



НАРОДНАЯ УКРАИНСКАЯ АКАДЕМИЯ

ЭКОНОМЕТРИКА

Практикум и методические рекомендации

Издательство НУА

НАРОДНАЯ УКРАИНСКАЯ АКАДЕМИЯ

ЭКОНОМЕТРИКА

Практикум и методические рекомендации
*для студентов факультета «Бизнес-управление»,
обучающихся по направлению подготовки
6.030504 «Экономика предприятия»*

Харьков
Издательство НУА
2017

УДК 004.383.1(072+078.5)

ББК 32.973.26-018.2р30

Э40

*Утверждено на заседании
кафедры информационных технологий и математики
Народной украинской академии.
Протокол № 10 от 16.05.2016*

Р е ц е н з е н т : канд. техн. наук *К. С. Барашев*

А в т о р ы - с о с т а в и т е л и : *С. Б. Данилевич, О. В. Дьячкова*

Э40 Эконометрика: метод. указания и материалы к самостоятельной работе для студентов фак. «Бизнес-управление», обучающихся по направлению подготовки 6.030504 «Экономика предприятия» / Нар. укр. акад. [каф. информ. технологий и математики ; авт.-сост. С. Б. Данилевич, О. В. Дьячкова]. – Харьков : Изд-во НУА, 2017. –76 с.

Методические рекомендации раскрывают содержание раздела «Эконометрика» дисциплины «Экономико-математические методы и модели». Издание включает программу и тематический план, указаны формы контроля и оценивание, приведены методические рекомендации по изучению тем и выполнению заданий, вопросы для самопроверки, основные формулы и справочные таблицы, описание необходимых модулей и функций MS Excel, словарь основных терминов, список рекомендуемой литературы.

Методичні рекомендації розкривають зміст розділу «Економетрика» дисципліни «Економіко-математичні методи та моделі». Видання включає програму і тематичний план, вказані форми контролю та оцінювання, наведено методичні рекомендації щодо вивчення тем і виконання завдань, питання для самоперевірки, основні формули та довідкові таблиці, опис необхідних модулів і функцій MS Excel, словник основних термінів, список рекомендованої літератури.

УДК 004.383.1(072+078.5)

ББК 32.973.26-018.2р30

© Народная украинская академия, 2017

ОБЩИЕ ПОЛОЖЕНИЯ

Раздел «Эконометрика» дисциплины «Экономико-математические методы и модели» изучается согласно учебному плану подготовки специалистов образовательно-квалификационного уровня «бакалавр» направления подготовки 6.030504 «Экономика предприятия».

Предмет эконометрики – количественная оценка зависимостей между экономическими величинами, то есть разработка моделей и их использование в научных исследованиях и управлении экономическими системами.

Цель дисциплины «Эконометрика» – научить подходам к формированию эконометрических моделей различных уровней и содержания, принципам эконометрического исследования.

Основные задачи эконометрики:

- построение эконометрических моделей;
- оценка параметров построенной модели, при которых она наиболее адекватна реальным данным;
- исследование качества найденных параметров модели и самой модели в целом;
- использование построенных моделей для объяснения поведения исследуемых экономических показателей, прогнозирования, осмысленного проведения экономической политики.

ОРГАНИЗАЦИОННО-МЕТОДИЧЕСКИЕ УКАЗАНИЯ

Раздел изучается на 3 курсе и состоит из разделов:

- раздел 1 «Парная регрессия»
- раздел 2 «Множественный регрессионный анализ»;
- раздел 3 «Моделирование динамических процессов».

В результате изучения дисциплины студент должен знать:

- цели и задачи эконометрики;
- область и степень ее применимости;
- основные положения, теоретические основы и методы;
- информационные системы поддержки эконометрических исследований и расчетов;
- основы регрессионного анализа;

- основы статистического оценивания и анализа точности параметров уравнения регрессии;
- предпосылки для применения классических регрессионных моделей;
- основы анализа и прогнозирования временных рядов;

уметь:

- решать типовые задачи;
- использовать основные приемы эконометрического исследования эмпирических данных;
- самостоятельно поработать учебно-методической литературой компьютерными программами при анализе применимости эконометрических методов;

иметь представление:

- об основных статистических инструментах, методах и способах обработки экономических данных.

Учебный материал базируется на теоретических знаниях и практических навыках, полученных в курсах информатики, теории вероятностей, математической статистики, твердых познаниях в области экономической теории.

Полученные в результате овладения материалом дисциплины умения и навыки могут быть применены в дальнейшем при изучении профессионально-ориентированных и специализированных дисциплин, при выполнении расчетных заданий, лабораторных, курсовых и дипломных работ.

Дисциплина имеет практическую направленность. Для достижения необходимого уровня практических знаний и умений предполагается проведение практических занятий в компьютерных аудиториях ХГУ «НУА», самостоятельное выполнение лабораторных и контрольных заданий. Для индивидуального изучения студентам выдаются список тем по дисциплине и вопросы к ним, список учебно-методической и справочной литературы, задания для самостоятельной работы.

Общий объем дисциплины:

	Наименование темы	Количество часов		
		Всего	Практические занятия	Самостоятельная работа
1.	Парная регрессия			
1.1.	Линейная регрессия	18	6	12
1.2.	Нелинейная регрессия	18	6	12
2.	Множественный регрессионный анализ			
2.1.	Многофакторные модели	18	6	12
2.2.	Особые случаи	18	8	12
3.	Динамические процессы			
3.1.	Автокорреляция	16	4	12
3.2.	Моделирование динамических процессов	18	6	12
	Всего за семестр	106	36	72

Проверка теоретических знаний и практических навыков предусматривает:

- текущий контроль во время занятий, при проверке выполнения практических работ;
- тестирование с использованием компьютерных обучающе-контролирующих программ;
- текущий контроль в виде контрольных заданий;
- итоговый экзаменационный контроль.

С целью обеспечения максимальной эффективности практического освоения материала дисциплины на каждом занятии и во время индивидуальной работы каждый студент группы обеспечивается в компьютерном классе индивидуальным рабочим местом.

Академические успехи студента определяются по рейтинговой 100-балльной шкале ХГУ «НУА», а итоговые оценки – по 5-балльной национальной шкале. Соотношения оценок национальной шкалы, рейтинговой шкалы ХГУ «НУА» и шкалы Европейской кредитно-трансферной и аккумулирующей системы (ECTS) выглядит следующим образом:

Оценка по шкале ECTS	Обозначение по шкале ECTS	Обозначение по национальной шкале	По шкале ХГУ НУА
A	ОТЛИЧНО – отличное выполнение лишь с незначительным количеством ошибок	5 (отлично)	85–100 баллов
B	ОЧЕНЬ ХОРОШО – выше среднего уровня с несколькими ошибками	4 (хорошо)	75–84 баллов
C	ХОРОШО – в общем правильная работа с определенным количеством значительных ошибок		65–74 баллов
D	УДОВЛЕТВОРИТЕЛЬНО – неплохо, однако с большим количеством недостатков	3 (удовлетворительно)	57–64 баллов
E	ДОСТАТОЧНО – выполнение удовлетворяет минимальным критериям		50–56 баллов
FX	НЕУДОВЛЕТВОРИТЕЛЬНО – необходимо поработать перед тем, как пересдать дисциплину	неудовлетворительно	25–49 баллов
F	НЕУДОВЛЕТВОРИТЕЛЬНО – необходима дальнейшая серьезная работа, с обязательным повторным изучением дисциплины		0–24 баллов

Общая сумма баллов за каждый раздел дисциплины складывается из баллов, полученных за соответствующие темы. Максимальное количество баллов за тему выставляется при безошибочном и своевременном выполнении практических работ и сдаче тестов с использованием компьютерных контролирующих программ на оценку «отлично». Сумма максимальных баллов по всем работам составляет 100 баллов.

Для успешного освоения и сдачи дисциплины необходимо набрать за семестр не менее 50 баллов по рейтинговой 100-балльной шкале при условии выполнения и успешной сдачи всех предусмотренных программой работ. Полученное количество баллов учитывается при итоговой аттестации студента по дисциплине.

Академическая аттестация учитывает результаты освоения и сдачи материала разделов дисциплины, а также итоговый контроль, проводимый в экзаменационную сессию. Итоговая аттестационная оценка складывается из трех компонент – количества баллов за текущую работу (вклад – 70%), экзаменационная оценка (вклад – 30%). Максимально возможное итоговое количество баллов для аттестации – 100 баллов.

ТЕМАТИЧЕСКИЙ ПЛАН

СОДЕРЖАНИЕ РАЗДЕЛА 1

«Парная регрессия»

Тема 1.1. Линейная регрессия.

Основы эконометрического моделирования. Статистическая база. Принципы построения эконометрических моделей. Парная линейная регрессия. Метод наименьших квадратов. Интерпретация уравнения регрессии. Коэффициент детерминации. Проверка на значимость коэффициентов регрессии и адекватность модели в целом.

Тема 1.2. Нелинейная регрессия.

Преобразование переменных (базисная процедура). Нелинейные модели, сводящиеся к линейным. Логарифмические преобразования. Коэффициент эластичности. Случайные ошибки. Сравнение моделей.

Тематический план раздела 1

Наименование темы		Количество часов		
		Всего	Практические занятия	Самостоятельная работа
1.1.	Линейная регрессия	18	6	12
1.2.	Нелинейная регрессия	18	6	12
	Всего	36	12	24

СОДЕРЖАНИЕ РАЗДЕЛА 2

«Множественный регрессионный анализ»

Тема 2.1. Многофакторные модели.

Теорема Гаусса – Маркова. Этапы и методы построения моделей множественной регрессии. Спецификация переменных в уравнениях регрессии. Многофакторные линейные модели. Множественная регрессия в нелинейных моделях. Свойства коэффициентов множественной регрессии. Качество оценки.

Тема 2.2. Особые случаи.

Обобщенные эконометрические модели. Особые случаи в эконометрических моделях (гетероскедастичность, мультиколлинеарность, автокорреляции). Фиктивные переменные. Множественные совокупности фиктивных переменных. Фиктивные коэффициенты для коэффициента наклона. Тест Чоу.

Тематический план раздела 2

Наименование темы		Количество часов		
		Всего	Практические занятия	Самостоятельная работа
2.1.	Многофакторные модели	18	6	12
2.2.	Особые случаи	18	8	12
	Всего	36	14	24

СОДЕРЖАНИЕ РАЗДЕЛА 3

«Динамические процессы»

Тема 3.1. Автокорреляция.

Эконометрические модели динамики. Автокорреляция уровней временного ряда. Автокорреляция в остатках. Критерий Дарбина – Уотсона. Оценивание коэффициентов регрессии при наличии автокорреляции в остатках.

Тема 3.2. Моделирование динамических процессов.

Моделирование тенденции, сезонных и циклических колебаний. Использование фиктивных переменных при наличии структурных изменений.

Тематический план раздела 3

Наименование темы		Количество часов		
		Всего	Практические занятия	Самостоятельная работа
3.1.	Автокорреляция	16	4	12
3.2.	Моделирование динамических процессов	18	6	12
	Всего	34	10	24

РАЗДЕЛ 1. Парная регрессия

ТЕМА 1.1. ЛИНЕЙНАЯ РЕГРЕССИЯ

Корреляционно-регрессионный анализ является основным в изучении взаимосвязей явлений. Данный метод содержит две составляющие части – корреляционный анализ и регрессионный анализ.

Корреляционный анализ – раздел математической статистики, изучающий тесноту связи (с помощью расчета коэффициентов корреляции) между переменными без их разделения на факторные и результативные.

Регрессионный анализ – раздел математической статистики, изучающий форму зависимости между факторными и результативными переменными.

Диаграмма рассеяния (корреляционное поле) – графическое представление корреляционной связи (где оси соответствуют обоим переменным).

Обобщенная эконометрическая модель:

$$y = f(\alpha, x) + \varepsilon$$

где y – результативный признак;

$f(\alpha, x)$ – функционал, выражающий вид и структуру взаимосвязей;

$x = (x_1, x_2, \dots, x_n)$ – факторные переменные;

$\alpha = (\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_n)$ – произвольные константы (*параметры модели*);

ε – ошибка модели, или возмущение.

Функционал $f(\alpha, x)$ задает теоретическое значение результативного признака (*Уоценочное*, или \hat{y}), тогда как y – его фактическое значение.

Регрессионные модели:

- **парная (простая) регрессия** – регрессия между двумя переменными – y и x , т. е. модель вида $y = f(x)$, где y – результативный признак, x – признак-фактор. Парная регрессия строится, если имеется доминирующий фактор – его и используют в качестве объясняющей переменной.
- **множественная регрессия** – регрессия результативного признака с двумя и большим числом факторов, т. е. модель вида $y = f(x_1, x_2, \dots, x_n)$.

Спецификация модели – запись конкретного вида модели, исходя из соответствующей теории связи между переменными.

Регрессионные модели с одним уравнением:

- линейная: $y = a + b \cdot x + \varepsilon$.
 - степенная: $y = a \cdot x^b \cdot \varepsilon$.
 - гиперболическая: $y = a + b/x + \varepsilon$.
 - полиномиальная: $y = a + b \cdot x + c \cdot x^2 + d \cdot x^3 + \dots + \varepsilon$
 - экспоненциальная: $y = e^{a+b \cdot x} \cdot \varepsilon$
- и другие.

Статистические характеристики – показатели, характеризующие последовательности $x = (x_1, x_2, \dots, x_n)$ и $y = (y_1, y_2, \dots, y_n)$ (см. Приложение 1).

Линейная регрессионная модель:

$$y = b_0 + b_1 \cdot x + \varepsilon$$

где y – результативный признак;

$x = (x_1, x_2, \dots, x_n)$ – факторные переменные;

b_0, b_1 – неизвестные параметры регрессионной модели;

ε – ошибка модели.

Задачи линейного регрессионного анализа: а) получить наилучшие оценки параметров b_0, b_1 ; б) проверить статистические гипотезы о параметрах модели; в) проверить адекватность модели исходным данным.

Наиболее распространенный метод нахождения коэффициентов – **метод наименьших квадратов (МНК)**. Оценки коэффициентов, найденных этим методом, имеют оптимальные свойства.

Для оценки адекватности принятой эконометрической модели экспериментальным данным используется **F-критерий Фишера**.

Для проверки значимости каждого параметра регрессии применяют **t-критерий Стьюдента**.

Формулы, а также функции MSExcel для их расчета, приведены в *Приложении 1*.

С помощью **надстройки MSExcel Пакет анализа данных** можно получить результаты регрессионной статистики, дисперсионного анализа и доверительных интервалов, а также остатки и графики подбора линии регрессии, остатков и нормальной вероятности и др. Описание применения надстройки приведено в *Приложении 2*.

Задания к теме 1.1. Линейная регрессия

Задание 1. Анализ зависимости между переменными

1. В файле *Фамилия ЭММ1* проанализируйте деятельность двух менеджеров, контролирующих остатки товаров в двух Ваших магазинах (см. табл. 1). Кто из них лучше управляет запасами либо оба работают одинаково?¹

Таблица 1

Стоки последних шести недель:

магазин 1	33	31	32	36	31	31
магазин 2	22	34	58	52	10	21

¹ Формулы для расчетов см. в Приложении 1.

а) Рассчитайте и сравните средние значения стоков. Как они характеризуют работу менеджеров?

б) Рассчитайте и сравните стандартные отклонения. Как они характеризуют работу менеджеров?

в) Рассчитайте и сравните коэффициенты вариации. Как они характеризуют работу менеджеров?

Сравните (письменно) работу обоих менеджеров.

2. Дана таблица производства и потребления:

Таблица 2

Поголовье КРС (млн т)	Объем производства молока (тыс. т)	Объем производства мяса (тыс. т)	Потребление мяса (кг на душу населения в год)
57	1,49	8,37	69
54,7	1,38	8,26	60
52,2	1,29	7,51	69
48,9	1,1	6,8	57
43,3	0,99	5,79	55
39,7	0,9	5,33	51
35,1	0,88	4,85	50

2.1. Проанализируйте потребление мяса:

а) Рассчитайте среднее, выборочную дисперсию, стандартное отклонение, коэффициент вариации. Интерпретируйте эти показатели и сделайте вывод (письменно) об однородности данных.

б) Вызовите *Пакет анализа / Описательная статистика*, выберите итоговую статистику. Сравните (устно) результаты с вашими расчетами.

2.2. Выясните, существует ли связь между поголовьем КРС и производством молока (письменно):

а) Рассчитайте для этого ковариацию и коэффициент корреляции. Интерпретируйте эти показатели.

б) Постройте диаграмму рассеяния.

в) Вызовите *Пакет анализа / Ковариация*, а затем *Корреляция* для тех же данных. Сравните (устно) результаты с вашими расчетами.

2.3. Проведите анализ зависимости между собой всех исходных данных таблицы 2 с помощью *Пакета анализа / Корреляция*.

Сделайте письменные выводы.

Результаты работы сохраните в своей папке и скопируйте в папку \$control.

Задание 2. Статистическая база эконометрических моделей

Создайте новую книгу Excel. Указитена первом (титульном) листе свою фамилию, группу, название работы, дату выполнения. Файл сохраните под названием *Фамилия ЭММ 2 Статистическая база* в своей папке.

1. Выберите исходные данные для своего варианта из Приложения 2. Скопируйте их в свой файл на лист *База*.
2. По исходным данным постройте корреляционное поле (=диаграмму рассеяния) с помощью *Мастер диаграмм / Точечная*. За переменную X возьмите потребительские расходы, Y – валовый внутренний продукт страны.
3. Для данных X и Y рассчитайте средние значения, стандартные отклонения, коэффициенты вариации.
4. Под исходными данными рассчитайте коэффициенты ковариации и корреляции X и Y .
5. Проведите анализ с помощью модулей *Ковариация* и *Корреляция* пакета анализа.
6. Сделайте (письменно) вывод о корреляционной связи между X и Y .

Результаты работы сохраните в своей папке и скопируйте в папку *\$control*.

Задание 3. Построение линейной эконометрической модели

Скопируйте Ваш файл предыдущей работы *Фамилия ЭММ2.xlsx*, назвав копию *Фамилия ЭММ3 Линейная регрессия*. В этом новом файле:

укажите на титульном листе текущую дату и тему работы (ниже предыдущих);

сделайте копию листа *База*, назвав его *Линейный тренд*.

1. На листе *Линейный тренд*:
 - 1.1. для имеющейся точечной диаграммы постройте линейный тренд. Задайте отображение уравнения на диаграмме. Настройте масштаб отображения шкалы на диаграмме, так чтобы обе линии занимали всю область построения диаграммы. Задайте для линий цвета: синий и красный.
 - 1.2. Ниже диаграммы выпишите (или скопируйте) уравнение полученного линейного тренда.
2. (доп.) Добавьте новый лист *Метод МНК*.
 - 2.1. Скопируйте на него исходные данные листа *База* (столбцы X и Y).
 - 2.2. Скопируйте из презентации *ЭММ2.pptx* формулы (изображения) расчета коэффициентов линейной регрессии.
 - 2.3. Рассчитайте по этим формулам значения коэффициентов.
 - 2.4. Ниже выпишите уравнение линейной регрессии с полученными коэффициентами.

3. (доп.) Добавьте новый лист *ЛИНЕЙН*.
 - 3.1. Скопируйте на него исходные данные (столбцы X и Y) с предыдущего листа.
 - 3.2. Рассчитайте коэффициенты линейной регрессии с помощью встроенной статистической функции *ЛИНЕЙН*. (Подсказка. Не забудьте, что это функция массивов, вводится с помощью *Ctrl+Shift+Enter*. Выделите предварительно две соседние ячейки строки и укажите в качестве последнего параметра функции значение *Ложь*). Каждый полученный коэффициент снабдите пояснительной надписью.
 - 3.3. Ниже выпишите уравнение линейной регрессии с полученными коэффициентами.
4. Добавьте новый лист *Линейная*. Задайте цвет ярлычку этого листа.
 - 4.1. Скопируйте на лист *Линейная* исходные данные (столбцы X и Y) с предыдущего листа.
 - 4.2. С помощью инструмента пакета анализа данных *Регрессия* проведите регрессионный анализ данных. Результаты расположите ниже на том же листе.
 - 4.3. Ячейки, содержащие рассчитанные значения коэффициентов регрессии, выделите цветом. Снабдите их примечанием с описанием каждого коэффициента (см. рис. 9).
 - 4.4. Ниже выпишите уравнение линейной регрессии с полученными коэффициентами. Укажите в соседней ячейке, что это за уравнение.
 - 4.5. Найдите среди результатов ячейку, содержащую значение корреляции, выделите ее цветом. Снабдите ее также примечанием с описанием (см. рис. 9). Задайте постоянное отображение примечаний.
5. Сравните ниже (письменно) значения коэффициентов регрессии и коэффициента корреляции, полученные на листе *Линейная* и на других листах.
Результат сохраните в своей папке и скопируйте в папку *\$control*.

Задание 4. Оценка значимости параметров и адекватности линейной эконометрической модели

Скопируйте Ваш файл предыдущей работы *Фамилия ЭММ3.xlsx*, назвав копию *Фамилия ЭММ4 Параметры регрессии*. В этом новом файле дополните титульный лист текущей датой и темой работы (ниже предыдущих).

1. На листе *Линейная* выделите цветом и снабдите примечаниями ячейки (см. рис. 9):
 - коэффициента детерминации;
 - F -критерия Фишера; его значимости;
 - рассчитанных коэффициентов b_0 , b_1 ;
 - $t(b_0)$, $t(b_1)$; их значимости.

