0=7$.

Результат тестирования возвращается в окне Equation, которое содержит имя уравнения, нулевую гипотезу и значения статистик напротив текста F-statistic и Chi-square (используются два различных варианта теста со статистиками Фишера и χ^2) вместе с соответствующими p -значениями. Будет ли нулевая гипотеза отвергнута или принята, зависит от выбранного уровня значимости при тестировании статистической гипотезы (см. п.1 данного занятия).

Обратите внимание, что результаты теста не сохраняются в отдельном объекте. Для сохранения результатов расчетов выберите в меню объекта Equation пункт Freeze, в результате будет создан новый объект Table, куда будет скопирован текущий текст результатов расчетов. Чтобы сохранить вновь созданный объект Table, выбрав в меню объекта пункт Name и введя допустимое имя без символов кириллицы. После нажатия кнопки «ОК» объект Table будет сохранен с введенным вами именем и отображен в рабочем окне *Eviews* в виде пиктограммы с изображением таблицы и надписью «Table».

Лабораторная работа 2. Обобщение и модификация задачи линейной регрессии.

Занятие 7. Мультиколлинеарность.

Задачи занятия: промоделировать полную и частичную мультиколлинеарность в линейной множественной регрессии.

1. Полная мультиколлинеарность в уравнении регрессии.

Явление полной мультиколлинеарности возникает при линейной зависимости переменных, включенных в правую часть оцениваемого уравнения регрессии. В этом случае оценок коэффициентов по методу наименьших квадратов не существует из-за вырожденности матрицы $X^T X$. Вспомните, что X – матрица, столбцы которой являются рядами данных, соответствующих независимым переменным регрессии, включая единичный вектор (если константа регрессии включена в состав регрессоров).

Промоделировать явление мультиколлинеарности в *Eviews* можно следующим образом. Создайте, например, новый ряд данных, который является линейной комбинацией уже существующих переменных с помощью уже знакомой вам команды *genr* (см. п.5 занятия 2). Это

можно сделать несколькими способами: во-первых, в окне ввода команд, во-вторых, выбрав в основном меню *Eviews* (или в меню рабочего окна) пункт *Procs, Generate Series*, или, в-третьих, нажав кнопку *Genr* в меню рабочего окна. В двух последних случаях в появившемся диалоговом меню *Generate Series by Equation* необходимо ввести команду *genr* и необходимое выражение.

Команда *genr z=5-7*x* означает, что ряд данных z является линейной комбинацией переменной x и единичного вектора (не забывайте, что если константа включена в уравнение регрессии, то единичный вектор является полноправным регрессором!).

Включите ряды данных z и x в качестве независимых переменных регрессии и оцените параметры уравнения регрессии методом наименьших квадратов (см. п.1 занятия 4). В данном случае объекта *Equation* создано не будет, и появится предупреждение об ошибке *Near singular matrix* (матрица вырождена), что говорит о несуществовании МНК-оценок.

В практике обработки статистических данных явление полной мультиколлинеарности в уравнении регрессии может возникнуть при включении в качестве независимых переменных одного и того же ряда данных, выраженного в различных единицах измерения. Такая ситуация может возникнуть и при работе с независимыми переменными, отражающими качественные характеристики (искусственные или *dummy*-переменные), что будет рассмотрен в п.4 занятия 7.

2. Частичная мультиколлинеарность в уравнении регрессии.

Явление частичной мультиколлинеарности на практике встречается гораздо чаще, чем случай полной мультиколлинеарности, и заключается в том, что регрессоры, включаемые в регрессионное уравнение, «почти» линейно зависимы. В этом случае оценки метода наименьших квадратов становятся неустойчивыми к начальным данным.

По каким признакам можно предполагать наличие частичной мультиколлинеарности в уравнении регрессии? Во-первых, должны насторожить слишком близкие к нулю значения стандартных отклонений оценок коэффициентов. Во-вторых, показателем наличия частичной мультиколлинеарности является сильная значимость уравнения в целом (высокое фактическое значение F -статистики) при одновременной незначимости отдельных коэффициентов регрессии кроме константы (низкие значения t -статистики).

Практическая рекомендация для анализа случая частичной мультиколлинеарности состоит в том, что необходимо построить матрицу корреляции для независимых переменных регрессии, *прежде* чем включать их в регрессионное уравнение. Можно использовать специальную процедуру последовательного присоединения регрессоров и методы факторного анализа для отбора независимых переменных. После оценивания коэффициентов регрессии необходимо проверить некоррелированность вектора остатков регрессии с независимыми переменными.

Занятия 8-9. Нелинейное уравнение регрессии. Искусственные (фиктивные) переменные.

Задачи занятия: рассмотреть приемы линеаризации в нелинейном регрессионном уравнении, дать экономическую интерпретацию коэффициентов в модели «в логарифмах», провести оценивание нелинейным методом наименьших квадратов. Сгенерировать искусственную переменную, включить ее в состав регрессоров, провести оценивание по методу наименьших квадратов и дать экономическую интерпретацию оцениваемых коэффициентов.

1. Нелинейное уравнение регрессии и его линеаризация.

Рассмотрим некоторые случаи нелинейных регрессионных уравнений, которые могут быть линеаризованы.

Заметьте, что регрессионные уравнения, нелинейные по независимым переменным, приводятся к стандартной постановке линейной множественной регрессии простой заменой переменных. Например, в

уравнениях $y = \beta_0 + \beta_1 \frac{1}{x} + \varepsilon$, $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \varepsilon$ можно сделать замену, введя

новую независимую переменную $z = \frac{1}{x}$ и $z = x^2$ соответственно. Тогда

будут получены линейные регрессионные уравнения $y = \beta_0 + \beta_1 z + \varepsilon$ и

$$y = \beta_0 + \beta_1 x + \beta_2 z + \varepsilon.$$

Уравнение регрессии с показательной зависимостью вида

$$y = \beta_0 \beta_1^{x_1} \beta_2^{x_2} \dots \beta_k^{x_k} + \varepsilon$$

после логарифмирования обеих частей приводится к виду $\ln y = \ln \beta_0 + x_1 \ln \beta_1 + x_2 \ln \beta_2 + \dots + x_k \ln \beta_k + \varepsilon$, а после замены переменной

вида $q = \ln y$ получим линейное уравнение множественной регрессии

$$q = \ln \beta_0 + x_1 \ln \beta_1 + x_2 \ln \beta_2 + \dots + x_k \ln \beta_k + \varepsilon.$$

Уравнение регрессии со степенной зависимостью вида $y = \beta_0 x_1^{\beta_1} x_2^{\beta_2} \dots x_k^{\beta_k} + \varepsilon$ после логарифмирования обеих частей приводится к линейному виду $\ln y = \ln \beta_0 + \beta_1 \ln x_1 + \beta_2 \ln x_2 + \dots + \beta_k \ln x_k + \varepsilon$.

После замены переменных вида $q = \ln y$, $z_i = \ln x_i$ получим линейное уравнение множественной регрессии $q = \ln \beta_0 + \beta_1 z_1 + \beta_2 z_2 + \dots + \beta_k z_k + \varepsilon$.

Величина ошибки ε считается аддитивной по определению регрессионного уравнения, однако, необходимо иметь в виду, что после такого преобразования нельзя считать неизменным вид распределения ошибок. Если ошибки в исходном уравнении предполагались независимыми, одинаково и нормально распределенными, то распределение ошибок в преобразованном уравнении нормальным не будет. Не забывайте проверить распределение ошибок на «нормальность» перед тестированием статистических гипотез, где существенно используется данное предположение.

2. Экономическая интерпретация оценок коэффициентов линеаризованной регрессии. Показатель эластичности.

На занятии 4 мы оценивали коэффициенты регрессионного уравнения $CS = C(1) \times GDP + C(2)$ методом наименьших квадратов (см. п.2 «Экономическая интерпретация оценок коэффициентов регрессии»), где CS – ряд данных для реального потребления, GDP – для реального ВВП в Великобритании для поквартальных данных за период с 1955 по 2001 г.

Построим и оценим уравнение с этими же переменными «в логарифмах», взяв натуральные логарифмы зависимой и независимой переменной и включив их в регрессионное уравнение. Дадим экономическую интерпретацию оценок коэффициентов получившихся регрессий.

Для регрессионного уравнения

$$\text{LOG}(CS) = C(1) \times \text{LOG}(GDP) + C(2),$$

где LOG – обозначение натурального логарифма в *Eviews*, оценка коэффициента $C(1)$ имеет смысл коэффициента эластичности, который имеет следующую экономическую интерпретацию: по имеющимся наблюдениям при росте (снижении) реального ВВП на $p\%$ реальное потребление увеличивается (снижается) в среднем на $C(1) \cdot p\%$ (обычно величину p выбирают равной 1).

Этот вывод справедлив для значений $p/100$, близких к нулю. Это легко увидеть, воспользовавшись тем, что при значениях x , близких к нулю, выполняется приближительное равенство $\ln(1+x) \approx x$ с точностью до линейных членов разложения в ряд Тейлора функции $\ln(1+x)$. При росте переменной GDP на $p\%$, значение регрессора $\ln(\text{GDP})$ увеличится по абсолютной величине примерно на $p/100$, а значит следует ожидать, что величина $\ln(\text{CS})$ изменится (в среднем) примерно на величину $C(1) \cdot p/100$. Таким образом, величина CS изменится (в среднем) примерно на $C(1) \cdot p\%$.

Аналогично можно получить экономическую интерпретацию оценки коэффициента $C(1)$ в регрессии

$$\text{CS} = C(1) \times \text{LOG}(\text{GDP}) + C(2),$$

а именно, по имеющимся наблюдениям при росте (снижении) реального ВВП на $p\%$ реальное потребление увеличится (уменьшится) в среднем примерно на $C(1) \cdot p$ единиц (по абсолютной величине).

Экономическая интерпретация коэффициента $C(1)$ в регрессии

$$\text{LOG}(\text{CS}) = C(1) \times \text{GDP} + C(2)$$

такова: по имеющимся наблюдениям при росте (снижении) реального ВВП на единицу реальное потребление увеличится (уменьшится) в среднем примерно на $C(1)$ процентов.

Насколько надежны полученные результаты с точки зрения тестирования гипотез, мы уже рассмотрели на занятии 5.

3. Нелинейный метод наименьших квадратов.

В случае, когда не удастся линеаризовать уравнение, используя простые приемы (см. п.1 данного занятия), остается возможность оценивания коэффициентов уравнения нелинейным методом наименьших квадратов (NLS, Nonlinear Least Squares).

Например, регрессионное уравнение, соответствующее CES-производственной функции вида $y = A(\delta K^{-\rho} + (1-\delta)L^{-\rho})^{-1/\rho}$, где y , K , L – валовой выпуск, используемый капитал и труд соответственно, с неизвестными параметрами A , δ , ρ нельзя линеаризовать методами п.1 данного занятия.

Для использования нелинейного метода наименьших квадратов необходимо проделать действия, изложенные в п.1 «Оценивание уравнения регрессии» занятия 4. Отличием будет другая лишь иная спецификация регрессионного уравнения в диалоговом окне Equation Specification. Для оценки параметров указанной выше CES-

производственной функции необходимо указать спецификацию регрессионного уравнения вида

$$y = C(1) * (C(2) * K^{(-C(3))} + (1 - C(2)) * L^{(-C(3))})^{-1/C(3)},$$

где $C(1)$ соответствует параметру A , $C(2)$ соответствует параметру δ , а $C(3)$ – параметру ρ производственной функции. Переменные y , K , L обозначают уже имеющиеся ряды данных для валового выпуска, используемого капитала и труда. Знак \wedge обозначает операцию возведения в степень, $*$ – операцию умножения.

При использовании нелинейного метода наименьших квадратов необходимо следить за возможностью неидентифицируемости коэффициентов уравнения. Например, в спецификации *Eviews* $y = a * \exp(b + c * x)$, неидентифицируемыми являются коэффициенты a и b . Это легко увидеть, взяв натуральный логарифм обеих частей уравнения, получим $\ln y = \ln a + b + cx$. Данное уравнение является линейным, однако, при спецификации такой линейной модели для оценивания, *Eviews* выдаст сообщение о вырожденности матрицы в силу полной мультиколлинеарности в модели. Действительно, в последнем уравнении в качестве независимых переменных участвуют два единичных вектора, соответствующие константам регрессии $\ln a$ и b , которые, очевидно, линейно зависимы.

4. Искусственные (фиктивные) переменные в регрессионном уравнении и их экономическая интерпретация.

Введем искусственную переменную, принимающую дискретный диапазон значений. Примером могут служить переменные, принимающие значения 0 или 1, которые называют бинарными. Искусственная переменная может описывать какой-то качественный признак изучаемого объекта или процесса: пол, цвет, этаж дома, наличие телефона, принадлежность к определенной профессии. Можно сгенерировать такую переменную, воспользовавшись уже знакомой вам командой *genr* (см п.5 занятия 2).

Создадим бинарную переменную y , значения которой будут равны 1 для наблюдений переменной x , значения которых больше некоторого наперед заданного числа, например 15, и для всех остальных наблюдений значения переменной y будут равны 0. Для этого необходимо ввести команду *genr* $y = (x > 15)$.

Экономическая интерпретация оценки коэффициента при бинарной переменной проста: по имеющимся наблюдениям, значениям искусственной переменной, равным 1, соответствует значение зависимой

переменной большее (при положительном значении коэффициента) или меньшее (при отрицательном значении коэффициента) в среднем на величину, равную оценке интерпретируемого коэффициента при бинарной переменной (в единицах измерения зависимой переменной).

Пусть изучается заработная плата мужчин и женщин одной профессии. Введем искусственную переменную, значение которой равно 1 для наблюдений, соответствующих заработной плате мужчины, и равно 0, если наблюдение соответствует заработной плате женщины. Тогда значение коэффициента при данной переменной, равное -3.5, в регрессии с зависимой переменной, описывающей среднедневную зарплату в рублях изучаемой группы людей, при значимости соответствующего коэффициента, позволяет сказать для выбранного уровня значимости, что среднедневная зарплата у мужчин по имеющимся наблюдениям в среднем меньше, чем у женщин на 3.5 руб.

Необходимо помнить, что включение в качестве факторов регрессии искусственных переменных может приводить к явлению полной мультиколлинеарности, которое мы рассматривали в п.1 занятия 6. В западной литературе ситуация полной мультиколлинеарности для искусственных переменных называется «dummy trap» (trap – «ловушка»).

В качестве примера рассмотрим четыре бинарные искусственные переменные, которые соответствуют четырем временам года: зиме, весне, лету и осени. Пусть значение первой переменной принимает значение 1, если соответствующее наблюдение относится к зиме, и 0, если оно относится к другому сезону. Пусть значение второй искусственной переменной, описывающей принадлежность к весеннему сезону, будет равно 1, если соответствующее наблюдение относится к весне, и 0, если оно относится к другому сезону. Аналогично составим искусственные переменные для лета и осени. Сумма таким образом составленных четырех искусственных переменных тождественно равна единичному вектору, размерность которого равна количеству наблюдений. Но единичный вектор соответствует константе регрессии, если она включена в регрессионное уравнение. Таким образом, все искусственные переменные и единичный вектор, соответствующий константе регрессии, линейно зависимы. В этом случае оценок коэффициентов уравнения регрессии по методу наименьших квадратов не существует из-за вырожденности соответствующей матрицы. Поэтому нельзя включать одновременно в уравнение регрессии все подобным образом связанные искусственные переменные.

Занятие 10. Линейные регрессионные модели с гетероскедастичными остатками.

Задачи занятия: Провести тест Уайта на гомоскедастичность остатков регрессии, сделать поправку оценки ковариационной матрицы ошибок в форме Уайта и Навье-Веста.

1. Тест Уайта на гомоскедастичность.

Нулевой гипотезой теста Уайта является гомоскедастичность остатков регрессии. *Eviews* позволяет работать с двумя вариантами данного теста: с включением попарных произведений регрессоров исходной модели (cross-terms) на втором шаге теста или без таковых (no cross-terms). Включение таких слагаемых позволяет учесть дополнительный источник возможного наличия гетероскедастичности остатков, и это приводит к повышению критического значения статистики χ^2 , поскольку увеличивается число степеней свободы.

Для работы с тестом Уайта выберем в меню объекта Equation пункт View, Residual Tests, White-Heteroscedasticity (cross-terms) или White-Heteroscedasticity (no cross-terms).

Результат тестирования возвращается в окне Equation, которое содержит значения статистики Фишера (F-statistic) и статистики χ^2 (Obs*R-squared) вместе с соответствующими *p*-значениями. Будет ли нулевая гипотеза отвергнута или принята, зависит от выбранного уровня значимости при тестировании статистической гипотезы (см. п.1 занятия 5). Далее возвращаются результаты оценивания второго шага теста Уайта, то есть регрессии квадрата остатков регрессии на соответствующие регрессоры.

Обратите внимание, что результаты теста не сохраняются в отдельном объекте. Для сохранения результатов расчетов выберите в меню объекта Equation пункт Freeze. В результате будет создан новый объект Table, куда будет скопирован текущий текст результатов расчетов. Можно сохранить вновь созданный объект Table, выбрав в меню объекта пункт Name и введя допустимое имя без символов кириллицы. После нажатия кнопки «ОК» объект Table будет сохранен с введенным именем и отображен в рабочем окне *Eviews* в виде пиктограммы с изображением таблицы и надписью «Table».

2. Поправка оценки ковариационной матрицы ошибок в форме Уайта и Навье-Веста.

Если нулевая гипотеза отвергнута для выбранного уровня значимости, то в регрессии присутствует гетероскедастичность остатков и поэтому необходимо сделать поправки оценок стандартных ошибок коэффициентов регрессии для ковариационной матрицы ошибок в форме Уайта или Навье-Веста. Для этого необходимо в диалоговом окне Equation Specification нажать кнопку Options. В появившемся диалоговом окне Options необходимо выбрать пункт Heteroscedasticity – White и затем повторить оценивание уравнения регрессии.

Лабораторная работа 3. Анализ временных данных в модели линейной регрессии и системы одновременных уравнений.

Занятия 11-12. Тест на структурную изменчивость и причинно-следственную связь.

Задачи занятий: Провести тест Чоу на структурную изменчивость коэффициентов регрессии и тест Гранжера на причинно-следственную связь между временными рядами.

1. Тест Чоу на структурную изменчивость коэффициентов регрессии.

Нулевой гипотезой теста Чоу является утверждение об отсутствии структурной изменчивости коэффициентов регрессии. Для работы с тестом Чоу выберите в меню объекта Equation пункт View Stability Tests, Chow Breakpoint Test. В появившемся диалоговом окне Chow Test необходимо ввести одну или несколько дат, разбивающих выборку на поддиапазоны.

Результат тестирования возвращается в окне Equation, которое содержит значения статистик F-statistic и Log likelihood ratio (используются два различных варианта теста) вместе с соответствующими p -значениями. Будет ли отвергнута или принята нулевая гипотеза, зависит от выбранного уровня значимости при тестировании статистической гипотезы (см. п.1 данного занятия).

Обратите внимание, что результаты теста не сохраняются в отдельном объекте. Для сохранения результатов расчетов выберите в меню объекта Equation пункт Freeze, в результате будет создан новый объект Table, куда будет скопирован текущий текст результатов расчетов. Чтобы сохранить вновь созданный объект Table, выберите в меню объекта пункт Name и введите допустимое имя без символов кириллицы. После нажатия кнопки «ОК» объект Table будет сохранен с

введенным вами именем и отображен в рабочем окне *Eviews* в виде пиктограммы с изображением таблицы и надписью «Table».

2. Тест Гранжера на причинно-следственную связь между временными рядами.