Задайте постоянное отображение примечаний.

2. Напишите вывод по коэффициенту корреляции и коэффициенту детерминации (см. рис. 9).
3. Оцените адекватность модели в целом (значимость коэффициента детерминации). Напишите вывод.
4. Оцените значимость параметров b_0 , b_1 . Напишите вывод. (В случае незначимости параметра b_0 провести ниже регрессионный анализ повторно, задав при этом опцию *Константа ноль*. Скопировать все примечания в новые результаты, написать заново выводы п. 2 и 3.) (Для копирования только примечаний используйте *Специальную вставку*.)
5. Добавьте к исходной таблице данных (столбцам X и Y) столбец $\hat{y} = y_{оцен.}$. Рассчитайте его значения по полученной формуле регрессии.
6. На одной диаграмме постройте корреляционное поле данных (x, y) и прямую $(x, y_{оцен.})$. Задайте отображение: для первого – маркеры, для второй – линию с маркерами.
Сравните с диаграммой на листе *Линейный тренд*(устно).
7. Добавьте внизу таблицы исходных данных строку, указав в ней: *Год* – текущий год, *X* – значение на 10% большее среднего всех значений *X*.
Сделайте линейный прогноз *Y* для этого значения *X* (=рассчитайте $y_{оцен.}$).
Результат сохраните в своей папке и скопируйте в папку *\$control*.

ТЕМА 1.2. НЕЛИНЕЙНАЯ РЕГРЕССИЯ

Нелинейная регрессия — частный случай регрессионного анализа, в котором рассматриваемая регрессионная модель есть функция, зависящая от параметров и от одной или нескольких свободных переменных. Зависимость от параметров предполагается нелинейной.

Выбор вида регрессии

Для парной регрессии применяют три способа выбора вида математической функции:

- графический
- аналитический (исходя из теории изучаемой взаимосвязи);
- экспериментальный

Виды нелинейных моделей

- регрессии, нелинейные по включенным в анализ объясняющим переменным, но линейные по оцениваемым параметрам;
- регрессии, нелинейные по оцениваемым параметрам.

Регрессии, нелинейные по переменным

Примеры моделей, нелинейных по переменным

- Полиномиальная: $y = b_0 + b_1x + b_2x^2 + \varepsilon$
 $y = b_0 + b_1x + b_2x^2 + b_3x^3 + \varepsilon$

Парабола 2-й степени обычно применяется, когда на некотором интервале меняется характер связи переменных – например, прямая связь меняется на обратную и наоборот.

- Гиперболическая: $y = b_0 + \frac{b_1}{x} + \varepsilon$

Гиперболу используют для характеристики связи удельных расходов сырья, материалов, топлива от объема выпуска, времени обращения товаров от товарооборота, прироста зарплаты от уровня безработицы (кривая Филлипса), расходов на непродовольственные товары от доходов или общей суммы расходов (кривые Энгеля) и др.

- Квадратного корня: $\hat{y} = b_0 + b_1 \cdot \sqrt{x} + \varepsilon$
- Полулогарифмическая: $\hat{y} = b_0 + b_1 \cdot \ln x + \varepsilon$

Полулогарифмические модели используются, когда надо определить темп роста или прироста экономических показателей (например, объема выпуска от относительных затрат ресурсов, дефицита бюджета от темпа роста ВВП, темпа роста инфляции от объема денежной массы и др.).

Линеаризация моделей, нелинейных по переменным

Для применения МНК (метода наименьших квадратов) должно выполняться условие линейности модели по ее коэффициентам. Поэтому, если такое возможно, нелинейные модели стараются преобразовать к линейным и применить МНК.

Регрессию, нелинейную по переменным, но линейную по параметрам, можно свести к линейной регрессии с помощью замены переменных:

$$\hat{y} = b_0 + b_1 \cdot \varphi(x) \quad \Leftrightarrow \quad z = \varphi(x) \quad \Leftrightarrow \quad \hat{y} = b_0 + b_1 \cdot z$$

Регрессии, нелинейные по параметрам

Некоторые модели, нелинейные по параметрам, также могут быть сведены к линейным. После этого можно применить МНК к преобразованным уравнениям.

Примеры нелинейных по параметрам моделей, внутренне линейных

- Степенная: $y = b_0 \cdot x^{b_1} \cdot \varepsilon$.

Степенная модель широко используется, т.к. b_1 имеет четкое экономическое объяснение – это коэффициент эластичности. К степенным относятся кривые спроса и предложения, производственные функции, зависимость ВВП от уровня занятости и др.

- Показательная: $y = b_0 \cdot b_1^x \cdot \varepsilon$

Показательную регрессию применяют, если показатели характеризуются приблизительно постоянным темпом относительного прироста во времени.

- Экспоненциальная: $y = e^{b_0 + b_1 x} \cdot \varepsilon$.
- Логистическая: $y = \frac{b_0}{1 + b_1 e^{-b_2 x + \varepsilon}}$

Логистическая модель применяется для показателей, имеющих уровни «насыщения»: зависимости спроса на товар от уровня дохода, роста численности населения и т.д.

- Обратная: $y = \frac{1}{b_0 + b_1 x + \varepsilon}$.

Примеры моделей, нелинейных по параметрам и внутренне нелинейных

Не приводятся к линейному виду.

$$y = b_0 + b_1 x^{b_2};$$

$$y = 1 - \frac{1}{1 - x^b} \text{ и др.}$$

Линеаризация моделей, нелинейных по параметрам, внутренне линейных

Нелинейная парная регрессия иногда может быть приведена к линейной с помощью логарифмирования:

шаг 1) логарифмирование:

$$\hat{y} = b_0 \cdot b_1^x \quad \Leftrightarrow \quad \ln \hat{y} = \ln b_0 + x \ln b_1 \quad (\text{для показательной})$$

$$\hat{y} = b_0 \cdot x^{b_1} \quad \Leftrightarrow \quad \ln \hat{y} = \ln b_0 + b_1 \ln x \quad (\text{для степенной})$$

шаг 2) последовательная замена величин и параметров:

$$\ln \hat{y} = \ln b_0 + x \ln b_1 \quad \Leftrightarrow \quad \hat{y}_1 = a_0 + a_1 x \quad (\text{для показательной})$$

$$\ln \hat{y} = \ln b_0 + b_1 \ln x \quad \Leftrightarrow \quad \hat{y}_1 = a_0 + b_1 z \quad (\text{для степенной})$$

шаг 3) и затем потенцирование для определения спецификации модели:

$$\hat{y} = e^{\ln \hat{y}} = e^{\hat{y}_1} = e^{a_0 + a_1 x} \quad (\text{для показательной})$$

$$\hat{y} = e^{\ln \hat{y}} = e^{\hat{y}_1} = e^{a_0 + b_1 z} \quad (\text{для степенной})$$

Коэффициент эластичности

Для парной нелинейной регрессии коэффициент эластичности показывает, насколько изменится $y=f(x)$ при изменении x на 1%. Рассчитывается как относительное изменение y к относительному изменению x :

$$\varepsilon = \left(\frac{dy}{y} \right) / \left(\frac{dx}{x} \right) = y'(x) \cdot \frac{x}{y}$$

Задания к теме 1.2. Нелинейная регрессия

Задание 5. Оценка параметров нелинейных эконометрических моделей

Скопировать Ваш файл предыдущей работы *Фамилия ЭММ4.xlsx*. Сохранить его под именем *Фамилия ЭММ5 Линейная и нелинейные модели*. Дополнить титульный лист текущей датой и темой работы (ниже предыдущих).

1. Добавить лист *Гипербола*. Скопировать на него исходные данные X и Y (нпр., с листа *База*). Добавить новый столбец с заголовком $1/X$ и рассчитать его. Построить и оценить линейную модель для данных Y и $1/X$ на том же листе. Написать выводы (по образцу линейной регрессии).
2. Добавить лист *Парабола*. Скопировать на него данные X и Y . Добавить и рассчитать столбец X^2 . Построить и оценить на том же листе линейную модель для данных Y и X^2 . Написать выводы.
3. Добавить лист *Корень*. Скопировать на него данные X и Y . Добавить и рассчитать столбец \sqrt{X} . Построить и оценить на том же листе линейную модель для данных Y и \sqrt{X} . Написать выводы.
4. Добавить лист *Логарифм*. Скопировать на него данные X и Y . Добавить и рассчитать столбец $\ln X$. Построить и оценить на том же листе линейную модель для данных Y и $\ln X$. Написать выводы.
5. Для каждой из полученных моделей на соответствующем ее листе:
 - 5.1. Добавить к таблице столбец с заголовком $\hat{y} = y_{оцен}$ и рассчитать его по полученной формуле регрессии.
 - 5.2. Добавить столбец с заголовком $(y - y_{оцен})^2$ и рассчитать его. Найти сумму этого столбца. Проанализировать ниже ее значение (письменно).
 - 5.3. Построить диаграмму с корреляционным полем данных (x, y) и оценочными значениями $(x, y_{оцен})$.

Результат сохранить в своей папке и скопировать в папку *\$control*.

Задание 6. Оценка параметров нелинейных эконометрических моделей.

Скопировать Ваш файл предыдущей работы *Фамилия ЭММ5.xlsx*. Сохранить его под именем *Фамилия ЭММ6 Линейная и нелинейные модели*. Дополнить титульный лист текущей датой и темой работы (ниже предыдущих).

1. Добавить новый лист *Показательная*. Скопировать на него исходные данные Y и X .
 - 1.1. Добавить к таблице столбец с заголовком $\ln y$ и рассчитать его. Построить и оценить на том же листе линейную модель $\ln y = a_0 + a_1x$ для данных $\ln y$ и X .

Написать выводы.

1.2. Добавить столбец $\ln y_{оцен}$ и рассчитать его по полученной формуле регрессии ($\ln y_{оцен} = a_0 + a_1 x$).

1.3. Добавить столбец $\ln y_{оцен}$ и рассчитать его по полученной формуле регрессии ($\ln y_{оцен} = a_0 + a_1 x$).

1.4. Добавить столбец $y_{оцен}$ и рассчитать его ($\hat{y} = y_{оцен} = e^{\ln y_{оцен}}$) (используйте функцию EXP).

1.5. Добавить столбец $(y - y_{оцен.})^2$ и рассчитать его. Найти сумму этого столбца.

Проанализировать ниже ее значение (письменно).

1.6. Построить диаграмму: корреляционное поле данных (x, y) и $(x, y_{оцен.})$.

2. Назвать новый лист *Степенная*. Скопировать на него исходные данные Y и X .

2.1. Добавить столбцы $\ln y$ и $\ln x$ и рассчитать их. Построить и оценить на том же листе линейную модель $\ln y = a_0 + b_1 \ln x$ для данных $\ln y$ и $\ln x$.

Написать выводы.

2.2. Добавить столбец $\ln y_{оцен}$ и рассчитать его по полученной формуле регрессии ($\ln y_{оцен} = a_0 + a_1 \ln x$).

2.3. Добавить столбец $y_{оцен}$ и рассчитать его ($\hat{y} = y_{оцен} = e^{\ln y_{оцен}}$).

2.4. Добавить столбец $(y - y_{оцен.})^2$ и рассчитать его. Найти сумму этого столбца.

Проанализировать ниже ее значение (письменно).

2.5. Построить графики: корреляционное поле данных (x, y) и $(x, y_{оцен.})$.

3. Добавить лист *Сравнение* в конец книги.

3.1. Создать на этом листе таблицу (см. ниже рис. 1) для сравнения данных (используйте ссылки на ячейки соответствующих листов).

3.2. Заполнить эту таблицу (по образцу – см. далее рис. 2).

4. Сделать общие выводы по построенным моделям, указать наилучшую из них (выписать ее спецификацию), обосновать ее выбор (письменно).

5. В соответствии с лучшей моделью сделать прогноз значения y на следующий период, если x возрастет на 10% по сравнению со средним значением x .

Результат сохранить в своей папке и скопировать в папку *\$control*.

	A	B	C	D	E	F	G
1	Фамилия Имя	Дата					
2	Сравнение моделей						
3		F	t(b ₀)	t(b ₁)	R ²	SSE	
4	Линейная						
5	Гипербола						
6	Парабола						
7	Корень						
8	Логарифм						
9	Показательная						
10	Степенная						

Рис. 1. Сравнение моделей

	A	B	C	D	E	F
1	Фамилия Имя	Дата				
2	Сравнение моделей					
3		F	t(b ₀)	t(b ₁)	R ²	SSE
4	Линейная	адекв.	не знач.	знач.	=Линейная!B17	=Линейная!C25
5	Гипербола	адекв.	знач.	знач.		
6	Парабола					
7	Корень					
8	Логарифм					
9	Показательная					
10	Степенная					
11						

Рис. 2. Образец сравнения

РАЗДЕЛ 2. Множественный регрессионный анализ

ТЕМА 2.1. МНОГОФАКТОРНЫЕ МОДЕЛИ

Цель – научиться строить, оценивать и сравнивать модели с большим числом факторов, определять при этом влияние каждого из них в отдельности, а также совокупное их воздействие на моделируемый показатель.

Многофакторная (множественная) регрессия – регрессия результативного признака с двумя или большим числом факторов, т. е. модель вида

$$y = F(x_1, x_2, \dots, x_k).$$

Для выбора спецификации модели необходимо провести:

- 1) отбор факторов для построения модели;
- 2) выбор уравнения модели.

Требования к факторам модели:

- должны быть количественными (если факторы качественные, следует придумать для них количественные соответствия);
- не должна наблюдаться высокая корреляция (и тем более точная функциональная связь) между объясняющими факторами.

Отбор факторов для модели множественной регрессии

Факторы должны объяснять вариацию резульативного признака. Долю объясненной ими вариации показывает коэффициент детерминации R^2 . Влияние других (не учтенных в модели) факторов оценивается как $1 - R^2$. Если в модель добавить еще один фактор, то коэффициент детерминации должен возрасти. Если же это не так, значит, добавленный фактор не улучшил модель и от него надо отказаться. (Лишние факторы приводят к статистической незначимости параметров регрессии по t-критерию Стьюдента.)

Обычно количество объясняющих факторов рекомендуют выбирать в 6-7 раз меньше, чем количество наблюдений. Иначе параметры уравнения регрессии и уравнение в целом будут статистически незначимы.

Условия Гаусса – Маркова для построения классических линейных регрессионных моделей (формулы – см. в *Приложении 1*):

- (1) случайная составляющая имеет нулевое среднее значение;
- (2) случайная составляющая имеет постоянную дисперсию (*гомоскедастичность*);
- (3) значения случайной составляющей не коррелированы друг с другом (отсутствует *автокорреляция*);
- (4) объясняющие переменные не коррелированы со случайной составляющей;
- (5) регрессионная модель правильно специфицирована, линейна по коэффициентам и содержит аддитивную случайную составляющую.

Для множественной линейной регрессии необходимо также условие:

- объясняющие переменные не являются линейными функциями других объясняющих переменных (отсутствует *мультиколлинеарность*);

Для проверки статистических гипотез и построения интервальных оценок важно также выполнение условия:

- случайная составляющая распределена нормально.

Теорема Гаусса – Маркова. Если выполнены условия (1)–(5), то оценки, полученные методом наименьших квадратов, являются несмещенными, состоятельными и эффективными.

В таком случае эти оценки, полученные по МНК, называются *BLUE-best linear unbiased estimators* (наилучшие линейные несмещенные оценки).

Мультиколлинеарность

Мультиколлинеарность – наличие высокой корреляции между объясняющими факторами (независимыми переменными) в задачах множественной регрессии.

Две переменные считаются *коллинеарными*, если коэффициент их корреляции $r_{xy} \geq 0,7$ (между этими переменными существует линейная зависимость). Для множественной регрессии – факторы считаются *мультиколлинеарными*, если коэффициент их корреляции $|r_{x_i, x_j}| \geq 0,7$.

Для выявления мультиколлинеарности рассчитывают матрицу парных корреляций (например, с помощью модуля *Корреляция* пакета анализа MS Excel).

Мультиколлинеарность особенно часто имеет место при анализе макроэкономических данных (например, анализе доходов, производства). Корреляция между факторами существует всегда; важно, чтобы она не была слишком сильной.

Если модель будет построена с включением мультиколлинеарных факторов, в ней могут быть не соответствующие экономической теории знаки или большие значения, все переменные могут оказаться незначимыми (а вся модель в целом при этом значимой) и т.д. Получаемые оценки – например, коэффициенты корреляции – оказываются нестабильными как в смысле статистической значимости, так и по величине и знаку, следовательно, они ненадежны.

Способы устранения мультиколлинеарности

- можно увеличить объем выборки (чем больше данных, тем меньше дисперсии оценок МНК). Однако обычно дополнительные данные нелегко найти;
- заменить мультиколлинеарные факторы или преобразовать их – сложить, разделить, заменить их средним и т.п., но при этом полученные факторные переменные должны иметь экономический смысл;
- зачастую исключают переменные, высококоррелированные с другими. Однако если какие-либо факторы были включены в модель на основе рекомендаций экономической теории, их не следует исключать.

При этом предпочтение в модели отдается не тому фактору, который теснее связан с результативным признаком, а тому, который при тесной связи с результативным признаком наименее тесно связан с другими факторами.

Методы построения моделей множественной регрессии

- метод включения;
- метод исключения.

Метод включения: строится модель с одной объясняющей переменной. Затем добавляется еще один фактор, строится новая модель, и обе модели сравниваются. Если модель улучшилась, добавляют еще один фактор. Если модель ухудшилась – добавленный фактор заменяют на другой. После перебора различных комбинаций объясняющих факторов получают наилучшую модель.

Метод исключения: строится и исследуется модель со всеми факторами. Затем удаляется один фактор, строится и исследуется новая модель и сравнивается с предыдущей. Выбирается лучшая модель, и процесс продолжается.

Для выбора удаляемого фактора:

- строится матрица парных коэффициентов корреляции;
- выбираются два фактора с наибольшим коэффициентом парной корреляции;
- из этих факторов выбирают тот, который имеет меньший коэффициент корреляции с результирующим признаком.

Выбор уравнения модели

Модели множественной регрессии (так же как и парной регрессии) могут быть *линейными* и *нелинейными* – как по переменным, так и по параметрам. Для оценки некоторых нелинейных моделей их сводят к линейным (в том числе с помощью логарифмирования).

Часто используются линейная и степенная модели регрессии. Они позволяют дать хорошую интерпретацию коэффициентов регрессии.

Например, в линейной множественной регрессии

$$y = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_n x_n$$

параметр b_i при x_i характеризует среднее изменение результата y при изменении фактора x_i на единицу и при неизменном значении других факторов, закрепленных на среднем уровне.

Оценка построенной модели регрессии

Качество построенной модели оценивают по коэффициенту детерминации R^2 . Для линейной регрессии он совпадает с квадратом коэффициента множественной корреляции. Для множественной регрессии R^2 всегда увеличивается при добавлении в модель новых факторов x – даже если они не оказывают существенного влияния на y . Поэтому его корректируют, чтобы учесть влияние числа факторов в модели.

Скорректированный (нормированный) коэффициент детерминации \tilde{R}^2 : (или R_{adj}^2):

$$R_{adj}^2 = 1 - (1 - R^2) \frac{n - 1}{n - m - 1} = R^2 - \frac{m}{n - m - 1} (1 - R^2)$$

(n – число факторов, m – число наблюдений).

Значимость уравнения множественной регрессии в целом оценивают по **критерию Фишера**:

$$F_{\text{набл}} = \frac{R^2(n - m - 1)}{(1 - R^2)m}$$

Оценка значимости коэффициентов регрессии осуществляется (как и в парной регрессии) по **t-критерию Стьюдента**.

Сравнение моделей

- Если количество факторов в моделях одинаковое, то модели сравнивают по коэффициенту детерминации.
- Если в модели включено разное количество факторов, то сравнение моделей проводят по нормированному коэффициенту детерминации.

Задания к теме 2.1. Многофакторные модели**Задание 7. Множественная регрессия**

Создайте новую книгу Excel *Фамилия_ЭММ7Множественная регрессия*. Укажите фамилию, группу, название работы, дату выполнения.

Задача 1. Реализуйте рассмотренный ниже пример.

Необходимо исследовать для муниципальных и частных автотранспортных предприятий зависимость прибыли за год (Y) от количества автобусов (X_1 , шт.), маршрутных такси (X_2 , шт.) и формы собственности этих предприятий (см. табл. 4).

Таблица 4

Порядковый № предприятия	Прибыль, млн грн.	Количество автобусов, шт.	Количество маршрутных такси, шт.	Форма собственности предприятия
	Y	X_1	X_2	
1	2241	48	18	муниципальное
2	3818	58	16	частное
3	2905	42	25	муниципальное
4	3237	58	27	муниципальное
5	4150	52	30	частное
6	5727	56	42	частное
7	1328	27	21	муниципальное
8	2905	52	21	муниципальное
9	5561	75	37	частное
10	3652	48	35	частное
11	4648	58	33	частное
12	3486	48	21	муниципальное

1.1. Введите для формы собственности предприятия фиктивную переменную X_3 : 0 – муниципальное АТП, 1 – частное АТП и используйте именно ее в дальнейшем анализе.