Для работы с тестом Гранжера необходимо последовательно выбрать левой кнопкой мыши несколько объектов *Series* (рядов данных) в рабочем окне *Eviews*, удерживая клавишу *Ctrl*. Затем, отпустив клавишу *Ctrl* и нажав правую кнопку мыши, необходимо выбрать в появившемся меню *Open as Group*. В меню объекта *Group* выберите пункт *View Granger causality*. В появившемся диалоговом окне *Lag Specification* необходимо указать количество лагов (*lags to include*) в качестве параметра теста. По умолчанию этот параметр равен 2.

Результат тестирования возвращается в окне *Group*, которое содержит формулировку нулевой гипотезы теста (например, переменная x не является причиной по Гранжеру для переменной y) значения статистики Фишера (*F-statistic*) и соответствующее p -значение. Будет ли нулевая гипотеза отвергнута или принята, зависит от выбранного уровня значимости при тестировании статистической гипотезы (см. п.1 занятия 5).

Обратите внимание, что результаты теста не сохраняются в отдельном объекте. Для сохранения результатов расчетов выберите в меню объекта *Group* пункт *Freeze*, в результате будет создан новый объект *Table*, куда будет скопирован текущий текст результатов расчетов. Чтобы сохранить вновь созданный объект *Table*, выберите в меню объекта пункт *Name* и введите допустимое имя без символов кириллицы. После нажатия кнопки «ОК» объект *Table* будет сохранен с введенным вами именем и отображен в рабочем окне *Eviews* в виде пиктограммы с изображением таблицы и надписью «Table».

Занятие 13. Тесты на автокорреляцию остатков. Оценивание при наличии автокорреляции остатков.

Задачи занятия: Провести тесты Дарбина-Уотсона и LM-тест Бреуша-Годфри на автокорреляцию первого порядка для остатков линейного уравнения регрессии. Провести оценивание линейного уравнения регрессии при наличии автокорреляции остатков.

Статистика Дарбина-Уотсона предназначена для анализа автокорреляции первого порядка остатков линейного уравнения регрессии. Значение статистики Дарбина-Уотсона в *Eviews* (Durbin-Watson stat) возвращается при оценивании коэффициентов уравнения методом наименьших квадратов в окне Equation. Заметьте, что статистика Дарбина-Уотсона не может применяться при авторегрессионной спецификации модели. К сожалению, не возвращаются значения для интервала неопределенности этого теста.

Тест LM (Lagrange Multipliers) Бреуша-Годфри позволяет тестировать наличие автокорреляции более высоких порядков и применяться в случае авторегрессионной спецификации модели.

Нулевой гипотезой теста LM является отсутствие автокорреляции остатков регрессии для всех порядков автокорреляции ниже или равной заданному. Для работы с тестом LM выберите в меню объекта Equation пункт View Residual Tests, Serial Correlation LM Test. В появившемся диалоговом окне Lag Specification необходимо указать количество лагов (lags to include) в качестве параметра теста, который задает максимальный порядок автокорреляции, до которого тестируется отсутствие автокорреляции остатков. По умолчанию этот параметр равен 2.

Результат тестирования возвращается в окне Equation, которое содержит значения статистик Фишера (F-statistic) и χ^2 (Obs*R-squared) вместе с соответствующими *p*-значениями. Будет ли нулевая гипотеза отвергнута или принята, зависит от выбранного уровня значимости при тестировании статистической гипотезы (см. п.1 занятия 5).

Обратите внимание, что результаты теста не сохраняются в отдельном объекте. Для сохранения результатов расчетов выберите в меню объекта Equation пункт Freeze, в результате будет создан новый объект Table, куда будет скопирован текущий текст результатов расчетов. Чтобы сохранить вновь созданный объект Table, выберите в меню объекта пункт Name и введите допустимое имя без символов кириллицы. После нажатия кнопки «ОК» объект Table будет сохранен с введенным вами именем и отображен в рабочем окне *Eviews* в виде пиктограммы с изображением таблицы и надписью «Table».

Для оценивания линейного уравнения регрессии при наличии автокорреляции остатков необходимо указать авторегрессионную спецификацию уравнения регрессии. В окне Equation

Specification (см. п.1 занятия 4) в конце спецификации необходимо указать специальное обозначение регрессора $ar(1)$, что соответствует регрессии, в которой остатки считаются автокоррелированными первого порядка с неизвестным параметром автокорреляции. В этом случае будет применяться специальный нелинейный двушаговый метод оценивания сначала коэффициента автокорреляции, а затем и основного уравнения регрессии.

Занятие 14. Оценивание модели полиномиальных лагов (модель распределенных лагов Алмон).

Задачи занятия: Оценить уравнения регрессий, используя для представления коэффициентов модели полиномиальных лагов.

Переменная x_{t-k} называется k -ым лагом (сдвигом), $k > 0$, для ряда наблюдений x_t . Если значения x_t заданы для наблюдений с номерами $1, \dots, m$, то первые $m-k$ значений ряда x_{t-k} будут равны значениям ряда x_t , начиная с номера $t=k+1$ до номера $t=m$, а значения для последующих номеров не будут определены. Такой ряд наблюдений моделирует запаздывающую переменную, т.е. такую, что ее реализация для текущего момента времени оказывает воздействие на значения зависимой переменной в последующие моменты времени.

Можно построить лаг x_{t+k} и говорить в этом случае об опережающей переменной. Первые k значений ряда x_{t+k} не будут определены, а начиная с номера $k+1$ они будут равны значениям ряда x_t , начиная с номера $t=1$ до номера $t=m-k$. Обратите внимание на отличие спецификации уравнения регрессии с лаговыми независимыми переменными от авторегрессионной модели.

Спецификация уравнения регрессии с p (запаздывающими) лагами независимой переменной x_t переменными называются моделью распределенных лагов, которая имеет вид

$$y_t = b + g_0 x_t + g_1 x_{t-1} + \dots + g_p x_{t-p} + e_t. \quad (*)$$

Недостаток такой спецификации регрессионного уравнения состоит в том, что если количество лагов p достаточно велико, то может возникнуть проблема частичной мультиколлинеарности. С точки зрения качества регрессии в целом увеличение числа регрессоров завышает величину коэффициента детерминации.

Модель полиномиальных лагов предполагает использование небольшого числа оцениваемых параметров относительно количества лагов p . Для уравнения регрессии (*) предполагается, что оцениваемые параметры g_i связаны с другими параметрами w_k соотношением

$g_i = \sum_{k=0}^s i^k w_k$. За счет выбора количества оцениваемых параметров w_k , равного $s+1$, существенно меньшего, чем количество лагов исходной модели, которое равно p , удается избежать проблем с оцениванием исходного уравнения с большим числом лагов.

Фактически в данной модели распределенных лагов оцениваются коэффициенты уравнения $y_t = b + w_0 z_0 + w_1 z_1 + \dots + w_s z_s + e_t$, где z_k – новые регрессоры, связанные с лагами x соотношением $z_k = \sum_{j=0}^p j^k x_{t-j}$.

Для оценивания модели полиномиального лага в *Eviews* используется функция *pdl*, которая имеет три аргумента. Первым аргументом является имя независимой переменной, лаги которой включаются в уравнение регрессии. Вторым аргумент функции отвечает за количество лагов независимой переменной в исходной модели p , третий аргумент задает степень полинома s в представлении коэффициентов g_i .

Для оценивания модели полиномиального лага по переменной x в *Eviews* необходимо в строке ввода команд или диалоговом окне Equation Specification указать в виде дополнительного регрессора выражение *pdl(x,8,3)*, что означает оценивание регрессионной зависимости в модели полиномиальных лагов с независимыми переменными $x_t, x_{t-1}, \dots, x_{t-8}$ и представлением коэффициентов g_i в виде

$$g_i = \sum_{k=0}^3 i^k w_k.$$

Eviews возвращает результаты оценивания не только параметров w_k , но и коэффициентов g_i исходного уравнения регрессии. В модель распределенных лагов (*) могут быть включены и другие регрессоры, не обязательно со своими лагами.

Занятие 15. Прогнозирование в регрессионных моделях.

Задачи занятия: Построение точечного и интервального прогноза в линейных регрессионных моделях.

При построении прогнозов в регрессионных моделях следует учитывать, что прогнозные значения объясняемой переменной, во-первых, могут быть вычислены для наблюдений внутри основного диапазона данных (range) или, во-вторых, для наблюдений, находящихся вне этого диапазона. В обоих случаях говорят о прогнозных значениях (fitted) объясняемой переменной, хотя понятие прогноза зачастую понимается более узко и связывается только со вторым случаем.

Eviews строит оценку среднеквадратической ошибки прогноза в предположении о случайности коэффициентов регрессионной модели и среднеквадратической ошибки остатков регрессии.

Если необходимо построить прогноз для значений объясняемой переменной за пределами исходного диапазона данных, необходимо сначала расширить основной диапазон данных (*range*) на дополнительное число наблюдений, для которых будет проводиться прогнозирование.

Вспомним (см. занятия 3-4, п.1), что увеличить или уменьшить основной диапазон данных (*range*) можно, сделав двойной щелчок по диапазону данных (*range*) в рабочем окне *Eviews* либо выбрав в рабочем окне *Eviews* пункт меню *Procs, Change Workfile range*. В случае расширения границ диапазона данных *Eviews* добавит ко всем рядам данных необходимое количество наблюдений. Значения рядов данных для вновь добавленных наблюдений будут не определены и помечены в таблице значений объекта *Series* как *NA (not available)* до тех пор, пока вы их не зададите явно или не присвоите им значения в результате вычислений, например, с помощью команды *genr*. В случае сокращения диапазона данных произойдет потеря значений для наблюдений во ВСЕХ рядах данных рабочего файла *Eviews*.

Далее необходимо задать значения независимых переменных для дополнительных наблюдений, для которых будет осуществляться прогнозирование. Эти значения, в свою очередь, могут быть получены в результате прогнозирования по простой трендовой модели или в самостоятельной регрессионной модели.

Для построения прогноза объясняемой переменной в *Eviews* необходимо выбрать пункт *Forecast* в меню объекта *Equation*. В появившемся диалоговом окне *Forecast* необходимо указать в поле *Forecast name* имя нового ряда данных, в котором будет сохранено прогнозное значение для объясняемой переменной (точечный прогноз). **ОСТОРОЖНО!** Если такое имя ряда данных уже есть в вашем рабочем файле, то его значения будут перезаписаны! По умолчанию *Eviews* предлагает создать имя ряда данных, состоящее из имени объясняемой переменной и дополнительной буквы *f* в конце имени. В поле *S.E. (optional)* можно указать имя нового ряда данных, в который будут сохранены значения ошибки прогноза для каждого номера наблюдения, которые можно использовать для построения интервального прогноза.

В поле *Sample range for forecast* указывается поддиапазон данных, для которых будут рассчитываться прогнозные значения. По

умолчанию здесь указан текущий поддиапазон данных (sample), который нужно изменить, если необходимо, чтобы прогнозные значения вычислялись только для вновь добавленных наблюдений.

Пункт Insert actuals for out-of-sample по умолчанию отмечен, что предполагает заполнение ряда данных прогнозных значений фактическими значениями объясняемой переменной для наблюдений внутри основного поддиапазона данных.

Если поле Do Graph отмечено, то в качестве результата будут построены графики точечного и интервального прогноза объясняемой переменной и показатели качества регрессии внутри существующего объекта Equation, которые можно после выбора пункта Freeze в меню объекта Equation сохранить в новый объект Graph. Если поле Do Graph не отмечено, будут сохранены только показатели качества регрессии внутри существующего объекта Equation, который можно сохранить в новый объект Table, выбрав пункт Freeze меню объекта Equation.

В случае авторегрессионной спецификации уравнения можно прогнозировать значения объясняемой переменной на несколько наблюдений вперед, указав пункт dynamic forecast в диалоговом окне Forecast и указав диапазон наблюдений для прогнозирования в поле Sample range for forecast. В этом случае будет использована оценка среднеквадратической ошибки прогноза с учетом автокоррелированности остатков регрессионной модели. Если значения объясняемой переменной для прогнозируемого диапазона значений известны, можно воспользоваться прогнозированием, применяемым в обычной регрессионной модели (static forecast).

Чтобы не учитывать авторегрессионные компоненты в правой части уравнения регрессии, необходимо выбрать пункт Structural (ignore ARMA) в диалоговом окне Forecast.

Занятия 16-17. Система линейных одновременных уравнений и методы ее оценивания.

Задачи занятия: Составить систему одновременных макроэкономических уравнений, выбрать инструментальные переменные и провести оценивание методами SUR, TSLS (2SLS) и 3SLS.

Для работы с системами одновременных уравнений в *Eviews* предусмотрены объекты System и Model.

Для оценивания параметров системы одновременных уравнений различными статистическими методами необходимо создать объект *System*. Для этого необходимо в меню рабочего окна *Eviews* выбрать пункт *Objects, New Object*. В появившемся диалоговом окне *New Object* необходимо выбрать тип объекта (*Type of Object*) «*System*» и указать его имя в поле *Name for Object*. По умолчанию вновь создаваемый объект имеет имя *Untitled*. После нажатия кнопки «*OK*» появится пустое окно объекта *System*, где необходимо указать спецификацию системы одновременных уравнений для *Eviews*. Для изменения имени объекта *System* необходимо выбрать пункт меню *Name*.

Рассмотрим в качестве примера систему макроэкономических уравнений

$$c_t = a_1 + b_1 y_t + g c_{t-1} + e_t, \quad (1)$$

$$x_t = a_2 + b_2 y_t + h z_t + u_t, \quad (2)$$

где c_t, y_t, x_t, z_t – ряды данных для потребления, ВВП, инвестиций и государственных расходов, e_t, u_t – ошибки наблюдений.

При оценивании неизвестных параметров методами *SUR, TSLS (2SLS) и 3SLS* в *Eviews* помимо уравнений необходимо указать список инструментальных переменных или, для краткости, инструментов. Для спецификации уравнений можно использовать способ, рассмотренный в п.1 занятия 5. Рассмотрим здесь другой способ спецификации уравнений.

Спецификация системы макроэкономических уравнений (1),(2) в *Eviews* имеет следующий вид:

$$cs = c(1) + c(2) * gdp + c(3) * cs(-1) @ gdp(-1) cs(-1)$$

$$inv = c(4) + c(5) * gdp + c(6) * gov @ gdp(-1) gov$$

Данную спецификацию уравнений необходимо ввести в пустое окно объекта *System*. Обратите внимание, что в объекте *System* не допускается спецификация тождеств (*identities*). Чтобы их учесть при оценивании неизвестных коэффициентов, необходимо сделать предварительное преобразование переменных.

Имена рядов данных cs, gdp, inv, gov соответствуют переменным c_t, y_t, x_t, z_t в уравнениях системы (1)-(2). Ошибки наблюдений e_t, u_t в спецификации уравнений не указываются.

После символа @ перечисляются инструментальные переменные при оценивании каждого уравнения системы. Количество инструментов вместе с константой не может быть меньше количества переменных в правой части оцениваемого уравнения. В качестве инструментальных

переменных можно выбирать не только экзогенные переменные, но и лаги для эндогенных переменных.

Для оценивания коэффициентов системы методами SUR, TSLS (2SLS) и 3SLS необходимо выбрать в меню объекта System пункт Estimate (второй способ: выбрать пункт меню Procs, Estimate). В появившемся диалоговом окне System Estimation необходимо выбрать метод для оценивания коэффициентов (Estimation Method). В разделе Iteration Control можно указать различные способы организации численных методов для вычисления оценок коэффициентов и матрицы весов.

Результаты оценивания возвращаются в рабочее окно объекта System. Для сохранения результатов расчетов выберите в меню этих объектов пункт Freeze, в результате будет создан новый объект Table, куда будет скопирован текущий текст результатов расчетов. Необходимо сохранить вновь созданный объект Table, выбрав в меню объекта пункт Name и введя допустимое имя без символов кириллицы. После нажатия кнопки «ОК» объект Table будет сохранен с введенным вами именем и отображен в рабочем окне *Eviews* в виде пиктограммы с изображением таблицы и надписью «Table».

Изменить спецификацию системы уравнений можно выбрав в меню объекта System пункт View, System Specification или пункт меню Spec. Для отображения результатов оценивания в меню объекта System выберите пункт меню View, Estimation Output или пункт меню Stats. Выбрав пункт меню View, Residuals, Graphs или Resids в меню объекта System, можно построить графики остатков регрессионных уравнений, входящих в систему. Графики остатков отображаются в окне объекта System, а после выбора пункта Freeze создается объект Graph, с которым можно работать отдельно от самого объекта System, например, сохранить его со своим именем, изменить настройки отображения графиков и т.п. в меню объекта Graph.

Для коэффициентов регрессионных уравнений системы можно тестировать набор линейных гипотез с помощью теста Вальда (подробнее см. п.4 занятия 6). Для этого необходимо в меню объекта System выбрать пункт меню View, Wald Coefficient Tests.

Объект Model используется для прогнозирования значений эндогенных переменных системы одновременных уравнений, когда значения коэффициентов уже оценены и считаются известными. Объект Model допускает использование тождеств (*identities*) в отличие от

объекта System. Для создания объекта Model с уравнениями регрессионных прямых (уравнениями с подставленными коэффициентами) из объекта System необходимо выбрать пункт меню Procs, Make Model в меню объекта System.

Лабораторные работы для самостоятельного выполнения

Лабораторная работа 1. Множественная линейная регрессия.

Исходные данные: по российским регионам из файла *rusreg.xls*.

Все расчеты сделать в *Eviews* 3.1 и в *Excel*. Сохранить результаты в рабочих файлах <имя_файла>.wfl и <имя_файла>.xls.

1. Проведите предварительную подготовку, обработку и проверку данных для работы с надстройкой «Анализ данных» в *Excel*. Импортируйте данные в *Eviews*.
2. Постройте гистограмму по рядам наблюдений. Постройте корреляционную матрицу для независимых переменных. Рассчитайте основные статистические показатели рядов данных (количество наблюдений, среднее, среднеквадратическое отклонение, минимальное и максимальное значение и т.д.).
3. Составьте уравнение регрессии, основываясь на известных моделях экономической теории. Рассчитать параметры линейного уравнения регрессии.
4. Оцените статистическую значимость параметров регрессионной модели с помощью *t*-статистики. Гипотезу о значимости уравнения в целом проверьте с помощью *F*-критерия.

Лабораторная работа 2. Обобщение и модификация задачи линейной регрессии.

Исходные данные: по российским регионам из файла *rusreg.xls*.

Все расчеты сделать в *Eviews* 3.1 и сохранить результаты в рабочем файле <имя_файла>.wfl.

1. Дать экономическую интерпретацию коэффициентов, оцененных в лабораторной работе 1. Построить уравнение «в логарифмах» и интерпретировать соответствующие коэффициенты регрессии. Имеют ли полученные коэффициенты экономический смысл?
2. В регрессии лабораторной работы 1 ввести переменную, соответствующую качественному признаку (*dummy*), дать экономическую интерпретацию оцениваемых коэффициентов.
3. В регрессии лабораторной работы 1 тестировать какую-либо гипотезу вида $H\hat{\beta} = r$ для вектора оценок коэффициентов $\hat{\beta}$.
4. Промоделировать полную и частичную мультиколлинеарность в регрессии лабораторной работы 1, проанализировать результаты и сделать выводы.

5. Тестировать гипотезу о гомоскедастичности остатков регрессии лабораторной работы 1 и оценить регрессию с поправками ковариационной матрицы ошибок в форме Уайта.

Лабораторная работа 3. Анализ временных данных в модели линейной регрессии и системы одновременных уравнений.

Исходные данные: из файла *UK.xls* (поквартальная динамика компонент расходов в реальном ВВП Великобритании, 1955-2001 гг.).

Все расчеты сделать в *Eviews 3.1* и сохранить результаты в рабочем файле <имя_файла>.wfl.