1.2. Проверьте факторы на коллинеарность – для этого рассчитайте матрицу парных коэффициентов корреляции (например, с помощью модуля *Корреляция* пакета анализа MS Excel):

	Y	X ₁	X ₂	X ₃
Y	1			
X ₁	0,804124189	1		
X ₂	0,773166796	0,443934029	1	
X ₃	0,781020866	0,550469012	0,633441653	1

Рис. 3. Корреляционная матрица

Ни один из коэффициентов корреляции между факторами не превышает по модулю 0,7, что свидетельствует об отсутствии мультиколлинеарности.

1.3. Постройте трехфакторную модель.

В полученной регрессии значимость коэффициента b_0 (свободного члена) 0,20492 превышает 0,05. Исключите его из рассмотрения и построьте регрессию еще раз, задав в диалоговом окне *Регрессия* опцию *Константа – ноль*.

Оцените полученное уравнение линейной регрессии ($\hat{y} = 38,262 \cdot x_1 + 938,546 \cdot x_2 + 44,808 \cdot x_3$).

Сделайте (письменно) выводы о смысловом значении каждого коэффициента. (Каждый дополнительный автобус приносит в среднем 38,262 млн прибыли, каждое маршрутное такси – 938,546 млн, при одинаковом составе автопарков прибыль частного ($X_3=1$) больше в среднем на 44,808.)

Выпишите на основании полученного уравнения еще два – отдельно для каждого типа предприятий (подставив $X_3=0$ и $X_3=1$).

1.4. Проанализируйте значения полученных коэффициентов детерминации ($R^2 = 0,987$; $\tilde{R}^2 = 0,873$). Сделайте выводы.

Результат сохраните в своей папке и скопируйте в папку *\$control*.

Задача 2

Строительная компания намерена за 2 года осуществить строительство 16-этажного жилого дома на 192 квартиры общей площадью 11600 кв. м. Оцените стоимость строительства на основании данных по аналогичным завершенным стройкам в том же районе. Известны (см. табл. 5):

Y – стоимость строительства дома (тыс. у.е.);

X₁ – общая площадь (кв. м);

X₂ – количество квартир в доме;

X₃ – количество этажей;

X₄ – продолжительность строительства (мес.).

Таблица 5

№	X ₁	X ₂	X ₃	X ₄	Y
1	39596	616	22	56	12998
2	7227	192	16	20	7899,5
3	19227	480	20	34	9184,8
4	20728	288	24	35	11289

№	X ₁	X ₂	X ₃	X ₄	Y
11	14088	384	16	28	7888,8
12	4553	144	18	17	7539,7
13	7045	224	14	20	6233,5
14	41605	672	24	58	13754

№	X_1	X_2	X_3	X_4	Y
5	10056	192	12	23	5984,8
6	14612	336	14	29	7124,2
7	23721	448	16	38	9898,6
8	22795	560	20	37	11116
9	35720	480	20	51	13257
10	3522	112	14	16	5686,1
21	20356	264	22	34	9095,6

№	X_1	X_2	X_3	X_4	Y
15	27976	616	22	43	11226
16	5869	96	12	19	6216,4
17	20212	480	20	35	10725
18	19856	336	12	34	7052,9
19	28490	480	20	43	9211,4
20	6981	128	16	20	7378,4
22	3984	96	12	17	5914,5

Предполагается линейная зависимость стоимости от факторов регрессии.

2.1. Организуйте данные в виде одной вертикальной таблицы.

Постройте матрицу коэффициентов корреляции, проверьте наличие или отсутствие корреляции и мультиколлинеарности. Поясните (письменно) полученные результаты.

2.2. Определите коэффициенты линейного многофакторного уравнения регрессии методом исключения.

(Обратите внимание, что при включении сразу всех объясняющих факторов в модель почти все параметры регрессии окажутся незначимы – хотя имеется сильная корреляция всех факторов с величиной y .)

Для каждого этапа построения регрессии определяйте значимость коэффициентов и адекватность модели (письменно). Указывайте кратко Ваши действия на каждом следующем этапе. (Используйте для анализа нормированный коэффициент детерминации.)

2.3. Выпишите вид полученной многофакторной зависимости. Поясните экономический смысл каждого параметра модели (письменно).

2.4. Оцените полученную модель на основе нормированного коэффициента детерминации.

2.5. Дайте оценку стоимости будущего строительства указанного дома.

Задача 3

Исследовать причины и выявить факторы, оказывающие наибольшее влияние на убытки производства шоколадных изделий кондитерским предприятием. Несущественные факторы исключить соответствующим методом построения множественной регрессии.

Данные предприятия (организовать их в виде одной таблицы):

X_1 – процент реализованной за месяц шоколадной продукции (%);

X_2 – стоимость 1 тонны сырья, в частности какао-бобов (грн.);

X_3 – расход сырья (какао-бобов) на одну тонну кондитерских изделий (%);

X_4 – затраты на 1 грн. произведенной продукции (грн.);

X_5 – расходы на электроэнергию (грн.);

Y – убытки предприятия от данного производства.)

Таблица 6

№ мес	1	2	3	4	5	6	7	8	9	10
Y	93376	72009	58908	70653	65531	100742	69252	86457	68589	66770
X1	47,6	64,9	83,2	63,1	70,6	46,9	67,1	47	70	71,1
X2	5000	4160	3000	4100	3900	5100	4000	4800	3900	3900
X3	0,78	0,68	0,653	0,68	0,678	0,78	0,68	0,67	0,67	0,67
X4	0,9	0,76	0,699	0,74	0,71	0,9	0,73	0,85	0,72	0,7
X5	1144	623	594	676	589	1171	667	884	598	658

№ мес	11	12	13	14	15	16	17	18	19	20
Y	90170	75987	79888	71673	90106	85259	91735	89593	50848	73742
X1	48,6	46,9	52,1	65,4	48,9	48	49,9	49,3	88,7	61,2
X2	4900	4180	4300	4150	4900	4500	5000	4900	2600	4170
X3	0,78	0,67	0,67	0,68	0,78	0,68	0,78	0,78	0,55	0,67
X4	0,87	0,78	0,8	0,75	0,89	0,82	0,9	0,86	0,689	0,77
X5	908	625	825	642	862	833	1055	872	928	672

№ мес	21	22	23	24	25	26	27	28	29	30
Y	80502	83589	91959	79944	67005	83238	100484	100398	110071	110534
X1	49,7	51,1	50,1	53,1	69	46,9	50,5	48,6	47,8	49,6
X2	4200	4700	4900	4600	3900	4400	5200	5300	5400	5500
X3	0,68	0,66	0,78	0,66	0,67	0,66	0,78	0,78	0,78	0,78
X4	0,79	0,84	0,88	0,83	0,71	0,81	0,9	0,9	0,9	0,9
X5	642	860	914	811	595	805	1134	1110	1131	1147

Задача 4 (дополн.)

В табл. 7 приведены данные по состоянию пенсионной системы Украины².

Таблица 7

t	P	PF	CF	N	AE	GDP	GDP-1	NPF	RR
1991	0,001	0,34	0,285	13100	0,005	2,99	0,06	0,055	0,22
1992	0,005	5,08	3,978	13600	0,065	50,33	0,97	1,105	0,08
1993	0,097	133,6	123,2	14200	0,163	1483	28,42	10,4	0,60
1994	2,921	998,9	892,8	14500	14,28	12038	231,8	106,1	0,20
1995	11,56	4409,1	4312,0	14500	73,0	54516	1058	97,1	0,16
1996	38,70	7539,7	7572,7	14488	126,0	81519	1595	-33,0	0,31

²Данные: Л.Якимова. Економетричне моделювання динаміки розвитку солідарної пенсійної системи України / Л.Якимова // Галицький економічний вісник. – 2011. – №2(31). – с.15–22.

<i>t</i>	<i>P</i>	<i>PF</i>	<i>CF</i>	<i>N</i>	<i>AE</i>	<i>GDP</i>	<i>GDP-1</i>	<i>NPF</i>	<i>RR</i>
1997	51,90	9614,2	9530,6	14487	143,0	93365	1842	83,6	0,36
1998	52,20	9749,2	9702,2	14535	153,0	102593	2040	47,0	0,34
1999	60,70	12038,4	12067,7	14520	178,0	130442	2614	-29,3	0,34
2000	68,90	15425,4	14293,0	14530	230,0	170070	3436	1132,4	0,30
2001	83,70	18823,1	18068,5	14447	311,0	204190	4195	754,6	0,27
2002	122,5	21732,3	22834,7	14423	376,0	225810	4685	-1102,4	0,33
2003	136,6	26786,2	24452,6	14376	462,0	267344	5591	2333,6	0,30
2004	182,2	39869,1	39245,7	14348	590,0	345113	7273	623,4	0,31
2005	316,2	65733,3	64064,9	14065	806,0	441452	9372	1668,4	0,39
2006	406,8	72197,1	74007,5	14050	1041,0	544153	11630	-1810,4	0,39
2007	478,4	101379,8	99940,5	13937	1351,0	720731	15496	1439,3	0,35
2008	751,4	143488,2	150349,1	13819	1806,0	948056	20495	-6860,9	0,42
2009	898,4	148372,7	165590,0	13750	1906,0	913345	19832	-17217	0,47
2010	999,0	157900,0	192300,0	13721	2239,0	951705	20789	-34400	0,45

где *P* – средний размер ежемесячной пенсии, грн.; *PF* – доходы ПФУ, млн грн.; *CF* – расходы ПФУ, млн грн.; *N* – количество пенсионеров, тыс. человек; *AE* – среднемесячная заработная плата, грн.; *GDP* – ВВП, млн грн.; *GDP-1* – ВВП на душу населения, грн.; *NPF* – дефицит/профицит ПФУ; *RR* – ставка замещения.

Постройте модель динамики развития солидарной пенсионной системы Украины. (Все выводы по модели – письменно, в том числе обоснование выбора экзогенных переменных)

На ее основе сформируйте прогноз на один год таких ее индикаторов:

а) средний размер пенсии, б) доходы и в) расходы ПФУ.

Результат сохраните в своей папке и скопируйте в папку \$control.

ТЕМА 2.2. ОСОБЫЕ СЛУЧАИ

Фиктивные переменные

Чтобы включить в рассмотрение качественные данные (образование, пол, район, сезон, тип продукции или предприятия и т.п.), надо преобразовать их в количественные переменные, которые называют **фиктивными**.

Обычно фиктивная переменная имеет вид:

$$D = \begin{cases} 1, & \text{если фактор имеет значение}_1 \\ 0, & \text{если фактор имеет значение}_2 \end{cases}$$

Если фактор может иметь *k* вариантов значений, следует вводить *k*–1 фиктивных переменных, принимающих значения 1 либо 0 – так чтобы в совокупности наборы этих 1 и 0 описывали все варианты значений факторного признака.

Значение качественного фактора, при котором все фиктивные переменные равны 0, называют *базовым значением*.

Фиктивные переменные могут входить в регрессионную модель как аддитивно (отдельными слагаемыми),

$$y = b_0 + b_1 \cdot x + b_2 \cdot D + \varepsilon \quad (\text{аддитивная фиктивная переменная})$$

так и мультипликативно.

$$y = b_0 + b_1 \cdot x + b_2 \cdot D + b_3 \cdot x \cdot D + \varepsilon \quad (\text{мультипликативная фиктивная переменная})$$

Мультипликативные фиктивные переменные оказывают влияние не только на свободный член, но и на коэффициент наклона регрессии и позволяют учесть влияние качественных факторов на количественные.

Тест Чоу

Чтобы выяснить, надо ли вводить в полученную модель регрессии дополнительные фиктивные переменные, применяется **метод, или тест Чоу**. Он используется в тех случаях, когда исходные данные можно бы разделить на части (подвыборки), или, наоборот, когда можно бы объединить две выборки данных в одну и рассматривать единую модель регрессии. Тогда тест Чоу позволяет проверить, лучше ли качество общей модели регрессии, чем качества частных моделей подвыборок, или нет.

Суть теста Чоу:

1. Выдвигается гипотеза о том, что лучше качество общей модели. Ее проверяют с помощью F-критерия Фишера: рассчитывают значение критерия

$$F_{\text{расч}} = \frac{SSE - (SSE_1 + SSE_2) \frac{n-2m-2}{m+1}}{SSE_1 + SSE_2},$$

где SSE , SSE_1 , SSE_2 – остаточные суммы квадратов для всей исходной выборки и для обеих ее частей соответственно;

$m+1$ – количество факторов;

$n-2m-2$ – количество наблюдений выборки.

2. Сравнивают полученный результат с критическим значением F-критерия $F_{\text{табл}}$ – его определяют по таблице распределения Фишера с уровнем значимости α (обычно 0,05) и двумя степенями свободы $m+1$ и $n-2m-2$.

Если $F_{\text{расч}} \leq F_{\text{табл}}$, то основная гипотеза принимается, базисная модель является оптимальной и разбивать общую регрессию на подвыборки не имеет смысла.

Если же наоборот, $F_{\text{расч}} > F_{\text{табл}}$, то основная гипотеза отклоняется и необходимо строить две модели. Построение регрессии для каждой подвыборки можно провести, добавив в общую модель набор фиктивных переменных, определяющих разбиение выборки на части.

Гомоскедастичность и гетероскедастичность

Гомоскедастичность – постоянство дисперсии случайной составляющей, независимо от наблюдения. Нарушение этого условия называется *гетероскеда-*

стичностью. Тестирование моделей на гетероскедастичность является одной из необходимых процедур при построении регрессионных моделей.

Гетероскедастичность приводит к неэффективности оценок, полученных с помощью метода наименьших квадратов. В этом случае классические МНК-оценки параметров оказываются смещенными и несостоятельными, и статистические выводы о качестве полученных оценок могут быть неадекватными. (Например, могут оказаться неверными выводы о значимости коэффициентов регрессии и детерминации на основании t - и F -статистик.) При обнаружении гетероскедастичности чаще всего следует переопределить переменные.

Различают *истинную* и *ложную* (смешанную) *гетероскедастичность*. Ложная гетероскедастичность возникает при неверной спецификации модели, в случае не включения в нее существенно влияющих переменных.

Так как гетероскедастичность может проявляться в модели самым разным образом, то нет единого универсального способа ее выявления. Существует множество методов, ни один из которых не однозначен.

Графический способ обнаружения гетероскедастичности: по оси абсцисс откладывают объясняющую переменную (либо их комбинацию, например \hat{y}), а по оси ординат – квадраты остатков ϵ_i^2 . В условиях гомоскедастичности будет равномерное распределение остатков, не зависящее от абсциссы. При наличии гетероскедастичности будут наблюдаться систематические закономерности.

Тест ранговой корреляции Спирмена

Предполагается, что дисперсия случайного члена будет увеличиваться или уменьшаться с ростом объясняющей переменной x , и поэтому в регрессии, оцениваемой с помощью МНК, абсолютные величины остатков ϵ_i значения x будут коррелированы.

Алгоритм теста:

1. Рассчитывают уравнение регрессии и вычисляют остатки (разность между фактическими и расчетными значениями результативного признака).

2. Данные упорядочивают по возрастанию переменной x , от которой предполагается зависимость дисперсии остатков. Каждому значению x ставят в соответствие порядковый номер (*rang* x).

3. Снова упорядочивают выборку, теперь по возрастанию остатков. Каждому значению остатков присваивают ранг (порядковый номер, *rang* ϵ).

4. Рассчитывают коэффициент ранговой корреляции Спирмена между рангами факторов и остатков ϵ .

Коэффициент ранговой корреляции Спирмена – это количественная оценка статистического изучения связи между явлениями, которая показывает, как отличается полученная при наблюдении сумма квадратов разностей между рангами от случая отсутствия связи.

$r_{x,\varepsilon} = 1 - 6 \frac{\sum d_i^2}{n(n^2-1)}$, где d_i – разность рангов x_i и ε_i , n – количество наблюдений в выборке.

5. Найденный коэффициент ранговой корреляции проверяют на значимость. Для этого вычисляют статистический критерий Стьюдента:

$t_{\text{факт}} = \frac{r_{x,\varepsilon}}{\sqrt{1-r_{x,\varepsilon}^2}} \cdot \sqrt{n-k-1}$, где k – число объясняющих переменных в модели.

6. Сравнивают полученное значение $t_{\text{факт}}$ с табличным значением $t_{\text{табл}}$ для уровня значимости α (обычно 0,05) и числа степеней свободы ($n-k-1$). Табличное значение $t_{\text{табл}}$ может быть:

- получено из справочной таблицы распределения Стьюдента (см. Приложение 5)
- или вычислено с помощью функции Excel СТЮДЕНТ.ОБР(γ ; $n-k-1$) (где $\gamma = 1-\alpha$).

Если $t_{\text{факт}} \geq t_{\text{табл}}$, то принимается гипотеза о наличии гетероскедастичности. Если $t_{\text{факт}} < t_{\text{табл}}$, то считают, что в модели наблюдается гомоскедастичность.

Параметрический тест Голдфельда–Квандта

Предполагается пропорциональность дисперсий отклонений и значений объясняющей переменной x . Тест используется для значительных по объему выборок. Количество наблюдений n должно по крайней мере в два раза превосходить количество переменных m .

Алгоритм теста:

1. Упорядочивают все n наблюдений по переменной x , от которой предполагается зависимость дисперсии остатков (или наиболее существенной для модели).

2. Разбивают всю выборку на три подвыборки и исключают срединную часть (вторую подвыборку) из k наблюдений. Обычно k принимают от $n/4$ до $n/3$ наблюдений. (Тогда первая и вторая подвыборки будут содержать по $(n-k)/2$ значений каждая.)

Если исходное предположение верно, то дисперсия регрессии по первой подвыборке будет существенно меньше дисперсии регрессии по третьей подвыборке.

3. Оценивают отдельные регрессии для первой и третьей подвыборок и вычисляют суммы квадратов их остатков.

4. Для сравнения соответствующих дисперсий рассчитывают соответствующую F -статистику:

$F_{\text{факт}} = \frac{S_3}{S_1}$, где $S_{1,3}$ – суммы квадратов соответствующих остатков.

5. Сравнивают полученное значение $F_{факт}$ с табличным значением критерия Фишера $F_{табл}$ (с выбранным уровнем значимости – например, 0,05 – и числом степеней свободы $\nu_1 = \nu_2 = \frac{n-k}{2} - m$). Табличное значение $F_{табл}$ может быть:

- получено из справочной таблицы распределения Фишера (см. Приложение 4)
- или вычислено с помощью функции Excel F.ОБР ($\gamma; (n-k)/2-m; (n-k)/2-m$) (где $\gamma = 1-\alpha$).

Если $F_{факт} \leq F_{табл}$, то принимается гипотеза о наличии гомоскедастичности. Если $F_{факт} > F_{табл}$, то считают, что в модели наблюдается гетероскедастичность.

Автокорреляция

Модель называется **автокоррелированной**, если не выполняется условие (3) Гаусса – Маркова: $\text{cov}(\varepsilon_i, \varepsilon_j) = 0$ для $i \neq j$ (т. е. между остатками есть зависимость).

Положительная автокорреляция означает постоянное действие неучтенных факторов на результирующую переменную в одном направлении (за положительным отклонением обычно следует положительное: $\text{cov}(\varepsilon_i, \varepsilon_j) > 0$); *отрицательная* – разнонаправленное действие (за положительным отклонением чаще всего следует отрицательное и наоборот: $\text{cov}(\varepsilon_i, \varepsilon_j) < 0$).

Основные причины, приводящие к автокорреляции, – ошибки в спецификации модели (не учтены какие-либо переменные либо неверно выбрана форма зависимости); сглаживание данных; объединение факторов; ошибки измерения; цикличность либо взаимозависимость экономических показателей и др.

Автокорреляция приводит к неэффективности оценок, занижению дисперсии и стандартных ошибок коэффициентов (из-за чего незначимые факторы могут быть признаны значимыми). В результате могут оказаться неверными выводы о значимости коэффициентов регрессии и детерминации на основании t - и F -статистик.

При наличии автокорреляции остатков полученное уравнение регрессии считают неудовлетворительным, даже при высоком значении R^2 и значимости коэффициентов. Следует попытаться, например, найти неучтенный фактор либо изменить формулу зависимости. Если автокорреляцию остатков не удастся устранить, возможно, она вызвана внутренними свойствами ряда.

Задания к теме 2.2. Особые случаи

Задание 8. Модель с фиктивными переменными

Создайте новую книгу Excel *Фамилия_ЭММ8Фиктивные переменные*. Укажите фамилию, группу, название работы, дату выполнения.

Задача 1.

Министерство инфраструктуры исследует стоимость железнодорожных грузоперевозок (C , у.е.) в зависимости от расстояния (S , км), при этом сравниваются расценки трех региональных управлений железной дороги (А, В, С). Имеются статистические данные на грузоперевозки по 36 договорам (заказам) (табл. 8).

Таблица 8

Регион	С	В	В	С	А	А	С	С	В
S	614	403	344	533	544	585	423	282	438
С	705,9	470,2	380,8	596,3	634,4	636,3	539,4	334	472,4

Регион	А	С	В	С	В	А	В	В	В
U	573	333	333	622	401	334	442	283	355
S	635,6	416,8	355,5	680,8	468,7	401,1	512,1	305,1	409,7

Регион	А	А	В	А	А	С	В	А	В
U	465	520	173	455	410	77	333	419	360
S	550,1	553	178,6	501,2	431,1	108,7	381,2	444,2	403,6

Регион	С	С	А	В	А	С	С	С	А
U	623	202	504	306	394	158	247	434	532
S	698,4	242,7	554,7	357	413,4	215,9	316,3	477	608

Необходимо построить линейные регрессионные модели зависимости урожайности от количества осадков – для каждого из трех регионов и общую модель.