1. Построить описательную статистику и графики компонентов ВВП, гистограмму ошибок основного макроэкономического тождества.
2. Провести тест на причинно-следственную связь между какой-либо парой переменных (тест Гранжера).
3. Составить какое-либо макроэкономическое уравнение для одной из компонентов расходов в ВВП, провести его оценивание методом наименьших квадратов и провести
 - а) тестирование статистической гипотезы о структурной изменчивости коэффициента регрессии при какой-либо независимой переменной (тест Чоу).
 - б) тестирование на наличие автокорреляции остатков в оцененной регрессии (тест Дарбина-Уотсона, LM-test).
4. Сделать безусловный поквартальный прогноз объясняемой переменной на 2002 г. в уравнении регрессии п. 3.
5. Составить какую-либо систему одновременных макроэкономических уравнений, выбрать инструментальные переменные и провести оценивание методами TSLS (2SLS) и 3SLS.

Практические задания и рекомендации к их выполнению

1. (Gujarati [15], 3.1) Докажите следующие утверждения о предположениях модели парной регрессии $Y_i = \beta_1 + \beta_2 X_i + u_i$.

а) Условие $E(u_i | X_i) = 0$ эквивалентно условию $E(Y_i | X_i) = \beta_1 + \beta_2 X_i$.

б) Условие $\text{cov}(u_i, u_j) = 0, i \neq j$ эквивалентно условию $\text{cov}(Y_i, Y_j) = 0, i \neq j$.

в) Условие $\text{var}(u_i | X_i) = \sigma^2$ эквивалентно условию $\text{var}(Y_i | X_i) = \sigma^2$.

Решение:

а)

$$Y_i = \beta_1 + \beta_2 X_i + u_i,$$

$$E(Y_i | X_i) = E(\beta_1 + \beta_2 X_i + u_i | X_i) = \beta_1 + \beta_2 X_i + E(u_i | X_i),$$

$$E(u_i | X_i) = 0 \Rightarrow E(Y_i | X_i) = \beta_1 + \beta_2 X_i + 0 = \beta_1 + \beta_2 X_i,$$

$$E(Y_i | X_i) = \beta_1 + \beta_2 X_i \Rightarrow E(u_i | X_i) = E(Y_i | X_i) - \beta_1 + \beta_2 X_i = 0.$$

б)

$$\text{cov}(Y_i, Y_j) = \text{cov}(\beta_1 + \beta_2 X_i + u_i, \beta_1 + \beta_2 X_j + u_j) =$$

$$= E(\beta_1 + \beta_2 X_i + u_i - E(\beta_1 + \beta_2 X_i + u_i))(\beta_1 + \beta_2 X_j + u_j - E(\beta_1 + \beta_2 X_j + u_j)) =$$

$$= E(\beta_1 + \beta_2 X_i + u_i - \beta_1 - \beta_2 X_i - E(u_i))(\beta_1 + \beta_2 X_j + u_j - \beta_1 - \beta_2 X_j - E(u_j)) =$$

$$= E(u_i - E(u_i))(u_j - E(u_j)) = \text{cov}(u_i, u_j).$$

в)

$$\text{var}(Y_i | X_i) = \text{var}(\beta_1 + \beta_2 X_i + u_i | X_i) =$$

$$= E(\beta_1 + \beta_2 X_i + u_i - E(\beta_1 + \beta_2 X_i + u_i | X_i) | X_i)(\beta_1 + \beta_2 X_i + u_i - E(\beta_1 + \beta_2 X_i +$$

$$+ u_i | X_i) | X_i) = E(\beta_1 + \beta_2 X_i + u_i - \beta_1 - \beta_2 X_i - E(u_i | X_i) | X_i)(\beta_1 + \beta_2 X_i + u_i -$$

$$- \beta_1 - \beta_2 X_i - E(u_i | X_i) | X_i) = E(u_i - E(u_i | X_i) | X_i)(u_i - E(u_i | X_i) | X_i) =$$

$$= \text{var}(u_i | X_i).$$

2. (Gujarati [15], 3.9) Рассмотрим два уравнения линейной парной регрессии $Y_i = \beta_1 + \beta_2 X_i + u_i$ и $Y_i = \alpha_1 + \alpha_2 (X_i - \bar{X}) + v_i$.

Совпадают ли оценки метода наименьших квадратов для коэффициентов β_1 и α_1 ? А их вариации? Совпадают ли оценки метода наименьших квадратов для коэффициентов β_2 и α_2 ? А их вариации?

Решение:

Представим второе уравнение в следующем виде:

$$Y_i = \alpha_1 + \alpha_2 (X_i - \bar{X}) + v_i = \alpha_1 - \alpha_2 \bar{X} + \alpha_2 X_i + v_i = \alpha_0 + \alpha_2 X_i + v_i.$$

По формулам метода наименьших квадратов получим

$$\tilde{\beta}_2 = \Sigma(X_i - \bar{X})(Y_i - \bar{Y}) / \Sigma(X_i - \bar{X})^2, \quad \tilde{\beta}_1 = \bar{Y} - \tilde{\beta}_2 \bar{X},$$

$$\tilde{\alpha}_2 = \Sigma(X_i - \bar{X})(Y_i - \bar{Y}) / \Sigma(X_i - \bar{X})^2, \quad \tilde{\alpha}_1 = \bar{Y} - \tilde{\alpha}_2 \bar{X} + \tilde{\alpha}_2 \bar{X} = \bar{Y}.$$

Имеем $\text{var}(\tilde{\beta}_2) = [\Sigma X_i^2 / (n \Sigma(X_i - \bar{X})^2)] \sigma^2$, $\text{var}(\tilde{\alpha}_1) = \sigma^2/n$.

Так как $\tilde{\beta}_2$ и $\tilde{\alpha}_2$ представляют одну и ту же случайную величину для генеральной совокупности, то есть $\tilde{\beta}_2 \equiv \tilde{\alpha}_2$, то $\text{var}(\tilde{\beta}_2) = \text{var}(\tilde{\alpha}_2)$.

Замечание. Из того, что $\tilde{\beta}_2 = \tilde{\alpha}_2$ для данной выборки еще не следует, что $\text{var}(\tilde{\beta}_2) = \text{var}(\tilde{\alpha}_2)$. Для фиксированной выборки $\tilde{\beta}_2$ и $\tilde{\alpha}_2$ представляют собой действительные числа, значения которых могут совпадать для одной выборки, и быть отличными для другой.

3. (Gujarati [15], 3.10) Рассмотрим уравнение регрессии в нормальной форме $y_i = \beta_1 + \beta_2 x_i + u_i$, где $y_i = Y_i - \bar{Y}$, $x_i = X_i - \bar{X}$. Чему равна оценка коэффициента β_1 по методу наименьших квадратов?

Решение:

По формулам метода наименьших квадратов получаем оценку $\tilde{\beta}_2 = \frac{\sum (y_i - \bar{y})(x_i - \bar{x})}{\sum (x_i - \bar{x})^2} = \frac{\sum (y_i - 0)(x_i - 0)}{\sum (x_i - 0)^2} = \frac{\sum (Y_i - \bar{Y})(X_i - \bar{X})}{\sum (X_i - \bar{X})^2}$, так как

$\bar{y} = \overline{Y - \bar{Y}} = \bar{Y} - \bar{Y} = 0$ и $\bar{x} = \overline{X - \bar{X}} = \bar{X} - \bar{X} = 0$. Поскольку $\tilde{\beta}_1 = \bar{y} - \tilde{\beta}_2 \bar{x}$, то, следовательно, $\tilde{\beta}_1 = 0 - \tilde{\beta}_2 \cdot 0 = 0$.

4. (Gujarati [15], 3.13) Пусть X_1, X_2, X_3 – попарно некоррелированные переменные с одинаковой дисперсией. Покажите, что коэффициент корреляции между $X_1 + X_2$ и $X_2 + X_3$ равен 1/2. Поясните, почему коэффициент корреляции не равен 0.

Решение:

Некоррелированность X_1, X_2, X_3 означает, что соответствующие коэффициенты ковариации (а значит и корреляции) равны 0, то есть

$$\text{cov}(X_1, X_2) = 0, \text{cov}(X_1, X_3) = 0, \text{cov}(X_2, X_3) = 0.$$

По определению коэффициента корреляции получаем, что

$$\begin{aligned} \text{corr}(X_1 + X_2, X_2 + X_3) &= \\ &= \text{cov}(X_1 + X_2, X_2 + X_3) / [\text{var}(X_1 + X_2) \text{var}(X_2 + X_3)]^{1/2} = \\ &= [\text{cov}(X_1, X_2) + \text{cov}(X_1, X_3) + \text{var}(X_2) + \text{cov}(X_2, X_3)] / \\ &= [(\text{var}(X_1) + \text{var}(X_2) - 2 \text{cov}(X_1, X_2))(\text{var}(X_2) + \text{var}(X_3) - 2 \text{cov}(X_2, X_3))]^{1/2} = \\ &= \text{var}(X_2) / [\text{var}(X_1 + X_2) \text{var}(X_2 + X_3)]^{1/2} = \sigma^2 / [2\sigma^2 2\sigma^2]^{1/2} = 1/2. \end{aligned}$$

Коэффициент корреляции не равен нулю, поскольку он включает корреляцию между X_1 и X_3 и вариацию X_2 . Тогда на вариацию X_2 влияют

две коррелированные переменные, что объясняет значение коэффициента корреляции $1/2$.

5. (Gujarati [15], 3.14) Предположите, что в регрессии $Y_i = \beta_1 + \beta_2 X_i + u_i$ каждое X_i умножено на некоторую константу, например, 2. Изменятся ли остатки регрессии (полученные после применения метода наименьших квадратов) и прогнозные значения для Y_i (\hat{Y}_i) ? Что произойдет, если каждое X_i будет увеличено на константу, например, число 2 ? Ответ поясните.