1. Скопируйте в свой файл исходные данные и преобразуйте их в одну вертикальную таблицу. Отсортируйте данные по регионам.
2. Постройте корреляционное поле для переменных *Урожайность*(U) и *Осадки* (S), изображая точки, соответствующие различным регионам, разными символами (на диаграмме должно быть 3 ряда данных).
3. Оцените уравнение регрессии $U = b_0 + b_1 \cdot S + \varepsilon$. Что оно отражает? Оцените качество построенной регрессии (сделайте выводы).
4. Создайте два новых (фиктивных) столбца D_1 и D_2 . Присвойте им значения

$$D_1 = \begin{cases} 1, & \text{если регион А,} \\ 0, & \text{в другом случае,} \end{cases} \quad D_2 = \begin{cases} 1, & \text{если регион В,} \\ 0, & \text{в другом случае} \end{cases}$$
 - 4.1. Оцените уравнение регрессии $U = b_0 + b_1 \cdot S + b_2 \cdot D_1 + b_3 \cdot D_2 + \varepsilon$, где D_1, D_2 – фиктивные переменные, отражающие различие трех фирм. Дайте интерпретацию построенной регрессии. Оцените качество построенной регрессии (сделайте выводы).

x7	1	1	1	1	0	1	1	1	1	1	0	1	0	0	0	1
x8	0	0	3	3	0	0	20	20	20	20	0	12	0	0	0	6
y	31,8	58	30,8	57,2	31,2	55,4	68,2	75,4	83,8	48,8	42,6	73,4	43	52,8	107,8	68,4

№ п/п	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
x1	2	3	1	2	3	1	2	3	4	1	2	3	2	3	3	3
x2	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	3
x3	71,1	68	38	93,2	117	42	62	89	132	40,8	59,2	65,4	60,2	82,2	98,4	76,7
x4	36,2	41	19	49,5	55,2	21	35	52,3	89,6	19,2	31,9	38,9	36,3	49,7	52,3	44,7
x5	13,3	8	7,4	14	25	10,2	11	11,5	11	10,1	11,2	9,3	10,9	13,8	15,3	8
x6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
x7	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1
x8	6	12	12	12	12	12	12	12	12	6	6	6	12	12	12	0
y	71,2	68	38	93,2	117	48,4	71,4	102,4	151,8	42,4	61,6	68	63,8	87,2	104,4	86,2

№ п/п	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64
x1	1	2	3	1	1	2	2	3	1	2	3	1	2	3	1	2
x2	3	3	3	3	3	3	3	3	4	4	4	4	4	4	3	3
x3	38,7	56,4	76,7	38,7	41,5	48,8	57,4	76,7	37	54	68	40,5	61	80	52	78,1
x4	20	32,7	44,7	20	20	28,5	33,5	44,7	17,5	30,5	42,5	16	31	45,6	21,2	40
x5	10,2	10,1	8	10,2	10,2	8	10,1	8	8,3	8,3	8,3	11	11	11	11,2	11,6
x6	1	1	1	1	1	1	1	1	0	0	0	0	0	0	1	1
x7	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1
x8	6	6	6	0	0	0	0	0	7	7	7	3	3	3	18	18
y	50	70,4	81,6	36,4	40,2	45,4	55,2	72	35,6	51,8	65,2	39,6	59,8	78,4	44,8	70,4

№ п/п	65	66	67	68	69
x1	3	1	2	3	4
x2	3	4	4	4	4
x3	91,6	39,9	56,2	79,1	91,6
x4	53,8	19,3	31,4	42,4	55,2
x5	16	8,4	11,1	15,5	9,4
x6	1	0	0	0	0
x7	0	1	1	1	1
x8	18	6	6	6	6
y	82,4	35,6	50	70,4	81,6

где: y – цена квартиры, тыс. у. е.;

x_1 – число комнат в квартире;

x_2 – район города (1 – Приморский; 2 – Нагорный; 3 – Юго-Запад; 4 – Павловка);

x_3 – общая площадь квартиры (m^2);

x_4 – жилая площадь квартиры (m^2);

x_5 – площадь кухни (m^2);

x_6 – тип дома (кирпичный / другой);

x_7 – наличие балкона (есть / нет);

x_8 – число месяцев до окончания срока строительства.

Проведите необходимые расчеты и дайте письменные ответы на вопросы. Не забывайте указывать номера заданий.

1. Определите факторы, формировавшие цену квартир в строящихся домах. (Указание. Сгенерируйте фиктивную переменную D , отражающую местоположение квартиры и позволяющую разделить всю совокупность квартир на две группы: квартиры на севере города ($D=1$ для Приморского района и Павловки) и на юге города ($D=0$ для Юго-Запада и Нагорного районов). Аналогично введите фиктивные переменные D_2, D_3 для других качественных переменных.

2. Составьте матрицу парных коэффициентов корреляции:

а) исходных переменных. Вместо переменной x_2 используйте фиктивную переменную D ;

б) логарифмов исходных переменных (кроме фиктивных переменных).

3. Постройте уравнение линейной множественной регрессии $y = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_m x_m + \varepsilon$, характеризующие зависимость цены от всех факторов. Установите, какие факторы мультиколлинеарны. Методом исключения уберите мультиколлинеарные переменные.

5. На том же листе сделайте выводы (письменно):

а) Какие факторы значимо воздействуют на формирование цены квартиры в этой модели?

б) Существует ли разница в ценах квартир, расположенных в северной и южной частях города? Если да, то какая?

в) является ли наличие балкона или лоджии преимуществом квартиры на рынке? Как Вы объясните этот факт?

6. Ниже выводов сделайте прогноз на стоимость квартиры в каждой части города: общая площадь – 64 кв.м, жилая площадь – 42 кв.м, площадь кухни – 9 кв.м, в кирпичном доме и при наличии балкона.

7. Скопируйте исходные данные на лист *Степенная*. Постройте аналогично степенную множественную регрессию ($\hat{y} = b_0 x_1^{b_1} x_2^{b_2} x_3^{b_3} \dots x_m^{b_m} \varepsilon$). Для этого:

- Рассчитайте логарифмы всех переменных (кроме фиктивных).
- Отобразите их корреляционную матрицу.
- Выявите мультиколлинеарность.
- Проведите метод исключения мультиколлинеарных переменных.
- Выпишите модель и сделайте ее анализ.
- Опишите смысл каждой переменной в ней.
- Сделайте прогноз стоимости для тех же данных, что и п. 6.

8. Сравните линейную и степенную модели (письменно). В какой модели мультиколлинеарность проявляется сильнее? (Если количество факторов в моделях одинаковое, то сравнивать по коэффициенту детерминации; разное – по нормированному коэффициенту детерминации.)

Результат сохраните в своей папке и скопируйте в папку \$control.

Задание 9. Гомо- и гетероскедастичность

Создайте новую книгу Excel *Фамилия_ЭММ9Гетероскедастичность*. Укажите фамилию, группу, название работы, дату выполнения.

Задача 1. Тест Голдфельда–Квандта

Дана таблица (см. табл. 10) для оценки зависимости уровня продаж торговых точек (y) от численности потенциальных покупателей (x). Требуется по методу Голдфельда–Квандта проверить модель на гетероскедастичность.

Указания.

1. Свелите данные в одну вертикальную таблицу.
Упорядочьте данные по возрастанию переменной x .
2. Исключите, например, девять центральных наблюдений, оставив две подвыборки по двенадцать наблюдений.
3. Для полученных подвыборок постройте уравнения регрессии, найдите остатки и рассчитайте их суммы квадратов.
4. Рассчитайте значение критерия F .
5. Сравните полученное значение с табличным значением критерия Фишера с уровнем значимости 0,05 (используйте функцию Excel F.ОБР).
6. Сделайте вывод о наличии гетероскедастичности ошибок модели.

Таблица 10

№ п/п	1	2	3	4	5	6	7	8
Доход y , грн	215838	475464	197158	610075	764135	182518	494200	482793
Численность потенциальных покупателей x , чел.	195132	202752	249978	165747	147550	242420	694045	251220

№ п/п	9	10	11	12	13	14	15	16
Доход y , грн	196992	216104	579152	493713	316692	307704	309972	635150
Численность потенциальных покупателей x , чел.	417200	151408	267684	831660	185853	200882	118386	419700

№ п/п	17	18	19	20	21	22	23	24
Доход y , грн	667020	501372	487544	403782	305874	548110	749420	196776
Численность потенциальных покупателей x , чел.	343480	449682	286930	740420	458080	158799	814000	118002

№ п/п	25	26	27	28	29	30	31	32	33
Доход y , грн	281582	405040	279034	345708	546996	210134	547488	234292	490614
Численность потенциальных покупателей x , чел.	380480	147600	567830	388250	467688	250248	919765	302285	325325

покупателей x , чел.									
------------------------	--	--	--	--	--	--	--	--	--

Задача 2. Тест Спирмена

На другом листе Excel проверьте по тем же исходным данным наличие гетероскедастичности для того же фактора с помощью теста ранговой корреляции Спирмена.

Для определения табличного значения t -критерия Стьюдента используйте функцию ExcelСТЮДЕНТ.ОБР.

Результатсохраните в своей папке и скопируйте в папку \$control.

РАЗДЕЛ3. Моделирование динамических процессов

ТЕМА 3.1. АВТОКОРРЕЛЯЦИЯ

Модели, которые включают текущие и некоторые предыдущие по времени значения объясняющих переменных, а также время t , называют *динамическими*.

Временной ряд – это совокупность значений какого-либо показателя за несколько последовательных моментов (периодов) времени.

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

Коэффициент автокорреляции 1-го порядка– это коэффициент корреляции между уровнями исходного временного ряда y_t и уровнями того же ряда, сдвинутыми на один момент времени y_{t-1} . При этом общее число пар наблюдений, по которым производится расчет, равно $(n-1)$.

Коэффициент автокорреляции порядка τ – это коэффициент корреляции $r_{t,t-\tau}$ между рядами y_t и $y_{t-\tau}$:

$$r_{t,t-\tau} = \frac{\text{cov}(y_t, y_{t-\tau})}{\sigma_t \cdot \sigma_{t-\tau}}$$

Число периодов τ , по которым рассчитывается коэффициент автокорреляции, называется *лагом* (или *запаздыванием*). Например, автокорреляция 4-го порядка может наблюдаться в ежеквартальных данных (за несколько лет).

Значение коэффициента автокорреляции, близкое по модулю к 1, говорит о наличии во временном ряде *тенденции* (тесной связи между текущими и предшествующими уровнями временного ряда).

Набор коэффициентов автокорреляции с последовательными значениями лагов задает **автокорреляционную функцию** (ACF) временного ряда. Ее график называется **коррелограммой** (по оси абсцисс – временной лаг, по оси ординат – значения коэффициентов автокорреляции). Наибольшее значение коэффициента $r_{t,t-\tau}$ позволяет определить лаг (периоды колебания ряда).

Модели временных рядов, в которых значения временного ряда в данный момент линейно зависят от предыдущих значений этого же ряда, называются **авторегрессионными**. Авторегрессионный процесс p -го порядка $AR(p)$ отражает зависимость от значений за p предыдущих периодов:

$$y_t = b_0 + b_1 y_{t-1} + \dots + b_p y_{t-p} + \varepsilon_t$$

Автокорреляция в остатках – корреляционная зависимость между остатками в последовательные моменты времени.

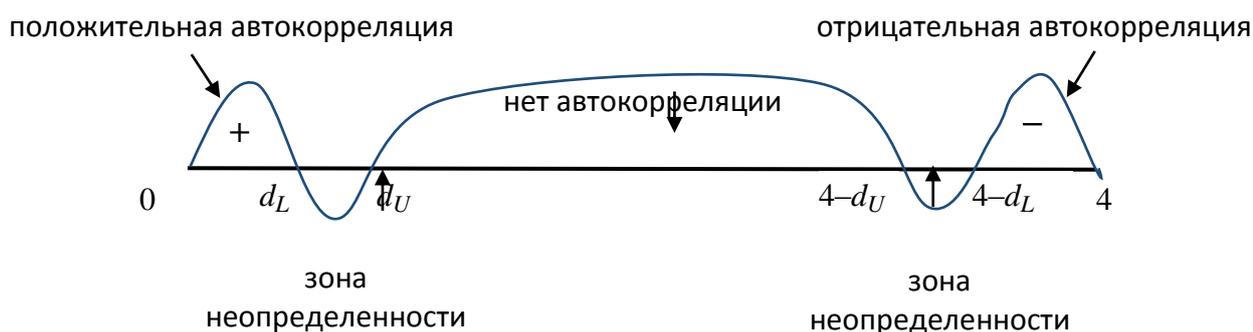
Критерий Дарбина – Уотсона обнаружения автокорреляции остатков.

1. По построенному уравнению регрессии вычисляют остатки.

2. Рассчитывают статистику Дарбина – Уотсона $DW = \frac{\sum_2^n (\varepsilon_t - \varepsilon_{t-1})^2}{\sum_1^n \varepsilon_t^2}$

3. По таблице критических точек распределения Дарбина – Уотсона находят нижнюю d_L и верхнюю d_U границу DW -статистики – для заданного уровня значимости α , количества наблюдений n и числа объясняющих переменных m .

4. Проверяют, в какой из 5 интервалов попала рассчитанная статистика DW :



Если DW попадает в интервал $[d_U; 4-d_L]$, то автокорреляция отсутствует. Значения DW в пределах $[0; d_L]$ и $[4-d_L; 4)$ подтверждают наличие автокорреляции (соответственно положительную и отрицательную). Для интервалов $[d_L; d_U]$ и $[4-d_U; 4-d_L]$ гипотеза об отсутствии автокорреляции не может быть ни принята, ни отклонена.

Можно оценить значение DW -статистики приблизительно, учитывая что $DW \approx 2 \cdot (1 - r_{\varepsilon_t \varepsilon_{t-1}})$, следовательно, $0 \leq DW \leq 4$. При $r_{\varepsilon_t \varepsilon_{t-1}} = 0$ (корреляция отсутствует) $DW \approx 2$. Приблизительно можно считать, что автокорреляция отсутствует, если $1,5 < DW < 2,5$.

Критерий Дарбина – Уотсона не применим для авторегрессионных моделей временных рядов с лагом в 1 период; требует большого размера выборки, наличия свободного члена в модели; предполагается, что остатки удовлетворяют схеме 1-го порядка $\varepsilon_t = k \cdot \varepsilon_{t-1} + v_t$.

Если автокорреляцию остатков не удается устранить, возможно, она вызвана внутренними свойствами ряда. В таком случае для ее устранения может быть использован метод скользящих средних.

ТЕМА 3.2. МОДЕЛИ ДИНАМИЧЕСКИХ ПРОЦЕССОВ

Классы динамических моделей:

- с лагами (содержащие только объясняющие переменные в предыдущие моменты времени) – зачастую их можно свести к уравнению множественной регрессии;
- авторегрессионные (включают также и значения зависимых переменных в предыдущие моменты времени).

Каждый уровень временного ряда можно считать сформированным из трех компонент: трендовой (Т), циклической (периодической) (S) и случайной (E). В зависимости от вида связи между этими компонентами может быть построена:

- **аддитивная модель временного ряда:** $Y(t) = T(t) + S(t) + E(t)$;
- **мультипликативная модель временного ряда:** $Y(t) = T(t) \cdot S(t) \cdot E(t)$.

Построение аналитической функции для моделирования тенденции (тренда) временного ряда называют *аналитическим выравниванием временного ряда*.

Выделение циклической компоненты осмысленно в случае подверженности данных циклическим (недельным, месячным, квартальным, годовым и т. п.) колебаниям. Циклическую компоненту часто называют **сезонной**.

Способы моделирования сезонности:

- авторегрессионная модель;
- модель с фиктивными переменными;
- модель скользящего среднего – аддитивная модель;
- модель скользящего среднего – мультипликативная модель и др.

Авторегрессия временного ряда

1. Рассчитывают значения автокорреляционной функции АСF временного ряда (т.е. коэффициенты корреляции между уровнями ряда и их предшествующими значениями с лагом 1, 2...) – для каждой пары наблюдений (y_t, y_{t-1}) ; общее число таких пар равно $(n - 1)$.

2. Определяют наибольшее значение функции и соответствующий ему лаг. Для визуального отображения строят коррелограмму (график автокорреляционной функции АСF).

3. Для данного лага строят регрессию вида $y_t = b_0 + b_1 y_{t-p} + \varepsilon_t$.

Применение фиктивных переменных при исследовании сезонных колебаний

Число сезонных фиктивных переменных в регрессионной модели всегда должно быть на единицу меньше сезонов внутри периода, т. е. должно быть равно величине $L - 1$, если L сезонов. Например, для поквартальных данных необходимы

3 фиктивные переменные, а общий вид регрессионной модели будет одной из разновидностей аддитивной модели временного ряда:

$$y_t = b_0 + b_1 t + b_2 D1 + b_3 D2 + b_4 D3 + \varepsilon_t$$

Подставив наборы значений фиктивных переменных, получим частные модели регрессии для каждого сезона.

Выбор между аддитивной и мультипликативной моделями:

- построить график ряда динамики (по оси абсцисс – временные промежутки, по оси ординат – уровни ряда);
- проанализировать наличие сезонной компоненты: если сезонные колебания имеют приблизительно одинаковую амплитуду, то для анализа временного ряда используют модель с аддитивной компонентой. Если амплитуда сезонных колебаний уменьшается или увеличивается, то для построения лучше подходит модель с мультипликативной компонентой.

Обычно сезонную компоненту рассчитывают и исключают из ряда (т. е. ряд корректируют на сезонность), затем выявляют тенденцию, т. е. трендовую компоненту.

При построении прогноза сначала находят точечный прогноз трендовой компоненты, а затем применяют сезонную поправку. Точечный прогноз случайной компоненты в аддитивной модели принимается равным нулю.

Этапы построения аддитивной и мультипликативной моделей:

- 1) выравнивание (сглаживание) исходного ряда методом скользящей средней;
- 2) расчет скорректированных значений сезонной компоненты S ;
- 3) устранение сезонной компоненты из исходных уровней ряда и получение выровненных данных: для аддитивной модели – вычитание сезонной компоненты из исходных значений; для мультипликативной – деление исходных значений на сезонную составляющую.
- 4) построение тренда T на основе полученных данных;
- 5) расчет значений соответственно $(T + S)$ или $(T * S)$;
- 6) расчет абсолютных и/или относительных ошибок;
- 7) расчет среднеквадратической ошибки – для обоснования соответствия модели исходным данным или для выбора из множества моделей наилучшей.

1) Методы сглаживания (выравнивания) временного ряда

Методы идентичны и для аддитивной, и для мультипликативной моделей: метод усреднения по двум половинам; метод укрупнения интервалов; метод экспоненциальной средней; метод скользящей средней.

Метод скользящей средней основан на расчете средних уровней ряда за определенный период. Если этот период включает L сезонов, надо усреднить

(сгладить) каждые L последовательных значений ряда. В том случае, если L оказалось четным, после этого полученные средние надо центрировать (т.е. найти среднее для каждых двух последовательных полученных результатов).

2) Расчет сезонных компонент

В аддитивной модели:

Сезонная компонента в аддитивной модели рассчитывается как разность между исходным значением ряда и найденным скользящим средним (центрированным для четных L), соответствующим ему.

Полученные сезонные оценки надо усреднить для каждого сезона.

Сумма полученных сезонных отклонений должна равняться нулю, т.к. в моделях с сезонной компонентой предполагается, что сезонные воздействия за период взаимопогашаются. Для этого надо скорректировать найденные усредненные сезонные отклонения: вычесть из каждого $1/L$ от их суммы.

В мультипликативной модели:

Сезонная компонента в мультипликативной модели рассчитывается как частное между исходным значением ряда и найденным скользящим средним (центрированным для четных L), соответствующим ему.

Полученные сезонные оценки надо усреднить для каждого сезона.

Сумма полученных сезонных отклонений должна равняться числу сезонов в периоде. Иначе надо провести их корректировку, умножив каждое на коэффициент: частное от L , деленного на их сумму.

3) Устранение сезонной компоненты

Найденные сезонные (скорректированные) компоненты надо устранить из исходных данных.

Для аддитивной модели: вычесть из исходного значения; для мультипликативной – разделить его на сезонную компоненту.

Полученные оценки можно использовать для построения тренда (предварительно можно построить для них корреляционное поле и визуально определить форму зависимости).

К найденному тренду необходимо прибавить (для аддитивной модели) либо умножить (для мультипликативной) скорректированные сезонные отклонения.

В заключение следует рассчитать ошибки, сделать выводы, в случае необходимости – построить прогноз. Поскольку сезонным колебаниям обычно отвечают небольшие временные интервалы, то возможно лишь краткосрочное прогнозирование по моделям с сезонной компонентой.

Задания к разделу 3. Моделирование динамических процессов**Задание 10-11. Временные ряды**

Создайте новую книгу Excel *Фамилия_ЭММ10-11 Временные ряды*. Укажите фамилию, группу, название работы, дату выполнения.

Задача 1. Моделирование тренда временного ряда

Имеются данные об объемах потребления электроэнергии жителями региона (Y_t) за 16 кварталов:

Таблица 11

№ кварта- тала	Y_t						
1	6,0	5	7,2	9	8,0	13	9,0
2	4,4	6	4,8	10	5,6	14	6,6
3	5,0	7	6,0	11	6,4	15	7,0
4	9,0	8	10,0	12	11,0	16	10,8

1. Постройте график исходных данных.

Имеется ли, на Ваш взгляд, восходящая или нисходящая тенденция? Имеются ли явно выраженные колебания объема продаж в зависимости от квартала?

2. Добавьте линию линейного тренда, отобразите уравнение тренда и коэффициент детерминации.

3. Получите уравнение линейной регрессии с помощью надстройки *Анализ данных*.

4. Выпишите оцененное уравнение. Проверьте значимость уравнения и параметров.

5. Проинтерпретируйте параметры (с оговоркой на значимость или незначимость).

Задача 2. Авторегрессионная модель временного ряда

Для данных задачи 1 построить авторегрессионную модель (регрессию ряда на самого себя с лагом). Для этого:

1. Рассчитайте значения автокорреляционной функции ACF временного ряда (т.е. коэффициенты корреляции между уровнями ряда и их предшествующими значениями с лагом 1, 2...). (*Указание*. Можно использовать Пакет анализа либо функцию Excel.) Постройте коррелограмму (график функции ACF).

2. Найдите наибольшее по модулю значение и построьте для соответствующего ему лага линейную регрессию. Оцените значимость коэффициента и регрессии в целом.

3. Постройте на диаграмме графики исходного и смоделированного ряда данных. Оцените визуально, удовлетворительно ли модель описывает данные?

4. Сделайте вывод о качестве построенной модели.

Задача 3. Фиктивные переменные

Для исходных данных задачи 1 постройте модель с сезонностью вида

$$y_t = b_0 + b_1 t + b_2 D1 + b_3 D2 + b_4 D3 + \varepsilon_t$$

где D_i – фиктивные переменные, принимающие значение 1, если наблюдение относится к i -му кварталу, и 0 в противном случае ($i = 1, 2, 3, 4$).

1. Оцените эту модель, предварительно добавив фиктивные переменные D_1, \dots, D_3 .

2. Выпишите оцененное уравнение регрессии. Проверьте его значимость (в целом и по отдельным параметрам). Если есть незначимый фактор, исключите его из уравнения и переоцените регрессию. Постройте и оцените визуально график остатков. Сделайте вывод о качестве построенной модели.

3. На основе последней оцененной модели выпишите четыре уравнения – отдельно для каждого квартала. Учитывая 95%-е доверительные интервалы для коэффициентов регрессии, сделайте вывод о том, в какие кварталы потребление больше.

4. Постройте прогноз на следующие два квартала.

Задача4. Скользящее среднее – аддитивная модель

Постройте аддитивную модель временного ряда. Для этого:

1. Сгладьте ряд скользящим средним с лагом 4.

2. Получите скорректированные оценки сезонных компонент отдельно для каждого квартала.

3. Исключите сезонную компоненту и постройте линейную регрессию для трендовой компоненты.

4. Выпишите итоговое уравнение регрессии. С его помощью постройте прогноз на следующие 2 квартала.

Задача5. Скользящее среднее –мультипликативная модель

Аналогично предыдущему заданию, для исходных данных задачи 1 постройте мультипликативную модель временного ряда. Для этого:

1. Сгладьте ряд скользящим средним с лагом 4.

2. Получите скорректированные оценки сезонных компонент отдельно для каждого квартала.

3. Исключите сезонную компоненту и постройте линейную регрессию для трендовой компоненты.

4. Выпишите итоговое уравнение регрессии. С его помощью постройте прогноз на следующие 2 квартала.

На отдельном листе Excel сравните 4 последние полученные модели.

СПИСОК УЧЕБНО-МЕТОДИЧЕСКОЙ ЛИТЕРАТУРЫ

Базовая

1. Михайленко С. В. Оптимизационные методы и модели: учеб. пособие для самостоят. изучения дисциплины / С. В. Михайленко, Е. В. Свищева ; Нар. укр. акад., [каф. информ. технологий и математики]. – Харьков : Изд-во НУА, 2012. – 184 с. – (Математика для экономистов).
2. Козьменко О. В. Економіко-математичні методи та моделі (економетрика): навч. посіб. / О. В. Козьменко, О. В. Кузьменко. – Суми : Унів. кн., 2014. – 405 с. – Бібліогр.: с. 399-400.
3. Економетрика : навч. посіб. / О. Є. Лугінін, В. М. Фомішина, О. М. Дудченко [та ін.]. – Херсон : ОЛДІ-ПЛЮС, 2014. – 320 с..
4. Доля В. Т. Економетрія [Текст] : навч. посіб. для студ. вищ. навч. закл. за спец. з екон. і менедж. / В. Т. Доля ; Харк. нац. акад. міськ. госп-ва. – Харьков: ХНАМГ, 2010. – 171 с. : рис., табл.

Дополнительная

5. Доугерти Кр. Введение в эконометрику [пер. с англ.] – М.: Инфра-М, 2009. – 479 с.
6. Эконометрика [учеб.] / под ред. И. И. Елисеевой. – М. : Проспект, 2011. – 288 с..
7. Практикум по эконометрике / под ред. И. И. Елисеевой. – М.: Финансы и статистика, 2006. – 344с.
8. Эконометрия: Опорные материалы к лекциям и практическим занятиям/ ХГИ «НУА»; Сост. А. Ю. Петрова. – Харьков, 2000. – 24 с.
9. Эконометрия: Метод. указания к практ. занятиям по теме «Парная регрессия»: Для студентов фак. «Бизнес-упр.» / НУА; Сост. А. Ю. Петрова. – Харьков, 2000. – 20 с.
10. Єгоршин О. О. Лабораторний практикум з економетрики навчальної дисципліни "Економіко–математичні методи та моделі" [навч.–практ. посібник] / О. О. Єгоршин, Л. М. Малярець. – Харків: Вид. ХНЕУ, 2011 – 148 с.
11. Бегун С. І. Економетрика : методичні вказівки для самостійної роботи / С. І. Бегун. – Луцьк: вид-во ПП Іванюк В.П., 2014. – 60 с.
12. Эконометрика в примерах и задачах для иностранных студентов : учебное пособие / Л. М. Малярець, Э. Ю. Железнякова, Л. А. Норик. – Харьков: Изд. ХНЭУ им. С. Кузнеця, 2014. – 268 с.

Приложение 1
ОСНОВНЫЕ ФОРМУЛЫ

Характеристика	Формула	Функция Excel
Среднее арифметическое (где n – количество наблюдений):	$\bar{x} = \sum_{i=1}^n \frac{x_i}{n}$	=СРЗНАЧ(х)
Дисперсия:	$D(x) = \sigma_x^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$	= ДИСП.Г(х)
Стандартное (среднее квадратическое) отклонение:	$\sigma_x = \sqrt{D(x)} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$	=СТАНДОТКЛОН.Г(х)
Коэффициент вариации:	$V = \frac{\sigma_x}{\bar{x}} \times 100\%$	
Коэффициент ковариации:	$\text{cov}(x, y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$	= КОВАР(у; х)
Коэффициент корреляции:	$r_{xy} = \frac{\text{cov}(X, Y)}{\sqrt{D(X) \cdot D(Y)}} = \frac{\text{cov}(X, Y)}{\sigma_x \cdot \sigma_y}$	= КОРРЕЛ(у; х)

Дисперсия отражает меру разброса данных вокруг средней величины.

Стандартное (среднее квадратическое) отклонение показывает абсолютное отклонение измеренных значений от среднеарифметического.

Коэффициент вариации отражает степень разбросанности значений независимо от их масштаба и единиц измерения (< 33% – совокупность данных считается однородной; > 33% – говорит о неоднородности информации и необходимости исключения самых больших и самых маленьких значений).

Коэффициент ковариации говорит о степени зависимости двух случайных величин (>0 – обе величины изменяются в одном направлении; <0 – изменения разнонаправленны).

Коэффициент корреляции r_{xy} – показатель взаимного влияния двух величин.

$r_{xy} > 0$ – связь прямая

$r_{xy} < 0$ – связь обратная

Коэффициент корреляции принимает значения от -1 до +1:

$0 < |r_{xy}| < 0,3$ – связь практически отсутствует;

$0,3 < |r_{xy}| < 0,5$ – связь слабая;

$0,5 < |r_{xy}| < 0,7$ – связь умеренная;

$0,7 < |r_{xy}| < 0,99\dots$ – связь сильная;

$|r_{xy}| = 1$ – связь функциональная (полная).

Общий вид эконометрической модели:

$$y = f(x_1, x_2, \dots, x_n, \varepsilon),$$

где y – зависимая, или эндогенная переменная;
 x_1, x_2, \dots, x_n – независимые (экзогенные) переменные, определяющие поведение y ;
 ε – случайная ошибка для генеральной совокупности.

Источники присутствия возмущения:

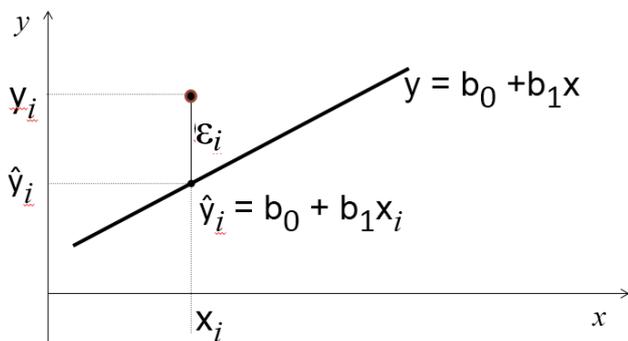
- ошибки спецификации модели;
- ошибки выборки исходных данных;
- ошибки измерения.

Выборочная линейная регрессионная модель:

$y = b_0 + b_1 x_1 + \dots + b_n x_n + \varepsilon$, где b_0, b_1, \dots, b_n – оценки неизвестных параметров обобщенной регрессионной модели, ε – случайная ошибка для выборки.

Однофакторная выборочная линейная регрессионная модель:

$y = b_0 + b_1 x + \varepsilon$, где ε – случайная ошибка (возмущение).



y_i – фактическое значение результирующего признака

\hat{y}_i – теоретическое значение результирующего признака;

ε_i – случайная ошибка.

Таким образом:

$$y_i = \hat{y}_i + \varepsilon_i = b_0 + b_1 x_i + \varepsilon_i$$

Метод наименьших квадратов (МНК) – самый распространенный метод нахождения коэффициентов линейной регрессии. Для применения метода должен быть выполнен ряд предпосылок – условия Гаусса – Маркова (см. ниже).

Мерой качества найденных оценок в методе МНК служит сумма квадратов

ошибок $\sum_{i=1}^n \varepsilon_i^2 \rightarrow \min$.

$$\sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - b_0 - b_1 x_i)^2 = g(b_0, b_1) \rightarrow \min$$

$$g'_{b_0} = -2 \sum_{i=1}^n (y_i - b_0 - b_1 x_i) = 0; \quad g'_{b_1} = -2 \sum_{i=1}^n x_i (y_i - b_0 - b_1 x_i) = 0$$

Отсюда коэффициенты уравнения линейной регрессии:

$$b_0 = \frac{1}{n} \sum y_i - b_1 \frac{1}{n} \sum x_i = \bar{y} - b_1 \bar{x}$$

$$b_1 = \frac{\sum x_i y_i - \frac{1}{n} \sum x_i \sum y_i}{\sum x_i^2 - \frac{1}{n} (\sum x_i)^2} = \frac{cov(x, y)}{D_x} = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\overline{x^2} - (\bar{x})^2}$$

Коэффициент b_1 – угловой коэффициент регрессии. Он показывает, на сколько единиц в среднем изменяется переменная y при увеличении переменной x на единицу.

Постоянная b_0 дает прогнозируемое значение переменной y при $x=0$ (если выборочные значения x находятся недалеко от $x=0$).

Коэффициент детерминации R^2 показывает степень близости корреляционной связи к точной, функциональной. Он характеризует качество уравнения регрессии и показывает, какая часть вариации результирующего признака объяснена вариацией независимого фактора:

$$R = \sqrt{1 - \frac{D_{ост}}{D(y)}} = \sqrt{1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

где $D_{ост}$ – дисперсия остаточная (необъясненная);

$D(y)$ – общая дисперсия результирующего признака.

Для R^2 верно: $0 \leq R^2 \leq 1$. Чем R^2 ближе к 1, тем в большей степени уравнение пригодно для прогнозирования.

Остатки, или отклонения, отдельных точек (наблюдений) от линии регрессии (предсказанных значений), могут быть представлены в виде:

$$(y_i - \bar{y}) = (\hat{y}_i - \bar{y}) + (y_i - \hat{y}_i), \text{ откуда:}$$

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n (y_i - \hat{y}_i)^2,$$

или

$$SST = SSR + SSE$$

где: $SSR = \sum (\hat{y}_i - \bar{y})^2$ (Sum of Squares due to Regression) – отклонение, которое можно пояснить, исходя из регрессионной прямой;

$SST = \sum (y_i - \bar{y})^2$ (Sum of Squares Total) – общее отклонение;

$SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2$ (Sum of Squares due to Error) – необъяснимое (остаточное) отклонение, которое нельзя пояснить, исходя из регрессионной прямой.

$$\text{Следовательно, } 1 = \frac{SSR}{SST} + \frac{SSE}{SST} = R^2 + \frac{SSE}{SST},$$

где $R^2 = \frac{SSR}{SST}$ – коэффициент детерминации.

Для линейной зависимости: $R^2 = r_{xy}^2$, где r_{xy} – коэффициент корреляции.

Для оценки адекватности принятой эконометрической модели экспериментальным данным используется **F-критерий Фишера (F-тест Фишера)**:

Проверка модели на адекватность с помощью F-критерия Фишера предусматривает выполнение определенных этапов:

1. Выдвигается нулевая гипотеза H_0 о том, что $R^2=0$.
2. Для оценивания истинности гипотезы рассчитывают значение F-критерия Фишера:

$$F = \frac{R^2}{1-R^2} \cdot \frac{n-m-1}{m}$$

шера:

где n —число наблюдений, m — число объясняющих факторов. Для парной регрессии $m=1$, следовательно, величина F :

$$F = \frac{R^2}{1-R^2} (n-2) = \frac{SSR}{SSE} (n-2)$$

3. Задают уровень значимости α (или $\alpha \cdot 100\%$). Например, если предполагают, что возможная ошибка α будет 0,05 (или 5%), т.е. что мы можем ошибиться не больше чем в 5 % случаев, а в 95 % случаев выводы будут правильными.
4. По статистическим таблицам F-распределения Фишера с $(m, n-m-1)$ степенями свободы и уровнем значимости α находят критическое значение $F_{табл}$.
5. Условие адекватности модели: $F > F_{табл}$.

Если рассчитанное значение $F > F_{табл}$, то гипотеза H_0 о незначимости коэффициента детерминации отвергается с риском ошибиться не больше чем в 5% случаев и признается статистическая значимость уравнения регрессии.

Поэтому если $F > F_{табл}$, то связь между рассматриваемыми признаками есть и результаты наблюдений не противоречат предположению о ее линейности. Построенная регрессионная модель адекватна реальной действительности. Если $F < F_{табл}$, то гипотеза H_0 не отвергается, уравнение регрессии считается статистически незначимым.

Проверка адекватности построенного уравнения регрессии наряду с проверкой значимости коэффициента детерминации включает и проверку значимости (значимого отличия от нуля) каждого коэффициента (параметра) регрессии.

Для проверки значимости каждого параметра регрессии b_i применяют **t-критерий Стьюдента**:

1. Выдвигается нулевая гипотеза H_0 : параметр $b_i=0$.
2. Рассчитывают значение t-критерия Стьюдента: $t_{n-m-1} = \frac{b_i}{S_{b_i}}$, где S_{b_i} — стандартная ошибка параметра b_i . Для парной регрессии $m = 1$, число степеней свободы $n-2$, следовательно,

$$i = 0, 1; S_{b_1} = \frac{S_{оцм}}{\sigma_x \sqrt{n}}; S_{b_0} = \frac{S_{оцм} \sqrt{\sum x_i^2}}{\sigma_x n}; S_{оцм}^2 = \frac{SSE}{n-2}.$$

где S_{b1} – стандартная ошибка параметра b_1 ,
 S_{b0} – стандартная ошибка параметра b_0 ,
 $S_{ост}$ – стандартная ошибка регрессии.

3. Определяют табличное значение $t_{табл}$ при определенном уровне значимости (обычно $\alpha = 0,05$) и числе степеней свободы $n - m - 1$.
4. Условие значимости параметра регрессии: $|t_b| > t_{табл}$
 Расчетное значение сравнивают с табличным. Если $|t_b| > t_{табл}$, то гипотеза H_0 отвергается, коэффициент b признается статистически значимым.
 Если $|t_b| < t_{табл}$, то коэффициент b признается статистически незначимым.

Доверительный интервал для параметров b_i : $b_i \pm S_{bi} \cdot t_{табл}$

Прогнозное значение $y_{прог}$ для парной регрессии при значении $x = x_{прог}$: $y_{прог} = b_0 + b_1 \cdot x_{прог}$

Ошибка прогноза:

$$S_y = S_{ост} \sqrt{1 + \frac{1}{n} + \frac{(x_{прог} - \bar{x})^2}{n \cdot \sigma_x^2}}$$

Доверительный интервал для прогноза $y_{прог}$: $y_{прог} \pm S_y \cdot t_{табл}$

Пригодность модели:

1. Модель по F-критерию Фишера в целом адекватна и все коэффициенты регрессии значимы \Rightarrow модель может быть использована для принятия решений и прогноза.

2. Модель по F-критерию Фишера адекватна, но часть коэффициентов регрессии незначима \Rightarrow модель пригодна для принятия некоторых решений, но не для производства прогнозов.

3. Модель по F-критерию Фишера адекватна, но все коэффициенты регрессии незначимы \Rightarrow на основе такой модели не принимаются решения и не осуществляются прогнозы.

Функции Excel линейной регрессии

Расчет параметра регрессии b_0 :	= ОТРЕЗОК(y; x)
Расчет коэффициента регрессии b_1 :	= НАКЛОН(y; x)
Расчет стандартной ошибки регрессии:	= СТОШУХ(y; x)
Расчет всех параметров регрессии, их стандартных ошибок, коэффициента детерминации, F-критерия, SSR, SSE, числа степеней свободы:	= ЛИНЕЙН(y; x; конст; статистика) (функция массива – для ввода использовать Ctrl+Shift+Enter, выделив диапазон из 5 строк и m+1 столбцов)

Функции Excel оценки значимости

Расчет табличного значения критерия Фишера $F_{табл}$:	= F.ОБР(α ; k_1 ; k_2)
Расчет значимости F-критерия (если значимость > стандартного уровня (обычно 0,05), модель адекватна, иначе – нет):	= F.РАСП(y ; x)
Расчет табличного значения критерия Стьюдента $t_{табл}$:	= СТЬЮДЕНТ.ОБР(α ; k)
Расчет значимости t-критерия Стьюдента:	= СТЬЮДЕНТ.РАСП(t ; k ; 2)

При экономической интерпретации уравнения регрессии широко используются частные **коэффициенты эластичности**, показывающие, на сколько процентов в среднем изменится значение результирующего признака при изменении значения соответствующего факторного признака на 1%.

Для парной линейной регрессии **эластичность** определяется как:

$$\varepsilon = \frac{dy/dx}{\hat{y}/x} = \frac{b_1}{\hat{y}/x} = \frac{b_1 x}{\hat{y}}.$$

Нелинейная однофакторная регрессия

Основные типы кривых:

экспоненциальная:	$y = \alpha \cdot e^x,$
степенная:	$y = \alpha \cdot x^\beta,$
гиперболическая:	$y = b_0 + b_1 \frac{1}{x},$
квадратичная:	$y = b_0 + b_1 x + b_2 x^2,$
модифицированная экспонента:	$y = \alpha \cdot \beta^x + \gamma,$
кривая Гомперца:	$y = e^{\alpha \beta^x + \gamma},$
логистическая кривая:	$y = \frac{1}{\alpha \beta^x + \gamma}.$

Формы экспоненциальной кривой:

$y = \alpha \cdot \beta^x$: основная форма, $\beta > 0$.
$y = \alpha \cdot e^{b_1 x}$: β заменяем на e^{b_1} (где $b_1 = \ln \beta$).
$y = \alpha \cdot (1-r)^x$: β заменяем на $(1-r)$ (где $r = \beta - 1$).
$y = e^{b_0 + b_1 x}$: α заменяем на e^{b_0} , β на e^{b_1} (где $b_0 = \ln \alpha$, $b_1 = \ln \beta$).
$y = 10^{b_0 + b_1 x}$: α заменяем на 10^{b_0} , β на 10^{b_1} (где $b_0 = \lg \alpha$, $b_1 = \lg \beta$).

Преобразования для экспоненциальной кривой:

$$y = \alpha \cdot \beta^x \quad : \quad \ln y = \ln \alpha + x \cdot \ln \beta$$

$$y = \alpha \cdot e^{b_1 x} \quad : \quad \ln y = \ln \alpha + x \cdot b_1$$

$$y = \alpha(1-r)^x \quad : \quad \ln y = \ln \alpha + x \ln(1-r).$$

$$y = e^{b_0 + b_1 x} \quad : \quad \ln y = b_0 + b_1 x.$$

$$y = 10^{b_0 + b_1 x} \quad : \quad \log y = b_0 + b_1 x.$$

Преобразования для степенной функции:

$$y = \alpha \cdot x^\beta \quad : \quad \ln y = \ln \alpha + \beta \ln x.$$

Преобразования для гиперболической функции:

$$y = b_0 + b_1 \frac{1}{x} \quad : \quad y = b_0 + b_1 z.$$

Преобразования для квадратичной функции:

$$y = b_0 + b_1 x + b_2 x^2 \quad : \quad y = b_0 + b_1 z_1 + b_2 z_2.$$

Многофакторная (множественная) регрессия

Многофакторная регрессия – регрессия результативного признака с двумя или большим числом факторов:

$$y = F(x_1, x_2, \dots, x_k).$$

Когда рассматривается несколько объясняющих факторов, может наблюдаться **мультиколлинеарность** – высокая корреляция между ними ($|r_{x_i, x_j}| \geq 0,7$). От этого явления необходимо избавляться увеличением размера выборки либо удалением / изменением / заменой мультиколлинеарных факторов.