Решение:

Рассмотрим две регрессии:

$$Y_i = \alpha_1 + \alpha_2 X_i + \varepsilon_i, \text{ и } Y_i = \beta_1 + \beta_2 Z_i + u_i, \text{ где } Z_i = 2X_i.$$

По формулам метода наименьших квадратов получим

$$\tilde{\beta}_2 = \frac{\sum (Z_i - \bar{Z})(Y_i - \bar{Y})}{\sum (Z_i - \bar{Z})^2} = \frac{\sum (2X_i - 2\bar{X})(Z_i - \bar{Z})}{\sum (2X_i - 2\bar{X})^2} = (1/2)\tilde{\alpha}_2,$$

$$\tilde{\beta}_1 = \bar{Y} - \tilde{\beta}_2 \bar{Z} = \bar{Y} - (1/2)\tilde{\alpha}_2(2\bar{X}) = \bar{Y} - \tilde{\alpha}_2 \bar{X} = \tilde{\alpha}_1.$$

Далее вычислим прогнозные значения для Y_i во второй регрессии $\hat{Y}_i = \tilde{\beta}_1 + \tilde{\beta}_2 Z_i = \tilde{\alpha}_1 + (1/2)\tilde{\alpha}_2(2\bar{X}) = \tilde{\alpha}_1 + \tilde{\alpha}_2 \bar{X}$, поэтому прогнозные значения остаются неизменными.

Для остатков регрессии имеем

$$\hat{u}_i = \hat{Y}_i - \tilde{\alpha}_1 - \tilde{\alpha}_2 X_i = \hat{Y}_i - \tilde{\alpha}_1 - (1/2)\tilde{\alpha}_2(2X_i) = \hat{Y}_i - \tilde{\beta}_1 - \tilde{\beta}_2 Z_i = \varepsilon_i,$$

поэтому вектор остатков не изменяется.

Если мы добавим константу, например 2, к каждому наблюдению X_i , то результаты оценивания также не изменятся.

По формулам метода наименьших квадратов получим, что $\tilde{\beta}_2 = \tilde{\alpha}_2$ и $\tilde{\beta}_1 = \tilde{\alpha}_1 - 2\tilde{\beta}_2 = \tilde{\alpha}_1 - 2\tilde{\alpha}_2$. Остатки и прогнозные значения не изменяются, поскольку по сути мы не добавляем при указанных линейных преобразованиях регрессоров какую-то новую информацию о взаимосвязи X и Y .

6. (Gujarati [15], 3.16) Объясните, является ли утверждение истинным, ложным или ответ неоднозначен.

а) Поскольку значение коэффициента корреляции между любыми рядами данных X и Y находится на отрезке от -1 до 1 , это означает, что ковариация $cov(X, Y)$ также равна 1 .

б) Если значение коэффициента корреляции между рядами данных равно 0 , это означает, что не существует какой-либо связи между ними.

в) Если провести регрессию оцененных значений \hat{Y} на фактические значения Y , то оценка константы регрессии по МНК будет равна 0, а оценка по МНК коэффициента при \hat{Y} будет равна 1.

Решение:

а) По определению коэффициента корреляции

$$\text{corr}(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X) \cdot \text{var}(Y)}},$$

откуда следует, что $|\text{cov}(X, Y)| \leq \sqrt{\text{var}(X) \cdot \text{var}(Y)}$, поэтому на вопрос а) ответ отрицательный.

б) Коэффициент корреляции $\text{corr}(X, Y)$ характеризует степень линейной зависимости между X и Y , поэтому условие $\text{corr}(X, Y) = 0$ не противоречит возможности наличия взаимосвязи, например, другого типа, между X и Y , поэтому на вопрос б) ответ отрицательный.

в) Рассмотрим две регрессии:

$$Y_i = \alpha_1 + \alpha_2 X_i + \varepsilon_i, \quad (*)$$

$$Y_i = \gamma_1 + \gamma_2 \hat{Y}_i + u_i,$$

где \hat{Y}_i – прогнозные значения Y_i в регрессии (*).

По формулам метода наименьших квадратов получаем

$$\begin{aligned} \tilde{\gamma}_2 &= \frac{\sum (\hat{Y}_i - \bar{\hat{Y}})(Y_i - \bar{Y})}{\sum (\hat{Y}_i - \bar{\hat{Y}})^2} = \frac{\sum (\tilde{\alpha}_1 + \tilde{\alpha}_2 X_i - (\tilde{\alpha}_1 + \tilde{\alpha}_2 \bar{X}_i))(Y_i - \bar{Y})}{\sum (\tilde{\alpha}_1 + \tilde{\alpha}_2 X_i - (\tilde{\alpha}_1 + \tilde{\alpha}_2 \bar{X}_i))^2} = \\ &= \frac{\tilde{\alpha}_2 \sum (X_i - \bar{X}_i)(Y_i - \bar{Y})}{\tilde{\alpha}_2^2 \sum (X_i - \bar{X}_i)^2} = \frac{\tilde{\alpha}_2 \tilde{\alpha}_2}{\tilde{\alpha}_2^2} = 1. \end{aligned}$$

Тогда

$$\tilde{\gamma}_1 = \bar{Y} - \tilde{\gamma}_2 \bar{\hat{Y}} = \bar{Y} - 1 \times (\tilde{\alpha}_1 + \tilde{\alpha}_2 \bar{X}) = 0.$$

Поэтому на вопрос в) ответ положительный.

7. (Gujarati [15], 3.17) Рассмотрим уравнение регрессии $Y_i = \beta_1 + u_i$. Используя метод наименьших квадратов, найдите оценку коэффициента β_1 , ее вариацию и показатель RSS . Теперь рассмотрим $Y_i = \beta_1 + \beta_2 X_i + u_i$. Стоит ли добавлять X в модель?

Решение:

По формулам метода наименьших квадратов получим, что

$$\tilde{\beta}_1 = \bar{Y}, \quad \text{Var}(\tilde{\beta}_1) = 1/n \text{Var}(Y_i) = 1/n \text{Var}(u_i) = \frac{\sigma^2}{n} \quad \text{и} \quad \text{RSS} = \sum (\bar{Y} - Y_i)^2.$$

Отсюда следует, что \bar{Y} – несмещенная оценка с минимальной дисперсией для ряда наблюдений Y_i . Можно интерпретировать этот результат следующим образом. Среднее значение является наилучшей оценкой метода наименьших квадратов в условиях полного отсутствия какой-либо информации о взаимосвязи Y с объясняющими переменными. Равенство $RSS=TSS$ влечет равенство коэффициента детерминации нулю, то есть $R^2=0$. Значит стоит добавить какой-либо регрессор X_i в модель, поскольку это повысит коэффициент детерминации R^2 , если, конечно, предполагается наличие экономической взаимосвязи между X и Y .

8. (Gujarati [15], 3.19) На годовых данных с 1980 по 1994 г. были получены следующие результаты оценивания линейной регрессионной модели, где Y – номинальный обменный курс немецкой марки за доллар США, X – отношение индекса потребительских цен в США к показателю в Германии. В скобках под значениями оценок коэффициентов указаны их среднеквадратические отклонения.

$$\hat{Y}_t = 6.682 - 4.318X_t, \quad R^2 = 0.528$$

(1.22) (1.333)

- а) Дайте экономическую интерпретацию регрессии в целом и величины коэффициента детерминации R^2 .
- б) Имеет ли экономический смысл отрицательное значение коэффициента при X_t ? Что об этом говорит экономическая теория?
- в) Допустим, мы переопределили X как отношение индекса потребительских цен в Германии по отношению к США. Как должен измениться знак оценки коэффициента при X и почему?

Решение:

а) Если отношение индекса потребительских цен в США по отношению к этому индексу в Германии увеличится на 1, тогда, в среднем, номинальный обменный курс немецкой марки по отношению к доллару США уменьшится на 4.318 марки за \$1 (то есть немецкая марка подорожает). Коэффициент детерминации R^2 говорит о том, что 52.8% вариации в данных о номинальных обменных курсах объясняется вариацией индекса потребительских цен в США и Германии.

б) Отрицательное значение коэффициента -4.318 имеет экономический смысл. Теория паритета покупательной способности (ППП) утверждает, что в долгосрочном периоде номинальный обменный курс немецкой

марки к доллару США будет пропорционален отношению индексов цен в Германии и США.

в) Следуя теории PPP в этом случае знак коэффициента изменится с отрицательного на положительный.

9. (Gujarati [15], 3.20) В таблице показаны данные по США в 1959-1997 гг. индексов средней производительности труда в деловом секторе экономики ($X1$) и несельскохозяйственном секторе ($X2$), реальной ставке заработной платы – $Y1$, $Y2$ соответственно. В качестве базового года для индексов выбран 1992 г.

а) Постройте графики зависимостей между X и Y .

б) Что говорит экономическая теория о зависимости между X и Y ?

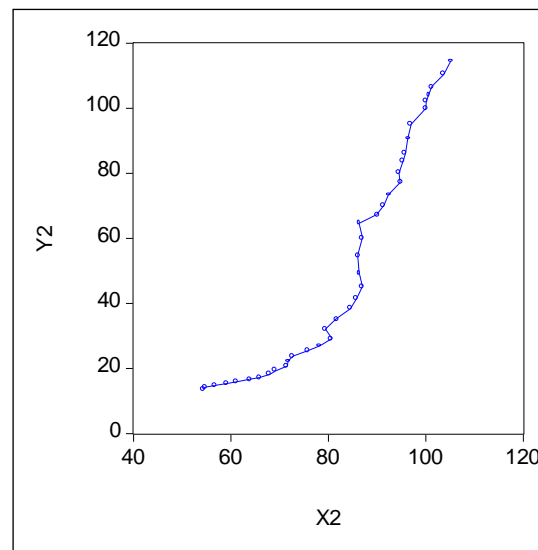
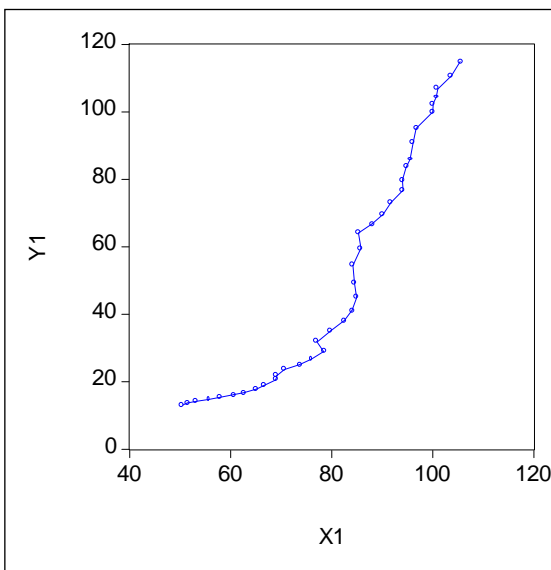
в) Оцените по методу наименьших квадратов регрессию Y на X .