Степень зависимости факторных переменных x_i и результата y отражает **коэффициент множественной корреляции R** .

Основным показателем качества модели множественной регрессии является **множественный коэффициент детерминации R^2** :

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Он показывает, какая доля изменений результата учтена в модели и обусловлена совместным влиянием включенных независимых факторов x_1, x_2, \dots, x_n . Чем ближе R^2 к единице, тем выше качество модели.

Свойство множественного коэффициента детерминации: введение нового $(n+1)$ -го фактора в регрессионную модель по крайней мере не уменьшает значение R^2 .

Поэтому, если добавить в модель много факторов, то может наблюдаться рост значения R^2 – но не за счет реального их влияния на результат y , а за счет их количества. Это может привести к ошибочному выводу о значимости влияния этих факторов на y .

Поэтому используется **скорректированный (нормированный) коэффициент детерминации \tilde{R}^2** : (или R_{adj}^2):

$$R_{adj}^2 = 1 - (1 - R^2) \frac{n - 1}{n - m - 1} = R^2 - \frac{m}{n - m - 1} (1 - R^2)$$

(n – число факторов, m – число наблюдений).

Коэффициент детерминации R^2 всегда увеличивается при добавлении в модель новых факторов, даже не оказывающих существенного влияния на y . Скорректированный (нормированный) коэффициент детерминации \tilde{R}^2 может при этом уменьшаться. Поэтому добавление факторов в модель по методу включения прекращается, когда R_{adj}^2 перестает расти.

Значимость коэффициентов множественной регрессии проверяют по t -критерию Стьюдента. Адекватность модели в целом проверяют по F -критерию Фишера:

$$F = \frac{R^2}{1 - R^2} \cdot \frac{n - m - 1}{m} \quad (n - \text{число факторов, } m - \text{число наблюдений}).$$

Условия Гаусса-Маркова – предпосылки МНК (метода наименьших квадратов):

1. $M(\varepsilon_i) = 0 (i = \overline{1, n})$

Математическое ожидание случайного отклонения ε_i равно нулю для всех наблюдений.

2. $D(\varepsilon_i) = D(\varepsilon_j) = \sigma^2 = \text{const} \quad (i, j = \overline{1, n})$

Дисперсия случайных отклонений ε_i постоянна для всех наблюдений (*гомоскедастичность*).

3. $\text{cov}(\varepsilon_i, \varepsilon_j) = M(\varepsilon_i, \varepsilon_j) = 0$ для $i \neq j$

Ковариация двух любых случайных отклонений равна нулю (отсутствует *автокорреляция*).

4. $\text{cov}(\varepsilon_i, x_i) = 0$

Случайные отклонения независимы от объясняющих переменных.

(Эта предпосылка не слишком критична для эконометрических моделей.

Если x_i – детерминированная (не случайная величина), то обычно это условие выполнено.)

5. $y = a + b \cdot x + \varepsilon$

Модель правильно специфицирована, линейна по параметрам и содержит аддитивную случайную составляющую.

Для множественной линейной регрессии необходимо также условие:

- Отсутствует *мультиколлинеарность* объясняющих переменных.

Для проверки статистических гипотез и построения интервальных оценок важно условие:

- $\varepsilon_i \sim N(0, \sigma^2)$ – случайные отклонения имеют нормальное распределение с нулевым математическим ожиданием и постоянной дисперсией.

Теорема Гаусса – Маркова. Если выполнены условия (1)–(5), то оценки, полученные методом наименьших квадратов, являются несмещенными, состоятельными и эффективными. (Эти оценки называются *BLUE–bestlinearunbiasedestimators*, наилучшие линейные несмещенные оценки.)

Приложение 2

ПРИМЕНЕНИЕ ПАКЕТА АНАЛИЗА MS EXCEL ПРИ РЕШЕНИИ ЗАДАЧ ЭКОНОМЕТРИКИ

Для подключения **надстройки *Пакет анализа*** – установить одноименный флажок в окне надстроек MS Excel (в меню *Файл / Параметры / Надстройки / Перейти*).

Пакет анализа содержит множество модулей (рис. 4).

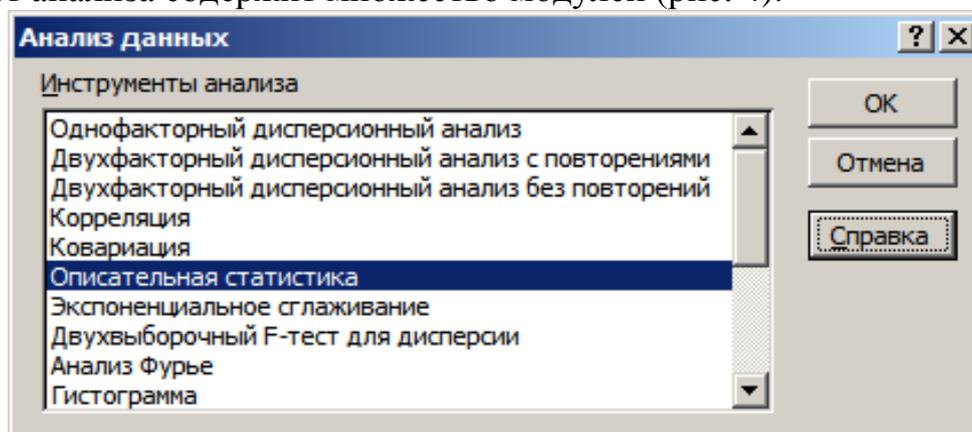


Рис. 4. Модули *Пакета анализа*

Модуль *Описательная статистика* применяется для получения статистического отчета с информацией об изменчивости входных данных (рис. 5).

Для проведения корреляционного и ковариационного анализа (расчета коэффициентов корреляции и ковариации) используются соответственно инструменты *Корреляция* и *Ковариация*. Для проведения регрессионного анализа в пакете анализа имеется инструмент *Регрессия*.

Для получения таблицы коэффициентов корреляции (или ковариации) надо выбрать в пакете анализа **модуль *Корреляция*** (или *Ковариация*) и выделить диапазон значений всех переменных (рис. 6). Желательно при этом выделять и заголовки

столбцов – тогда надо установить флажок *Метки в первой строке*. Выбрав опцию *Выходной интервал*, надо указать ячейку, начиная с которой будут размещены результаты.

Среднее	1,147142857
Стандартная ошибка	0,091488076
Медиана	1,1
Мода	#Н/Д
Стандартное отклонение	0,242054697
Дисперсия выборки	0,058590476
Эксцесс	-1,782147976
Асимметричность	0,29120209
Интервал	0,61
Минимум	0,88
Максимум	1,49
Сумма	8,03
Счет	7

Рис. 5. Результат модуля *Описательная статистика*

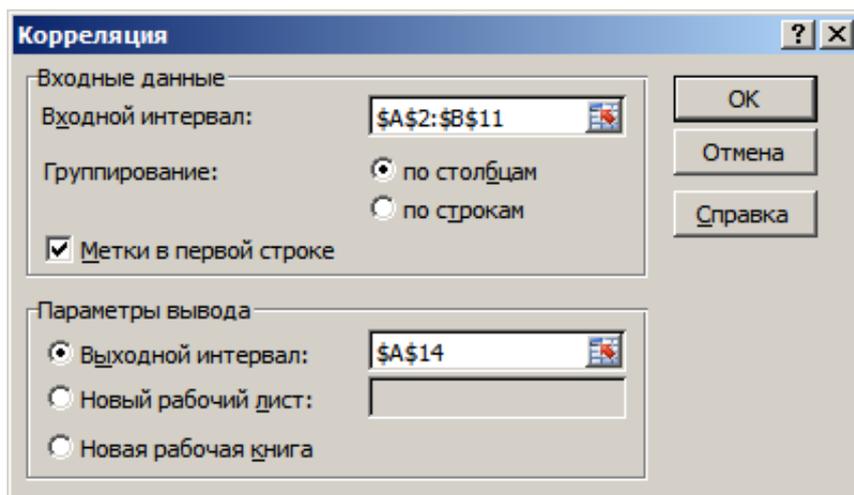


Рис. 6. Диалоговое окно модуля *Корреляция*

Результаты работы инструментов *Корреляция* и *Ковариация*:

Корреляция:		Ковариация:	
	X	Y	
X	1		дисперсия X
Y	0,921407556	1	дисперсия Y
			коэф. ковариации X и Y

Рис. 7. Результат применения инструментов *Корреляция* и *Ковариация*

Для построения линейной модели вида $y = b_0 + b_1x$ служит модуль *Регрессия*. В его диалоговом окне надо отдельно выделить диапазоны значений

результативного признака Y и одного или нескольких независимых признаков X . Желательно при этом выделять и заголовки столбцов (тогда надо установить флажок *Метки в первой строке*). Выбрав опцию *Выходной интервал*, надо указать ячейку, начиная с которой будут размещены результаты.

Диалоговое окно *Регрессия* (рис. 8) содержит несколько разделов. В разделе *Входные данные* указываются такие параметры:

- *Входной интервал Y* – диапазон с данными результативного признака;
- *Входной интервал X* – диапазон с данными факторов;
- *Метки* – флажок, указывающий, что первые строки диапазонов содержат названия столбцов;
- *Константа – ноль* – флажок, указывающий на отсутствие свободного члена в уравнении.

В параметрах раздела *Параметры вывода* указывают:

- *Выходной интервал* – левую верхнюю ячейку будущего диапазона результатов (на текущем листе);
- *Новый рабочий лист* – произвольное имя нового листа с результатами;
- *Новая рабочая книга* – имя новой рабочей книги с результатами.

Раздел *Остатки* содержит флажки, которые следует установить, если необходимо получить значения и графики остатков либо график нормального распределения вероятности.

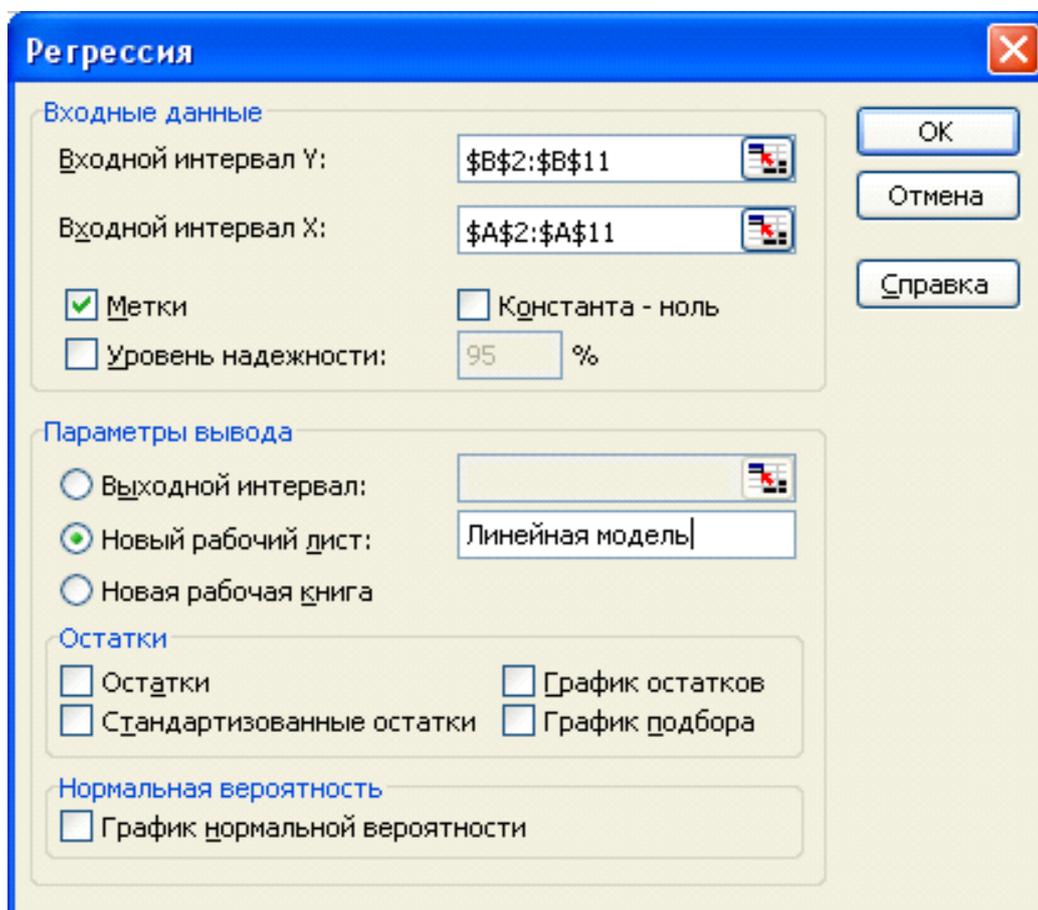


Рис. 8. Диалоговое окно модуля *Регрессия*

Результаты регрессионного анализа представлены на рис. 9. Они состоят из 3-х блоков-таблиц.

Первый блок – *Регрессионная статистика*:

- *Множественный $R(r)$* – это коэффициент корреляции. (На рис. $r = 0,924$ показывает, что связь между Y и фактором X высокая, то есть степень зависимости Y от X достаточно велика.)
- *R -квадрат (R^2)* – коэффициент детерминации. (На рис. $R^2 = 0,854$, что означает, что ~85% изменений Y связано с линейным влиянием X .)
- *Стандартная ошибка* – ошибка, возникающая при замене фактических наблюдений Y_i оценками \hat{Y}_i , рассчитываемых по линейному уравнению регрессии.

Вторая таблица в выводе итогов модуля *Регрессии* – это *Дисперсионный анализ*. В ней рассчитаны F -критерий Фишера и его значимость. Если значимость $F > \alpha$ (уровень значимости α – вероятность совершить ошибку при принятии гипотезы), то принимают гипотезу о том, что уравнение регрессии не значимо. В противном случае эта гипотеза отбрасывается, и уравнение регрессии будет значимо.

Если значимость F меньше 0,01, полученный результат является высоко значимым. Если значимость F от 0,01 до 0,05 – результат значим; больше 0,05 – результат значимым не является. Обычно уровень значимости принимают $\alpha = 0,05$.

На рис. 9 значимость $F = 0,00037 < 0,05$. Поэтому уравнение регрессии значимо, т.е. **модель адекватна**.

В третьем блоке-таблице отображены оценки параметров (коэффициентов) уравнения регрессии. (На рис. 9 в столбце параметров отображены $b_0 = 1,98797$; $b_1 = 0,16373$.)

Для каждого параметра рядом рассчитаны t -статистика и ее значимость (*P-значения*). Если *P-значение* $< \alpha$ (уровень значимости), то соответствующий **коэффициент значим**. Обычно уровень значимости принимают $\alpha = 0,05$. (Оба параметра на рис. 9 значимы, следовательно, уравнение регрессии будет иметь вид $y = 1,98797427 + 0,16372990 \cdot x$.)

Если какой-либо коэффициент окажется незначимым, следует исключить соответствующую ему переменную из формулы. Для этого надо снова вызвать модуль *Регрессия* пакета анализа и выделить входной диапазон X с меньшим числом факторов.

Если незначимым окажется свободный член уравнения – коэффициент b_0 (*Y-пересечение*), надо также вызвать построение регрессии повторно и в ее диалоговом окне установить флажок *Константа ноль* (см. рис.8).

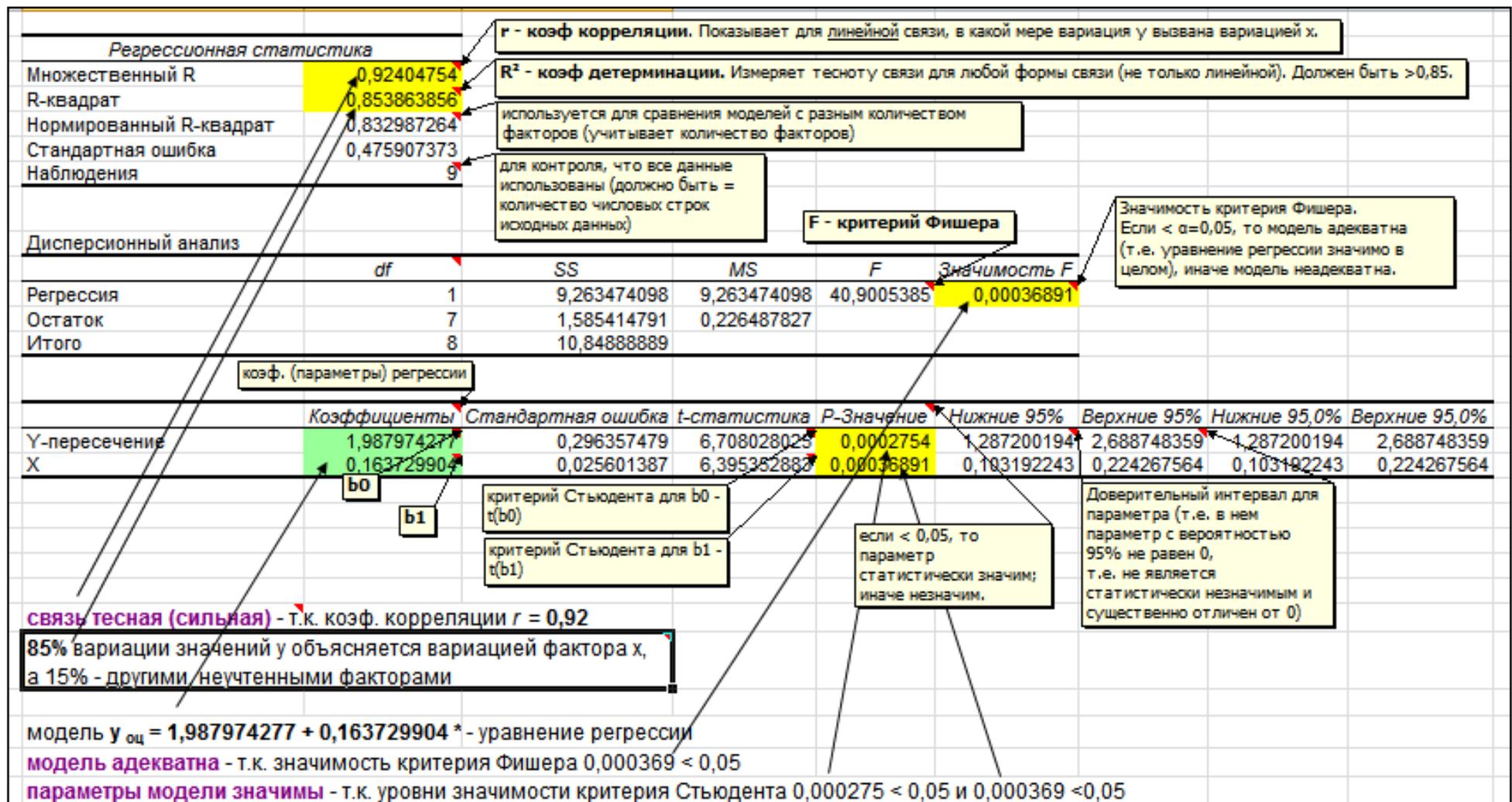


Рис. 9. Результат применения инструмента Регрессия

Приложение 3

Валовый внутренний продукт (ВВП) и потребительские расходы (ПР), 1970–2013³
(млрд долл., в ценах 1970 г.)