год	$X1$	$X2$	$Y1$	$Y2$
1959	50.5	54.2	13.1	13.7
1960	51.4	54.8	13.7	14.3
1961	53.2	56.6	14.2	14.8
1962	55.7	59.2	14.8	15.4
1963	57.9	61.2	15.4	15.9
1964	60.6	63.8	16.2	16.7
1965	62.7	65.8	16.8	17.2
1966	65.2	68	17.9	18.2
1967	66.6	69.2	18.9	19.3
1968	68.9	71.6	20.5	20.8
1969	69.2	71.7	21.9	22.2
1970	70.6	72.7	23.6	23.8
1971	73.6	75.7	25.1	25.4
1972	76	78.3	26.7	27
1973	78.4	80.7	29	29.2
1974	77.1	79.4	31.8	32.1
1975	79.8	81.6	35.1	35.3
1976	82.5	84.5	38.2	38.4
1977	84	85.8	41.2	41.5
1978	84.9	87	44.9	45.2
1979	84.5	86.3	49.2	49.5
1980	84.2	86	54.5	54.8
1981	85.8	87	59.6	60.2
1982	85.3	86.3	64.1	64.6
1983	88	89.9	66.8	67.3
1984	90.2	91.4	69.7	70.2
1985	91.7	92.3	73.1	73.4

1986	94.1	94.7	76.8	77.2
1987	94	94.5	79.8	80.1
1988	94.7	95.3	83.6	83.7
1989	95.5	95.8	85.9	86
1990	96.1	96.3	90.8	90.7
1991	96.7	97	95.1	95.1
1992	100	100	100	100
1993	100.1	100.1	102.5	102.2
1994	100.7	100.6	104.4	104.2
1995	101	101.2	106.8	106.7
1996	103.7	103.7	110.7	110.4
1997	105.4	105.1	114.9	114.5

Решение:

а)



б) Как известно из курса микроэкономики, реальная заработная плата определяется предельной производительностью труда. В определенных условиях реальная заработная плата будет пропорциональна средней производительности труда с коэффициентом пропорциональности, отражающим долю затрат на оплату труда в общих затратах на факторы производства. В то же время в построенной регрессии зависимой переменной является производительность, а не реальная заработная плата, что искажает причинно-следственную связь.

в) Результаты оценивания в *Eviews* приведены в таблицах.

Dependent Variable: Y1
 Method: Least Squares
 Date: 10/18/03 Time: 18:48
 Sample: 1959 1997
 Included observations: 39

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-109.3833	9.711982	-11.26271	0.0000
X1	2.003875	0.117652	17.03226	0.0000
R-squared	0.886884	Mean dependent var	53.00769	
Adjusted R-squared	0.883827	S.D. dependent var	33.88065	
S.E. of regression	11.54795	Akaike info criterion	7.780813	
Sum squared resid	4934.138	Schwarz criterion	7.866124	
Log likelihood	-149.7259	F-statistic	290.0977	
Durbin-Watson stat	0.070662	Prob(F-statistic)	0.000000	

Dependent Variable: Y2
 Method: Least Squares
 Date: 10/14/03 Time: 22:20
 Sample: 1959 1997
 Included observations: 39

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-123.6000	11.01983	-11.21614	0.0000
X2	2.138592	0.131231	16.29638	0.0000
R-squared	0.877715	Mean dependent var	53.26154	
Adjusted R-squared	0.874410	S.D. dependent var	33.68158	
S.E. of regression	11.93630	Akaike info criterion	7.846966	
Sum squared resid	5271.585	Schwarz criterion	7.932277	
Log likelihood	-151.0158	F-statistic	265.5721	
Durbin-Watson stat	0.076707	Prob(F-statistic)	0.000000	

10. (Gujarati [15], 3.23) В таблице указаны данные о валовом внутреннем продукте США в 1959-1997 гг.

а) Постройте графики ВВП в текущих и постоянных (т.е. для базового 1992 года) ценах.

б) Обозначьте за Y значение ВВП, а за X – время (т.е. $X_1=1, X_2=2, \dots, X_{39}=39$) и оцените модель $Y_i = \beta_1 + \beta_2 X_i + \varepsilon_i$ для ВВП в текущих и постоянных ценах.

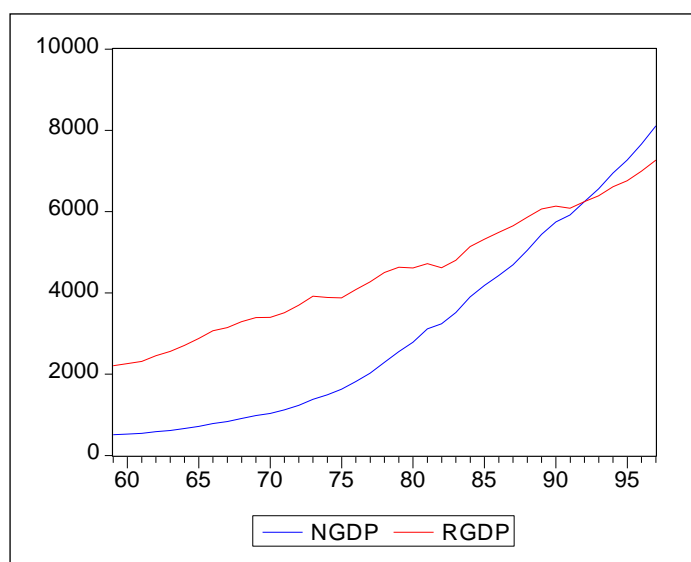
в) Как интерпретировать коэффициент β_2 ?

г) Если присутствует различие между оценками для β_2 в моделях для ВВП в текущих и постоянных ценах, то что объясняет это различие?

Год	ВВП в текущих ценах	ВВП в постоянных ценах	Год	ВВП в текущих ценах	ВВП в постоянных ценах
1959	507.2000	2210.200	1979	2557.500	4630.600
1960	526.6000	2262.900	1980	2784.200	4615.000
1961	544.8000	2314.300	1981	3115.900	4720.700
1962	585.2000	2454.800	1982	3242.100	4620.300
1963	617.4000	2559.400	1983	3514.500	4803.700
1964	663.0000	2708.400	1984	3902.400	5140.100
1965	719.1000	2881.100	1985	4180.700	5323.500
1966	787.7000	3069.200	1986	4422.200	5487.700
1967	833.6000	3147.200	1987	4692.300	5649.500
1968	910.6000	3293.900	1988	5049.600	5865.200
1969	982.2000	3393.600	1989	5438.700	6062.000
1970	1035.600	3397.600	1990	5743.800	6136.300
1971	1125.400	3510.000	1991	5916.700	6079.400
1972	1237.300	3702.300	1992	6244.400	6244.400
1973	1382.600	3916.300	1993	6558.100	6389.600
1974	1496.900	3891.200	1994	6947.000	6610.700
1975	1630.600	3873.900	1995	7269.600	6761.700
1976	1819.000	4082.900	1996	7661.600	6994.800
1977	2026.900	4273.600	1997	8110.900	7269.800
1978	2291.400	4503.000			

Решение:

а) Графики номинального, то есть в текущих ценах (NGDP) и реального, т.е. в постоянных ценах (RGDP) ВВП показаны на рисунке.



б) Результаты оценивания в *Eviews* уравнений регрессий для реального и номинального ВВП показаны в следующих таблицах.

Dependent Variable: RGDP
 Method: Least Squares
 Date: 10/26/03 Time: 14:46
 Sample: 1959 1997
 Included observations: 39

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1907.715	45.13298	42.26875	0.0000
T	128.7820	1.966646	65.48306	0.0000
R-squared	0.991445	Mean dependent var	4483.354	
Adjusted R-squared	0.991214	S.D. dependent var	1474.662	
S.E. of regression	138.2259	Akaike info criterion	12.74558	
Sum squared resid	706937.3	Schwarz criterion	12.83089	
Log likelihood	-246.5387	F-statistic	4288.031	
Durbin-Watson stat	0.470276	Prob(F-statistic)	0.000000	

Dependent Variable: NGDP
 Method: Least Squares
 Date: 10/26/03 Time: 14:47
 Sample: 1959 1997
 Included observations: 39

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-986.3317	212.6816	-4.637597	0.0000
T	201.9772	9.267489	21.79417	0.0000
R-squared	0.927732	Mean dependent var	3053.213	
Adjusted R-squared	0.925779	S.D. dependent var	2390.907	
S.E. of regression	651.3667	Akaike info criterion	15.84594	
Sum squared resid	15698308	Schwarz criterion	15.93125	
Log likelihood	-306.9959	F-statistic	474.9858	
Durbin-Watson stat	0.040102	Prob(F-statistic)	0.000000	

в) Коэффициент β_2 может быть интерпретирован как абсолютный прирост долларах США реального и номинального ВВП.

г) Различие между оценками для β_2 в моделях для ВВП в текущих и постоянных ценах объясняется изменением цен на входящие в состав ВВП товары и услуги.