Год	Вариант 1		Вариант 2		Вариант 3		Вариант 4		Вариант 5		Вариант 6	
	США		Китай		Япония		Германия		Великобритания		Франция	
	ВВП	ПР	ВВП	ПР	ВВП	ПР	ВВП	ПР	ВВП	ПР	ВВП	ПР
1970	1075,9	901,9	91	59,3	209,1	124	215	154,8	129,7	100,1	148,9	106,4
1971	1167,8	923,4	98,1	62,9	236,2	134,1	249	162,6	146,7	104,5	166,6	112,7
1972	1282,4	964,8	111,6	66,4	312,7	147	298,7	171,8	168,6	110,4	204,3	117,6
1973	1428,5	1001,4	136,1	70,6	424,9	158,4	396,9	180,6	190,3	117	265,4	124,2
1974	1548,8	1012,2	141,5	72,8	471,6	161,1	443,6	185,1	205,6	118,2	286,5	131,1
1975	1688,9	1028,6	160,3	76,8	512,9	176,3	488,8	192,2	245,7	116,6	362	133,7
1976	1877,6	1071,6	150,9	80,3	576,4	183,5	517,8	200,1	236,1	118,4	373,4	140,1
1977	2086	1110,5	171,5	83	709,4	192	598,2	207,5	266,7	117,8	411,5	145,1
1978	2356,6	1153,5	214,2	89,1	996,7	201,4	737,7	212,4	337,9	122,1	508,2	151
1979	2632,1	1183,9	263,2	99,3	1037,5	215,6	878	221,9	439,2	126,9	615,9	156,6
1980	2862,5	1204,2	306,5	109	1087	222,6	946,7	228,5	563,3	125,8	703,5	161,3
1981	3211	1217,3	293,9	117,6	1201,5	230	797,4	232,9	535,3	126,9	617,6	167,7
1982	3345	1230,9	295,4	127	1116,8	242,6	773,6	230,8	510,3	129,2	586,8	173,4
1983	3638,1	1298,3	314,6	140,6	1218,1	253,4	767,8	234	484	134,8	561,9	175,3
1984	4040,7	1365,2	317,4	160,6	1294,6	261,6	722,4	239,4	456,1	137,5	532,7	178,3
1985	4346,7	1445,4	309,1	182,5	1384,5	273,2	729,8	242,9	482,6	141	555,2	182,5
1986	4590,2	1511,1	304,3	195,6	2051,1	279,8	1042,3	243,4	592,9	148,9	774,6	184,8
1987	4870,2	1568,7	329,9	213,7	2485,2	292,9	1293,3	248,9	728,2	156	938,3	190,7

³<http://www.be5.biz/makroekonomika/index.html>

Год	Вариант 1		Вариант 2		Вариант 3		Вариант 4		Вариант 5		Вариант 6	
	США		Китай		Япония		Германия		Великобритания		Франция	
1988	5252,6	1626,7	413,4	239,2	3015,4	307,6	1395,9	255,7	884,8	165,4	1023,5	196
1989	5657,7	1678,9	459,8	251	3017,1	321,5	1393,7	262,6	893,4	169	1030,1	202,1
1990	5979,6	1728	404,5	252,6	3103,7	338,8	1764,9	272,1	1059,6	168,8	1275,3	207,9
1991	6174	1736,5	424,1	275,4	3536,8	348	1861,9	283,3	1108,5	169,5	1275,5	211
1992	6539,3	1795,5	499,9	314,6	3852,8	356,9	2123,2	291,3	1149,9	172,5	1408,7	216,2
1993	6878,7	1840,9	641,1	340,6	4415	364,2	2068,5	291,3	1036,2	177,2	1330,2	218,8
1994	7308,8	1897,5	582,7	378,3	4850,3	374,8	2205,9	297,1	1122,6	182	1401,6	222,3
1995	7664,1	1946,9	757	418,9	5333,9	385,2	2590,3	301,5	1235,6	183	1609,8	226,1
1996	8100,2	2009,5	892	469,5	4706,2	397,4	2501,6	306,8	1304,6	187,6	1614,3	231,3
1997	8608,5	2078	985	510,8	4324,3	403,3	2215,9	310,9	1438,7	193,8	1460,7	233,7
1998	9089,2	2171,8	1045,2	557,1	3914,6	401,8	2239,9	314,1	1529,1	202,5	1510,8	239,1
1999	9660,6	2286	1100,8	614	4432,6	410,2	2196,6	322,4	1558,4	212,6	1500,2	246,3
2000	10284,8	2400,4	1192,8	679,3	4731,2	418,1	1947,2	331,9	1548,6	220,6	1368,4	255,8
2001	10621,8	2465,3	1317,2	725	4159,9	428,8	1947,9	337,9	1529,2	229,1	1382,2	261,3
2002	10977,5	2539,7	1455,6	768	3980,8	436,6	2076,4	336,9	1674,4	236,6	1500,4	265,7
2003	11510,7	2623,3	1650,5	806	4302,9	442,5	2502,2	338,7	1943,8	246,7	1848,1	270,4
2004	12274,9	2714,5	1944,7	848,8	4655,8	450,5	2815,6	339,1	2298,1	254,6	2124,2	277,1
2005	13093,7	2795,9	2287,2	920,7	4571,9	459,9	2857,6	342,9	2412,1	261,9	2203,6	283,2
2006	13855,9	2870,8	2793,2	993,8	4356,7	467,3	2998,5	350,2	2582,8	268,3	2324,9	288,5
2007	14477,6	2938,3	3504,4	1108,5	4356,3	473,6	3435,5	350,3	2963,3	274,3	2663	293,5
2008	14718,6	2988,4	4547,3	1189,7	4849,2	477,4	3747,1	356,5	2791,9	277,6	2923,6	295,9
2009	14418,7	2947,5	5105,5	1298,7	5035,1	469,4	3412,8	353,6	2308,9	272	2693,7	296,3
2010	14964,4	3011	5949,8	1423,8	5495,4	485,2	3412	360,2	2407,9	274,3	2646,8	301,6
2011	15517,9	3056,3	7314,4	1585,9	5905,6	493,5	3752,1	369,6	2591,8	275,7	2862,7	305,5
2012	16163,2	3087,8	8229,4	1752,2	5937,9	503,7	3533,2	372,8	2614,9	279,1	2686,7	306,1

Год	Вариант 7		Вариант 8		Вариант 9		Вариант 10		Вариант 11		Вариант 12	
	Бразилия		Италия		Канада		Индия		Австралия		Испания	
	ВВП	ПР	ВВП	ПР	ВВП	ПР	ВВП	ПР	ВВП	ПР	ВВП	ПР
1970	35,2	28,1	113,1	84	87,8	66,9	61,5	55	45,2	30,6	39,8	29,1
1971	39,3	31,6	124,3	87,3	99,1	69,7	65,9	56,9	50,5	32,2	45,2	30,6
1972	50,6	35,3	144,8	91,6	112,9	73,5	71,7	56,3	59,4	33	57,4	32,8
1973	71,7	39,6	175	97,2	131,1	76,5	85,5	56,8	85,8	33,8	76,3	35,2
1974	91,9	44	199	100,7	160,2	78,3	96,6	58,8	102,3	36,6	94,4	37,8
1975	108,1	44,4	227	101,2	173,6	82,9	100,4	62	109,1	38,1	111,3	38,3
1976	127,7	49,8	224,1	106,2	206,3	86,4	101,2	61,3	117,5	39,8	115	40,8
1977	148,7	51,7	256,8	108,9	211,3	90,6	117,4	66	116,5	40,9	128,5	41,7
1978	168,8	54	314,1	111,4	218,3	94,2	135,8	70,1	135,9	41,7	155,8	41,9
1979	186,1	58,2	392,5	118,9	242,7	95,3	150,3	66	150,4	42,4	208,2	42,5
1980	191,1	67,1	475,8	125,4	273,4	97,6	184,8	71,7	173,5	44,3	225,8	44,5
1981	225,6	58,6	429,4	128,8	305,7	99,8	197,1	74,9	202,1	45,8	196,7	45,3
1982	238,8	61	426	129,7	313,4	100	201,2	77,5	192,1	47,1	190,1	45,8
1983	170,4	61,3	441,7	130,6	340,3	103,1	219,6	83,8	192,6	47,7	165,8	46,4
1984	175,5	62,7	436,6	134,6	353,9	107,3	217,5	86,7	206,6	49,5	167	46,3
1985	187,4	64,9	450,9	138,6	363	113,6	226,5	89,4	182,2	52,3	175,4	47,4
1986	225,3	72,4	638,5	142	375,7	118,9	248,1	95	191,6	53,9	243,8	48,3
1987	247,1	71,1	803,4	147,5	429	122,6	274,6	98,8	228	55,1	309,2	51,3
1988	276	68,7	888,9	153,5	506	126,8	303,8	105,7	288,7	56,1	364,9	53,3
1989	375,5	70,8	925,9	159,3	563,6	131,1	300,7	109,7	320	59	402,4	56,4
1990	402,1	74,6	1177,8	161,5	592	135	326,8	113,2	323,8	61,5	520,5	58,4
1991	374,2	79,5	1242,5	164,8	608,2	137,2	289,7	113,6	329,6	63,8	559,9	60,4
1992	358,4	78,3	1316,3	167,8	590,1	139,8	290,9	118	326,1	65,8	612	62,2
1993	402,2	81,1	1061,8	165	574,8	142,2	284	123,6	317,3	67,6	509,4	61,9

Год	Вариант 7		Вариант 8		Вариант 9		Вариант 10		Вариант 11		Вариант 12	
	Бразилия		Италия		Канада		Индия		Австралия		Испания	
	ВВП	ПР	ВВП	ПР	ВВП	ПР	ВВП	ПР	ВВП	ПР	ВВП	ПР
1994	573,7	85,6	1096	168,2	576	144,8	325,3	129,7	362,6	70	514,7	62,8
1995	769	91,5	1171,3	169,9	602	145,5	369,2	137,8	392,3	72,4	596,2	63,8
1996	839,7	94,9	1309,5	171,7	627	147,3	389,2	150,7	435,8	74,5	622,1	65,1
1997	871,2	98,1	1239,5	176,4	651	153,1	422,6	155,9	437,2	78,2	572,5	67
1998	843,8	98,4	1266,8	180,4	631,4	159,8	425,3	169,7	389,8	83,1	600,7	69,7
1999	587,1	98,7	1249	184,9	674,3	165,2	453,4	184,1	426,6	85,8	617,6	73
2000	644,7	101,1	1142,2	192,3	739,5	169,5	467,8	189,4	409,1	88,2	580,3	76,8
2001	554,2	102,1	1162,8	194,8	732,7	174,2	483	200,4	389,9	90,7	608,9	78,9
2002	506	103,6	1267,1	194,7	752,5	181,7	504,9	203,4	435,3	94,2	686,3	80,3
2003	552,4	103,5	1570,3	196,6	887,8	185	591,3	214,3	559	97,4	883,8	82,2
2004	663,7	106,3	1799,2	199,2	1018,4	186,9	715,5	222,9	678,6	100,1	1044,6	85,7
2005	882	111,3	1853,5	203	1164,2	190,9	837,5	240,4	762,4	101,5	1130,8	88,9
2006	1089,3	116	1943,4	206,9	1310,8	196,3	947,9	258,2	819	104,8	1236,4	92,1
2007	1366,9	122,7	2204	207,9	1457,9	201,8	1206,1	280,4	986,5	108,3	1441,4	95,7
2008	1653,5	127,4	2392	207,9	1542,6	204,7	1294,1	297,3	1055,8	107,8	1593,4	97,8
2009	1620,2	132,1	2186,1	202,1	1370,8	213,2	1338,2	324,9	1011	112,2	1454,3	95,6
2010	2143	139,5	2126,6	206	1614,1	217	1704,8	349,3	1290,3	112,3	1384,8	97,1
2011	2476,7	143,6	2278,2	206,4	1788,7	219,1	1930,5	378,2	1532,9	116,1	1454,5	97,7
2012	2248,8	150,4	2091,8	200,3	1832,7	222,9	1892,6	398,3	1578	121,4	1322,5	95,7

Приложение 4
F-распределение Фишера
(уровень значимости $\alpha = 0,05$)

df2/df1	1	2	3	4	5	6	7	8	9	10
1	161,448	199,5	215,707	224,583	230,162	233,986	236,768	238,883	240,543	241,882
2	18,5128	19	19,1643	19,2468	19,2964	19,3295	19,3532	19,371	19,3848	19,3959
3	10,128	9,5521	9,2766	9,1172	9,0135	8,9406	8,8867	8,8452	8,8123	8,7855
4	7,7086	6,9443	6,5914	6,3882	6,2561	6,1631	6,0942	6,041	5,9988	5,9644
5	6,6079	5,7861	5,4095	5,1922	5,0503	4,9503	4,8759	4,8183	4,7725	4,7351
6	5,9874	5,1433	4,7571	4,5337	4,3874	4,2839	4,2067	4,1468	4,099	4,06
7	5,5914	4,7374	4,3468	4,1203	3,9715	3,866	3,787	3,7257	3,6767	3,6365
8	5,3177	4,459	4,0662	3,8379	3,6875	3,5806	3,5005	3,4381	3,3881	3,3472
9	5,1174	4,2565	3,8625	3,6331	3,4817	3,3738	3,2927	3,2296	3,1789	3,1373
10	4,9646	4,1028	3,7083	3,478	3,3258	3,2172	3,1355	3,0717	3,0204	2,9782
11	4,8443	3,9823	3,5874	3,3567	3,2039	3,0946	3,0123	2,948	2,8962	2,8536
12	4,7472	3,8853	3,4903	3,2592	3,1059	2,9961	2,9134	2,8486	2,7964	2,7534
13	4,6672	3,8056	3,4105	3,1791	3,0254	2,9153	2,8321	2,7669	2,7144	2,671
14	4,6001	3,7389	3,3439	3,1122	2,9582	2,8477	2,7642	2,6987	2,6458	2,6022
15	4,5431	3,6823	3,2874	3,0556	2,9013	2,7905	2,7066	2,6408	2,5876	2,5437
16	4,494	3,6337	3,2389	3,0069	2,8524	2,7413	2,6572	2,5911	2,5377	2,4935
17	4,4513	3,5915	3,1968	2,9647	2,81	2,6987	2,6143	2,548	2,4943	2,4499
18	4,4139	3,5546	3,1599	2,9277	2,7729	2,6613	2,5767	2,5102	2,4563	2,4117
19	4,3807	3,5219	3,1274	2,8951	2,7401	2,6283	2,5435	2,4768	2,4227	2,3779
20	4,3512	3,4928	3,0984	2,8661	2,7109	2,599	2,514	2,4471	2,3928	2,3479
21	4,3248	3,4668	3,0725	2,8401	2,6848	2,5727	2,4876	2,4205	2,366	2,321
22	4,3009	3,4434	3,0491	2,8167	2,6613	2,5491	2,4638	2,3965	2,3419	2,2967

df2/df1	1	2	3	4	5	6	7	8	9	10
23	4,2793	3,4221	3,028	2,7955	2,64	2,5277	2,4422	2,3748	2,3201	2,2747
24	4,2597	3,4028	3,0088	2,7763	2,6207	2,5082	2,4226	2,3551	2,3002	2,2547
25	4,2417	3,3852	2,9912	2,7587	2,603	2,4904	2,4047	2,3371	2,2821	2,2365
26	4,2252	3,369	2,9752	2,7426	2,5868	2,4741	2,3883	2,3205	2,2655	2,2197
27	4,21	3,3541	2,9604	2,7278	2,5719	2,4591	2,3732	2,3053	2,2501	2,2043
28	4,196	3,3404	2,9467	2,7141	2,5581	2,4453	2,3593	2,2913	2,236	2,19
29	4,183	3,3277	2,934	2,7014	2,5454	2,4324	2,3463	2,2783	2,2229	2,1768
30	4,1709	3,3158	2,9223	2,6896	2,5336	2,4205	2,3343	2,2662	2,2107	2,1646
40	4,0847	3,2317	2,8387	2,606	2,4495	2,3359	2,249	2,1802	2,124	2,0772
60	4,0012	3,1504	2,7581	2,5252	2,3683	2,2541	2,1665	2,097	2,0401	1,9926
120	3,9201	3,0718	2,6802	2,4472	2,2899	2,175	2,0868	2,0164	1,9588	1,9105
INF	3,8415	2,9957	2,6049	2,3719	2,2141	2,0986	2,0096	1,9384	1,8799	1,8307

df2/df1	12	15	20	24	30	40	60	120	INF
1	243,906	245,95	248,013	249,052	250,095	251,143	252,196	253,253	254,314
2	19,4125	19,4291	19,4458	19,4541	19,4624	19,4707	19,4791	19,4874	19,4957
3	8,7446	8,7029	8,6602	8,6385	8,6166	8,5944	8,572	8,5494	8,5264
4	5,9117	5,8578	5,8025	5,7744	5,7459	5,717	5,6877	5,6581	5,6281
5	4,6777	4,6188	4,5581	4,5272	4,4957	4,4638	4,4314	4,3985	4,365
6	3,9999	3,9381	3,8742	3,8415	3,8082	3,7743	3,7398	3,7047	3,6689
7	3,5747	3,5107	3,4445	3,4105	3,3758	3,3404	3,3043	3,2674	3,2298
8	3,2839	3,2184	3,1503	3,1152	3,0794	3,0428	3,0053	2,9669	2,9276
9	3,0729	3,0061	2,9365	2,9005	2,8637	2,8259	2,7872	2,7475	2,7067
10	2,913	2,845	2,774	2,7372	2,6996	2,6609	2,6211	2,5801	2,5379
11	2,7876	2,7186	2,6464	2,609	2,5705	2,5309	2,4901	2,448	2,4045

df2/df1	12	15	20	24	30	40	60	120	INF
12	2,6866	2,6169	2,5436	2,5055	2,4663	2,4259	2,3842	2,341	2,2962
13	2,6037	2,5331	2,4589	2,4202	2,3803	2,3392	2,2966	2,2524	2,2064
14	2,5342	2,463	2,3879	2,3487	2,3082	2,2664	2,2229	2,1778	2,1307
15	2,4753	2,4034	2,3275	2,2878	2,2468	2,2043	2,1601	2,1141	2,0658
16	2,4247	2,3522	2,2756	2,2354	2,1938	2,1507	2,1058	2,0589	2,0096
17	2,3807	2,3077	2,2304	2,1898	2,1477	2,104	2,0584	2,0107	1,9604
18	2,3421	2,2686	2,1906	2,1497	2,1071	2,0629	2,0166	1,9681	1,9168
19	2,308	2,2341	2,1555	2,1141	2,0712	2,0264	1,9795	1,9302	1,878
20	2,2776	2,2033	2,1242	2,0825	2,0391	1,9938	1,9464	1,8963	1,8432
21	2,2504	2,1757	2,096	2,054	2,0102	1,9645	1,9165	1,8657	1,8117
22	2,2258	2,1508	2,0707	2,0283	1,9842	1,938	1,8894	1,838	1,7831
23	2,2036	2,1282	2,0476	2,005	1,9605	1,9139	1,8648	1,8128	1,757
24	2,1834	2,1077	2,0267	1,9838	1,939	1,892	1,8424	1,7896	1,733
25	2,1649	2,0889	2,0075	1,9643	1,9192	1,8718	1,8217	1,7684	1,711
26	2,1479	2,0716	1,9898	1,9464	1,901	1,8533	1,8027	1,7488	1,6906
27	2,1323	2,0558	1,9736	1,9299	1,8842	1,8361	1,7851	1,7306	1,6717
28	2,1179	2,0411	1,9586	1,9147	1,8687	1,8203	1,7689	1,7138	1,6541
29	2,1045	2,0275	1,9446	1,9005	1,8543	1,8055	1,7537	1,6981	1,6376
30	2,0921	2,0148	1,9317	1,8874	1,8409	1,7918	1,7396	1,6835	1,6223
40	2,0035	1,9245	1,8389	1,7929	1,7444	1,6928	1,6373	1,5766	1,5089
60	1,9174	1,8364	1,748	1,7001	1,6491	1,5943	1,5343	1,4673	1,3893
120	1,8337	1,7505	1,6587	1,6084	1,5543	1,4952	1,429	1,3519	1,2539
INF	1,7522	1,6664	1,5705	1,5173	1,4591	1,394	1,318	1,2214	1