Англо-русский глоссарий

Данные эконометрические термины часто встречаются в иностранных научных публикациях и учебниках и имеют в литературе на русском языке соответствующие эквиваленты.

adjusted R^2 – скорректированный коэффициент детерминации

autocorrelation function (ACF)– автокорреляционная функция

autoregressive conditional heteroscedasticity (ARCH) – авторегрессионная модель с условной гетероскедастичностью

autoregressive model (AR) – авторегрессионная модель

autoregressive integrated moving average model (ARIMA) –

интегрированная модель авторегрессии-скользящего среднего

best linear unbiased estimator (BLUE) – наилучшая оценка в классе линейных несмещенных оценок (эффективная оценка)

binary variable – дискретная бинарная переменная, которая может принимать значения 0 или 1

censored model – модель, основанная на цензурированной выборке, в которой зависимая переменная ограничена некоторым пороговым значением

classical normal regression (CNR) – классическая регрессионная модель, ошибки которой имеют совместное нормальное распределение

classical regression (CR) – классическая регрессионная модель

coefficient of determination (R-squared) – коэффициент детерминации

conditional distribution – условное распределение

confidence interval – доверительный интервал

consistent estimator – состоятельная оценка

convergence in distribution – сходимость по распределению

correlation – корреляция

correlation coefficient – коэффициент корреляции

count data – счетные данные

covariance – ковариация

cross-section data – пространственные данные

density function – функция плотности распределения

dependent variable – зависимая (объясняемая) переменная

distributed lags model – модель распределенных лагов

distribution – распределение (функция распределения)

dummy variable – фиктивная (искусственная) переменная

duration model – модель “времени жизни”

efficient estimator – эффективная оценка

endogenous variable – эндогенная переменная, т.е. переменная, определяемая внутри модели

error – ошибка регрессии

estimator – оценка

exogenous variable – экзогенная, внешняя по отношению к модели переменная

explanatory variables – объясняющие (независимые) переменные

exponential smoothing – экспоненциальное сглаживание

fitted value – прогнозное значение

forecast – прогноз

generalized least squares (GLS) estimation – обобщенный метод наименьших квадратов

goodness-of-fit– качество подгонки кривой

hazard rate – интенсивность отказов

heteroscedasticity– гетероскедастичность

homoscedasticity – гомоскедастичность

idempotent matrix – идемпотентная матрица

independent variable – независимая переменная

index function – индексная функция

indirect least squares – косвенный метод наименьших квадратов

information matrix – информационная матрица

instrumental variable (IV) – инструментальная переменная

intercept – свободный член (константа регрессии)

joint distribution – совместное распределение

lag operator – оператор лага (оператор сдвига по времени)

lagged variable – запаздывающая (лаговая) переменная

latent variable – скрытая, ненаблюдаемая переменная

law of large numbers – закон больших чисел

likelihood function – функция правдоподобия

linear probability model – линейная модель вероятности

linear regression model – линейная регрессионная модель

logit model – логит-модель, нелинейная модель дискретной (бинарной) зависимой переменной, основанная на логистическом распределении ошибки

loglikelihood function – логарифм функции правдоподобия

marginal distribution – предельное распределение, распределение одной или нескольких компонент случайного вектора

maximum likelihood (ML) – метод максимального правдоподобия

maximum likelihood estimate – оценивание по методу максимального правдоподобия

maximum likelihood estimator – оценка метода максимального правдоподобия

maximum score estimator (MSCORE) – оценивание по методу максимального счета

mean – математическое ожидание (среднее значение)

mean absolute deviation – среднее абсолютное отклонение

mean absolute percentage error – среднее относительное отклонение

mean squared error – среднеквадратическая ошибка

model for binary choice – модель бинарного выбора

model for multiple choice – модель множественного выбора

model specification – спецификация модели

moving average – скользящее среднее

moving average (MA) model – модель скользящего среднего

multicollinearity – мультиколлинеарность

multiple regression model – модель множественной регрессии

normal (Gaussian) distribution – нормальное (гауссовское) распределение

OLS-estimator – оценка метода наименьших квадратов

omitted variable – пропущенная (невключенная в модель) независимая переменная

ordered data – упорядоченные данные

ordinary least squares (OLS) method – метод наименьших квадратов

panel data – панельные (матричные) данные, сочетание временных и пространственных данных

partial autocorrelation function (PACF) – частная автокорреляционная функция

partial correlation coefficient – частный коэффициент корреляции

probit model – пробит-модель, нелинейная модель дискретной (бинарной) зависимой переменной, основанная на нормальном распределении ошибки

qualitative variable – качественная переменная

random utility model – модель случайной полезности

random walk – процесс случайного блуждания

ranking variable – ординальная, порядковая, ранговая, переменная

reduced form of the model – приведенная форма модели

residuals – остатки регрессии

restricted regression – регрессия с ограничениями на параметры

sample – выборка

sample mean (variance, covariance, moment etc.) – выборочное среднее (дисперсия, ковариация, момент и т. д.)

seemingly unrelated regression (SUR) – система внешне несвязанных уравнений

selection model – модель, основанная на случайно усеченной выборке

serial correlation – корреляция между показателями, относящимися к разным моментам времени

series – ряд данных (наблюдений)

significance level – уровень значимости

simultaneous equations – одновременные уравнения

slope – коэффициент при регрессоре в парной регрессии

standard deviation – среднеквадратическое отклонение (квадратный корень из дисперсии)

stationary time series – стационарный временной ряд

strictly stationary process – строго стационарный процесс, стационарный в узком смысле процесс

time-series data – временные данные

truncated model – модель, построенная для усеченной выборки (из которой исключены некоторые наблюдения)

two-stage least squares (TSLS, 2SLS) – двушаговый метод наименьших квадратов

unbiased estimator – несмещенная оценка

unrestricted regression – модель без ограничений на параметры

variance – дисперсия

variance (covariance) matrix – ковариационная матрица

weighted least squares – взвешенный метод наименьших квадратов

white noise – «белый шум», процесс случайного блуждания с независимыми одинаково распределенными значениями с нулевыми средними

Список литературы

Основная литература

1. (Айвазян) Айвазян С.А., Мхитарян В.С. Прикладная статистика. Основы эконометрики. Т. 2. – М.: ЮНИТИ, 2001.
2. (Берндт) Берндт Э. Практика эконометрики. – М.: ЮНИТИ-ДАНА, 2005.
3. (Магнус) Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика. Начальный курс. 6-е изд. – М.: Дело, 2004.

Задачники, практикумы

4. Катышев П.К., Магнус Я.Р., Пересецкий А.А. Сборник задач к начальному курсу эконометрики. 2-е изд. – М.: Дело, 2003.
5. Практикум по эконометрике: учеб. Пособие / Под ред. И.И. Елисеевой. – М.: Финансы и статистика, 2003.
6. Методические материалы по экономическим дисциплинам для преподавателей средних школ и вузов: программы, тесты, задачи, решения / Под общ.ред. Л.С. Гребнева. - М.: ГУ ВШЭ, 2000.
7. Колеников С.О. Прикладной эконометрический анализ в статистическом пакете Stata / Российская экономическая школа. – М., 2001.
8. Молчанов И.Н., Герасимова И.А. Компьютерный практикум по начальному курсу эконометрики (реализация на *Eviews*) / Ростовский государственный экономический университет. – Ростов-н/Д., 2001.

Дополнительная литература

9. Арженовский С.В., Федосова О.Н. Эконометрика / Ростовский

- государственный экономический университет. – Ростов-н/Д., 2002.
10. Доугерти К. Введение в эконометрику. – М.: ИНФРА-М, 1999.
 11. Носко В.П. Эконометрика для начинающих / Институт экономики переходного периода. – М., 2000.
 12. Тихомиров Н.П., Дорохина Е.Ю. Эконометрика / Российская экономическая академия имени Г.В. Плеханова. – М., 2002.
 13. Уотшем Т., Паррамоу К. Количественные методы в финансах: Учеб. пособие для вузов. – Гл. 5-7. – М.: Финансы, ЮНИТИ, 1999.
 14. Эконометрика: учебник / Под ред. И.И. Елисеевой. – М.: Финансы и статистика, 2002.
 15. Gujarati D. Basic econometrics. 4th ed. McGraw Hill, 2003. – Режим доступа: <http://www.mhhe.com/economics/gujarati4/>.
 16. Maddala G. Introduction to econometrics. 3rd ed. Wiley, 2001. – Режим доступа: <http://www.wiley.co.uk/maddala/>.
 17. Verbeek M. A guide to modern econometrics. Wiley, 2000.
 18. Wooldridge J. Introductory econometrics. 2nd ed. South-Western, 2003. – Режим доступа: <http://www.swlearning.com/economics/wooldridge/wooldridge2e/wooldridge2e.html>.

Литература продвинутого уровня

19. Greene W. Econometric analysis. 5th ed. Prentice Hall, 2003. – Режим доступа: <http://pages.stern.nyu.edu/~wgreene/Text/econometricanalysis.htm>.
20. Handbook on econometrics. Vols. 1-5. North-Holland, 1983-2000.
21. Wooldridge J. Econometric analysis of cross section and panel data. MIT Press, 2002. – Режим доступа: <http://www.msu.edu/~ec/faculty/wooldridge/book2.htm>.

Содержание

Предисловие.....	3
Учебная программа и содержание курса.....	5
Рейтинг-план дисциплины.....	9
Вопросы к теоретическому экзамену.....	10
Вопросы для самоконтроля к теоретической части курса.....	11
Лабораторные работы и рекомендации к их выполнению.....	13
Лабораторные работы для самостоятельного выполнения.....	50
Практические задания и указания к их выполнению.....	52
Англо-русский глоссарий.....	62
Список литературы.....	68

Учебное издание

Величко Андрей Сергеевич

**ИЗУЧАЕМ ЭКОНОМЕТРИКУ.
НАЧАЛЬНЫЙ КУРС**

Учебное пособие

В авторской редакции

Подписано в печать 26.09.2007.

Формат 60 × 84 $\frac{1}{16}$. Усл. печ. л. 4,18. Уч.-изд. л. 4,20.

Тираж 100 экз.

Издательство Дальневосточного университета

690950, г. Владивосток, ул. Алеутская, 56

Отпечатано на множительной технике ИМКН ДВГУ

690950, г. Владивосток, ул. Октябрьская, 27

ДЛЯ ЗАМЕТОК