Приложение 5
Распределение Стьюдента

df	0.2	0.1	0.05	0.04	0.03	0.025	0.02	0.01	0.005	0.0005
1	1.376	3.078	6.314	7.916	10.579	12.706	15.895	31.821	63.657	636.619
2	1.061	1.886	2.920	3.320	3.896	4.303	4.849	6.965	9.925	31.599
3	0.978	1.638	2.353	2.605	2.951	3.182	3.482	4.541	5.841	12.924
4	0.941	1.533	2.132	2.333	2.601	2.776	2.999	3.747	4.604	8.610
5	0.920	1.476	2.015	2.191	2.422	2.571	2.757	3.365	4.032	6.869
6	0.906	1.440	1.943	2.104	2.313	2.447	2.612	3.143	3.707	5.959
7	0.896	1.415	1.895	2.046	2.241	2.365	2.517	2.998	3.499	5.408
8	0.889	1.397	1.860	2.004	2.189	2.306	2.449	2.896	3.355	5.041
9	0.883	1.383	1.833	1.973	2.150	2.262	2.398	2.821	3.250	4.781
10	0.879	1.372	1.812	1.948	2.120	2.228	2.359	2.764	3.169	4.587
11	0.876	1.363	1.796	1.928	2.096	2.201	2.328	2.718	3.106	4.437
12	0.873	1.356	1.782	1.912	2.076	2.179	2.303	2.681	3.055	4.318
13	0.870	1.350	1.771	1.899	2.060	2.160	2.282	2.650	3.012	4.221
14	0.868	1.345	1.761	1.887	2.046	2.145	2.264	2.624	2.977	4.140
15	0.866	1.341	1.753	1.878	2.034	2.131	2.249	2.602	2.947	4.073
16	0.865	1.337	1.746	1.869	2.024	2.120	2.235	2.583	2.921	4.015
17	0.863	1.333	1.740	1.862	2.015	2.110	2.224	2.567	2.898	3.965
18	0.862	1.330	1.734	1.855	2.007	2.101	2.214	2.552	2.878	3.922
19	0.861	1.328	1.729	1.850	2.000	2.093	2.205	2.539	2.861	3.883
20	0.860	1.325	1.725	1.844	1.994	2.086	2.197	2.528	2.845	3.850
21	0.859	1.323	1.721	1.840	1.988	2.080	2.189	2.518	2.831	3.819
22	0.858	1.321	1.717	1.835	1.983	2.074	2.183	2.508	2.819	3.792
23	0.858	1.319	1.714	1.832	1.978	2.069	2.177	2.500	2.807	3.768
24	0.857	1.318	1.711	1.828	1.974	2.064	2.172	2.492	2.797	3.745
25	0.856	1.316	1.708	1.825	1.970	2.060	2.167	2.485	2.787	3.725
26	0.856	1.315	1.706	1.822	1.967	2.056	2.162	2.479	2.779	3.707
27	0.855	1.314	1.703	1.819	1.963	2.052	2.158	2.473	2.771	3.690
28	0.855	1.313	1.701	1.817	1.960	2.048	2.154	2.467	2.763	3.674
29	0.854	1.311	1.699	1.814	1.957	2.045	2.150	2.462	2.756	3.659
30	0.854	1.310	1.697	1.812	1.955	2.042	2.147	2.457	2.750	3.646

df	0.2	0.1	0.05	0.04	0.03	0.025	0.02	0.01	0.005	0.0005
31	0.853	1.309	1.696	1.810	1.952	2.040	2.144	2.453	2.744	3.633
32	0.853	1.309	1.694	1.808	1.950	2.037	2.141	2.449	2.738	3.622
33	0.853	1.308	1.692	1.806	1.948	2.035	2.138	2.445	2.733	3.611
34	0.852	1.307	1.691	1.805	1.946	2.032	2.136	2.441	2.728	3.601
35	0.852	1.306	1.690	1.803	1.944	2.030	2.133	2.438	2.724	3.591
36	0.852	1.306	1.688	1.802	1.942	2.028	2.131	2.434	2.719	3.582
37	0.851	1.305	1.687	1.800	1.940	2.026	2.129	2.431	2.715	3.574
38	0.851	1.304	1.686	1.799	1.939	2.024	2.127	2.429	2.712	3.566
39	0.851	1.304	1.685	1.798	1.937	2.023	2.125	2.426	2.708	3.558
40	0.851	1.303	1.684	1.796	1.936	2.021	2.123	2.423	2.704	3.551
41	0.850	1.303	1.683	1.795	1.934	2.020	2.121	2.421	2.701	3.544
42	0.850	1.302	1.682	1.794	1.933	2.018	2.120	2.418	2.698	3.538
43	0.850	1.302	1.681	1.793	1.932	2.017	2.118	2.416	2.695	3.532
44	0.850	1.301	1.680	1.792	1.931	2.015	2.116	2.414	2.692	3.526
45	0.850	1.301	1.679	1.791	1.929	2.014	2.115	2.412	2.690	3.520
46	0.850	1.300	1.679	1.790	1.928	2.013	2.114	2.410	2.687	3.515
47	0.849	1.300	1.678	1.789	1.927	2.012	2.112	2.408	2.685	3.510
48	0.849	1.299	1.677	1.789	1.926	2.011	2.111	2.407	2.682	3.505
49	0.849	1.299	1.677	1.788	1.925	2.010	2.110	2.405	2.680	3.500
50	0.849	1.299	1.676	1.787	1.924	2.009	2.109	2.403	2.678	3.496
60	0.848	1.296	1.671	1.781	1.917	2.000	2.099	2.390	2.660	3.460
70	0.847	1.294	1.667	1.776	1.912	1.994	2.093	2.381	2.648	3.435
80	0.846	1.292	1.664	1.773	1.908	1.990	2.088	2.374	2.639	3.416
90	0.846	1.291	1.662	1.771	1.905	1.987	2.084	2.368	2.632	3.402
100	0.845	1.290	1.660	1.769	1.902	1.984	2.081	2.364	2.626	3.390
120	0.845	1.289	1.658	1.766	1.899	1.980	2.076	2.358	2.617	3.373
140	0.844	1.288	1.656	1.763	1.896	1.977	2.073	2.353	2.611	3.361
180	0.844	1.286	1.653	1.761	1.893	1.973	2.069	2.347	2.603	3.345
200	0.843	1.286	1.653	1.760	1.892	1.972	2.067	2.345	2.601	3.340
500	0.842	1.283	1.648	1.754	1.885	1.965	2.059	2.334	2.586	3.310
1000	0.842	1.282	1.646	1.752	1.883	1.962	2.056	2.330	2.581	3.300

Приложение 6

Граничные значения для статистик Дарбина – Уотсона
(уровень значимости $\alpha = 0,05$)

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

Приложение 7

ГЛОССАРИЙ

Автокорреляция – корреляция между временной переменной и лаговой переменной, составленной от той же переменной.

Авторегрессия – регрессия зависимой временной переменной от лаговой переменной, составленной от той же переменной.

Временной ряд (ряд динамики, динамический ряд) – упорядоченная во времени последовательность численных, характеризующих уровни развития изучаемого явления в последовательные моменты или периоды времени.

Гетероскедастичность – неоднородность относительно дисперсии.

Гомоскедастичность – однородность относительно дисперсии.

Корреляционная связь – зависимость, при которой изменение среднего значения результативного признака обусловлено изменением факторных признаков.

Корреляционный анализ – раздел математической статистики, изучающий наличие и тесноту связи между переменными без разделения переменных на зависимые и объясняющие.

Коэффициент детерминации R^2 – представляет собой квадрат коэффициента корреляции, показывает, какая часть дисперсии результативного признака объяснена уравнением регрессии.

Коэффициент эластичности \mathcal{E} – показывает, на сколько процентов в среднем изменится результативный признак при изменении фактора на 1 % от своего номинального значения.

Критерий Стьюдента (t-критерий) – применяется для оценки статистической значимости коэффициентов полученного уравнения регрессии.

Критерий Фишера (F-критерий) – применяется для оценки статистической значимости полученного уравнения регрессии в целом.

Лаг – задержка. Обычно лаг рассматривают между причиной и следствием.

Линейная аналитическая зависимость – зависимость, определяемая соотношением $y = a + b_1 \cdot x_1 + b_2 \cdot x_2 + \dots + b_n \cdot x_n + \varepsilon$.

Линейный коэффициент корреляции – характеризует тесноту линейной связи

между переменными и равен
$$r_{xy} = \frac{\frac{1}{n} \sum_i (x_i + \bar{x})(y_i + \bar{y})}{\sigma_x \sigma_y} = \frac{\text{cov}(x, y)}{\sigma_x \sigma_y}.$$

Ложная корреляция – высокое значение коэффициентов корреляции при отсутствии истинной зависимости между переменными.

Метод наименьших квадратов (МНК) – метод оценки параметров уравнения регрессии, основанный на минимизации суммы квадратов отклонений расчетных значений (по уравнению регрессии) зависимой переменной от ее наблюдаемых значений.

Мультиколлинеарность – множественная линейная связь.

- Показательная аналитическая зависимость** – зависимость, определяемая соотношением $\hat{y} = a \cdot b^x$.
- Поле корреляций** – совокупность точек на координатной плоскости, изображающих наблюдения.
- Предмет эконометрики** – количественная оценка взаимосвязи между случайными событиями, признаками, показателями, факторами переменных экономических объектов.
- Регрессионный анализ** – раздел математической статистики, изучающий форму зависимости характеристик стохастического процесса от одного или нескольких факторов.
- Регрессия** – зависимость среднего значения какой-либо величины от некоторой другой величины или от нескольких величин.
- Регрессия множественная** – зависимость с несколькими зависимыми переменными.
- Регрессия парная** – зависимость между двумя переменными.
- Результативный признак** – признак, изменяющийся под действием факторных признаков
- Степенная аналитическая зависимость** – зависимость, определяемая соотношением $y = a \cdot x_1^{b_1} \cdot x_2^{b_2} \cdot \dots \cdot x_n^{b_n} \cdot \varepsilon$.
- Тенденция** – изменение, определяющее общее направление развития, основную тенденцию временного ряда.
- Тест Голдфельда–Квандта** – применяется для проверки остатков регрессии на гетероскедастичность
- Тест ранговой корреляции Спирмена** – применяется для проверки остатков регрессии на гетероскедастичность. Проверяет наличие монотонной зависимости между дисперсией ошибки и величиной фактора.
- Тренд** – изменение, определяющее общее направление развития, основную тенденцию временного ряда.
- Факторный признак** – признак, обуславливающий изменение другого, связанного с ним признака.
- Фиктивные переменные** – применяются для учета в модели влияния качественного фактора.
- Функциональная связь** – связь, при которой определенному значению факторного признака соответствует одно и только одно значение результативного признака.
- Экзогенные (внешние, независимые) переменные** – переменные, значения которых определяются вне рассматриваемой модели, для которой они являются заданными.

Эконометрика – раздел науки, изучающий конкретные количественные и качественные взаимосвязи экономических объектов и процессов с помощью математических и статистических методов и моделей.

Экспоненциальная аналитическая зависимость – зависимость, определяемая соотношением $y = e^{a \cdot x_1^{b_1} \cdot x_2^{b_2} \cdot \dots \cdot x_n^{b_n} \cdot \varepsilon}$.

Экстраполяция – прогноз, получение расчетных значений при условии, что значения аргумента выходят за пределы области определения функции.

Эндогенные (внутренние, зависимые) переменные – переменные, значения которых определяются внутри рассматриваемой модели.

Приложение 8

КРАТКИЙ АНГЛО-РУССКИЙ СЛОВАРЬ ЭКОНОМЕТРИЧЕСКИХ ТЕРМИНОВ

Adjusted R^2 – скорректированный R^2

Autocorrelation function (ACF) – автокорреляционная функция

Autoregressivemodel order p , AR(p) – авторегрессионная модель порядка p

Autoregressiveintegratedmovingaveragemodel (ARIMA) – интегрированная модель авторегрессии и скользящего среднего

Autoregressivemovingaveragemodel (ARMA) – модель авторегрессии и скользящего среднего

Bestlinearunbiasedestimator (BLUE) – наилучшая (с минимальной дисперсией) оценка в классе несмещенных линейных оценок

Box-Jenkinsmodel = ARIMA – модель Бокса-Дженкинса = интегрированная модель авторегрессии и скользящего среднего

Central limittheorem (CLT) – центральная предельная теорема

Classicalnormalregression (CNR) – модель классической регрессии, в которой ошибки имеют совместное нормальное распределение

Classicalregression (CR) – модель регрессии, в которой ошибки независимы, одинаково распределены, имеют нулевое среднее значение и постоянную дисперсию

Conditionaldistribution – условное распределение

Conditionalexpectation – условное среднее, условное математическое ожидание

Confidenceinterval – доверительный интервал

Consistentestimator – состоятельная оценка

Convergence in distribution (law) – сходимость по распределению

Convergence in probability – сходимость по вероятности

Correlation – корреляция

- Correlogramm**– график (выборочной) автокорреляционной функции
- Correlationcoefficient**– коэффициент корреляции
- Covariance**– ковариация
- Cross-sectiondata**– данные, не имеющие временной природы, порядок их расположения не существен
- Curvefitting**– подгонка кривой
- Densityfunction** – плотность распределения
- Dependent (endogenous) variable** – зависимая (эндогенная) переменная
- Determination coefficient (R-squared)** – коэффициент детерминации R^2
- Distribution** – распределение
- Distributionfunction** – функция распределения
- Dummyvariable** – фиктивная независимая переменная, принимающая, как правило, два значения – 0 или 1
- Dummytrap** – ситуация, когда сумма нескольких фиктивных переменных, включенных в регрессию, равна константе, также включенной в регрессию
- Endogenous (dependent) variable** – эндогенная (зависимая) переменная
- Estimation** – величина оценки при заданных выборочных значениях
- Estimator** – метод оценивания, функция выборочных значений
- Exogenous (independent) variable** – экзогенная (независимая) переменная, регрессор
- Expectation (mean)** – среднее значение, математическое ожидание
- Explanatoryvariables**– объясняющие переменные, регрессоры, независимые переменные
- Explained (unexplained) variance** – объясняемая (необъясняемая) дисперсия
- Exponential smoothing** – экспоненциальное сглаживание
- Fitted value** – прогнозное значение
- Firstordercondition (FOC)** – необходимые условия экстремума
- Generalizedleastsquareestimations (GLS)** – обобщенный метод наименьших квадратов
- Goodnessoffit**– качество приближения данных моделью
- Heteroscedasticity**– гетероскедастичность
- Homoscedasticity**– гомоскедастичность
- Idempotentmatrix**– идемпотентная матрица
- Independent (exogenous) variable** – независимая (экзогенная) переменная
- Indirectleastquares**– косвенный метод наименьших квадратов

Instrumentalvariable (IV) – инструментальная переменная

Instrumentalvariableestimator (IV-estimator) – метод оценивания с помощью инструментальных переменных

Intercept– свободный член в уравнении регрессии

Jointdistribution– совместное распределение

Laggedvariable– лагированная переменная (для временных рядов), переменная с запаздыванием, значения которой в предыдущие моменты времени входят в уравнение для текущего момента времени

Law of large numbers (LLN) – законбольшихчисел

Level of statistical significance (significance level) – уровеньзначимостистатистическогокритерия

Likelihoodfunction– функция правдоподобия

Lineofthebestfit– линия, которая в определенном смысле наилучшим образом приближает наблюдения

Linearregressionmodel– линейная регрессионная модель

Lossfunction– функция потерь, измеряющая степень отклонения данных от «желаемых» значений

Marginaldistribution– маргинальное распределение, т.е. распределение одной или нескольких компонент случайного вектора

Maximumlikelihoodmethod (ML) – метод максимального правдоподобия

Maximumlikelihoodestimator– оценивание с помощью метода максимального правдоподобия

Maximumlikelihoodestimation– оценка максимального правдоподобия

Meanabsolutedeviation– среднее абсолютное отклонение

Meanabsolutepersentageerror– среднее относительное отклонение

Mean-squareerror– среднеквадратичная ошибка

Modelspecification– спецификация модели

Movingaverage– скользящее среднее

Movingaveragemodel, MA(q)– модель скользящего среднего (порядка q)

Multicollinearity – мультиколлинеарность

Multipleregressionmodel– многомерная регрессионная модель

Normal (Gaussian) distribution– нормальное (гауссовское) распределение

OLS-estimator, OLS-estimation– оценивание с помощью метода наименьших квадратов и значения этих оценок

Omittedvariables– пропущенные переменные (независимые переменные, не включенные в модель)

Ordinary least squares method (OLS) – метод наименьших квадратов, МНК

Outliers– выбросы (данные, имеющие большие отклонения от большинства других)

Partial autocorrelation function (PACF) – частная автокорреляционная функция

Partial correlation coefficient– коэффициент частной корреляции

Reduced form of the model – приведенная форма модели

Residuals – остатки

Sample – выборка

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

Seemingly unrelated regression (SUR) – система внешне не связанных между собой уравнений

Serial correlation– (для временных рядов) наличие корреляции между ошибками, относящимися к разным моментам времени

Significance level– уровень значимости

Simultaneous equations– одновременные уравнения

Slopes– коэффициенты при независимых переменных в уравнении регрессии

Standard deviation– стандартное отклонение (корень из дисперсии)

Stationary time series– стационарный временной ряд (статистические свойства которого не зависят от времени)

Testing hypothesis– проверка гипотез

Time series data– временные ряды, параметры системы, наблюдаемые в последовательные моменты времени

Two stage least squares (2SLS) – двухшаговый метод наименьших квадратов

Unbiased estimator– несмещенное оценивание (оценка)

Under-, overestimation– оценка параметра, имеющая отрицательное (соответственно положительное) смещение

Variance– дисперсия

Variance (covariance) matrix– матрица ковариаций случайного вектора

Weighted least squares– взвешенный метод наименьших квадратов

*Приложение 9***КОНТРОЛЬНЫЕ ВОПРОСЫ**

1. Что является предметом изучения эконометрии?
2. Связь эконометрии с другими науками.
3. Статистическая база эконометрических моделей.
4. В чем различие между функциональной и стохастической зависимостью?
5. Что собой представляет корреляционная связь?
6. Какими статистическими методами исследуются функциональные и корреляционные связи?
7. Какие основные задачи решают с помощью корреляционного и регрессионного анализа?
8. Дайте определение модели.
9. Этапы эконометрического анализа.
10. Охарактеризуйте основные проблемы и правила построения однофакторной линейной регрессионной модели.
11. В чем состоит значение уравнения регрессии?
12. Что характеризуют коэффициенты регрессии?
13. Метод определения параметров уравнения регрессии.
14. Предпосылки использования метода наименьших квадратов.
15. Зачем необходима проверка адекватности регрессионной модели?
16. Как осуществляется проверка значимости коэффициентов?
17. Какими показателями измеряется теснота корреляционной связи?
18. Какое значение имеет расчет коэффициента детерминации?
19. Линейные коэффициенты корреляции и детерминации, их смысл и значение.
20. Проверка существенности показателей тесноты связи как необходимое условие распространения выводов по результатам выборки на всю генеральную совокупность. Как она осуществляется?
21. Как экономически охарактеризовать однофакторную регрессионную модель?
22. Какой экономический смысл имеют коэффициенты эластичности?
23. Когда регрессионная модель считается линейной? Нелинейной?
24. Основные проблемы и правила построения многофакторной корреляционной модели.
25. Сущность и назначение парных и частных коэффициентов корреляции.
26. Сущность и значение совокупного коэффициента множественной корреляции и совокупного коэффициента детерминации.
27. Как проверить адекватность в целом? Значимость коэффициента регрессии? Какие критерии для этого можно использовать?
28. Как экономически интерпретировать многофакторную регрессионную модель?
29. Построение доверительных интервалов.
30. Производственная функция Кобба-Дугласа.

31. Какие выводы можно сделать на основе коэффициентов эластичности производственной функции?
32. В чем причины мультиколлинеарности?
33. Тестирование на наличие мультиколлинеарности и методы ее исключения.
34. Когда возникает гетероскедастичность? Тесты на гетероскедастичность и методы ее устранения.
35. Метод Спирмена тестирования гетероскедастичности.
36. Метод Гольдфельда-Квандта тестирования гетероскедастичности.
37. Модели с фиктивными переменными. Трактовка коэффициентов при фиктивных переменных.
38. Проверка значимости параметров в модели с фиктивными переменными.
39. Авторегрессионные модели.
40. В чем состоит автокорреляция? Тест на автокорреляцию.
41. Компоненты временных рядов.
42. Автокорреляция уровней ряда. Коэффициент автокорреляции. Лаг.
43. Автокорреляционная функция временного ряда. Коррелограмма.
44. Аналитическое выравнивание временного ряда (тренд).
45. Аддитивная модель временного ряда (скользящие средние).
46. Аддитивная модель временного ряда (фиктивные переменные)
47. Мультипликативная модель временного ряда (скользящие средние).
48. Мультипликативная модель временного ряда (фиктивные переменные).
49. Дистрибутивно-лаговые модели.
50. Система одновременных уравнений.
51. Методы оценивания неизвестных параметров в системах одновременных уравнений.
52. Область применения косвенного метода наименьших квадратов и двухшагового метода наименьших квадратов.

СОДЕРЖАНИЕ

ОБЩИЕ ПОЛОЖЕНИЯ	3
ОРГАНИЗАЦИОННО-МЕТОДИЧЕСКИЕ УКАЗАНИЯ.....	3
ТЕМАТИЧЕСКИЙ ПЛАН	6
СОДЕРЖАНИЕ РАЗДЕЛА 1 «ПАРНАЯ РЕГРЕССИЯ».....	6
СОДЕРЖАНИЕ РАЗДЕЛА 2 «МНОЖЕСТВЕННЫЙ РЕГРЕССИОННЫЙ АНАЛИЗ».....	7
СОДЕРЖАНИЕ РАЗДЕЛА 3 «ДИНАМИЧЕСКИЕ ПРОЦЕССЫ»	7
РАЗДЕЛ 1. Парная регрессия	8
ТЕМА 1.1. ЛИНЕЙНАЯ РЕГРЕССИЯ.....	8
ТЕМА 1.2. НЕЛИНЕЙНАЯ РЕГРЕССИЯ	13
РАЗДЕЛ 2. Множественный регрессионный анализ	18
ТЕМА 2.1. МНОГОФАКТОРНЫЕ МОДЕЛИ.....	18
ТЕМА 2.2. ОСОБЫЕ СЛУЧАИ	26
РАЗДЕЛ 3. Моделирование динамических процессов	36
ТЕМА 3.1. АВТОКОРРЕЛЯЦИЯ.....	36
ТЕМА 3.2. МОДЕЛИ ДИНАМИЧЕСКИХ ПРОЦЕССОВ.....	38
СПИСОК УЧЕБНО-МЕТОДИЧЕСКОЙ ЛИТЕРАТУРЫ	44
Приложение 1.....	45
ОСНОВНЫЕ ФОРМУЛЫ.....	45
Приложение 2.....	53
ПРИМЕНЕНИЕ ПАКЕТА АНАЛИЗА MS EXCEL.....	53
Приложение 3.....	58
ВАЛОВЫЙ ВНУТРЕННИЙ ПРОДУКТ (ВВП) И ПОТРЕБИТЕЛЬСКИЕ РАСХОДЫ (ПР), 1970–2013	58
Приложение 4.....	62
F-РАСПРЕДЕЛЕНИЕ ФИШЕРА	62
Приложение 5.....	65
РАСПРЕДЕЛЕНИЕ СТЬЮДЕНТА	65
Приложение 6.....	67
ГРАНИЧНЫЕ ЗНАЧЕНИЯ ДЛЯ СТАТИСТИК ДАРБИНА – УОТСОНА.....	67
Приложение 7.....	68
ГЛОССАРИЙ	68
Приложение 8.....	70
КРАТКИЙ АНГЛО-РУССКИЙ СЛОВАРЬ	70
Приложение 9.....	74
КОНТРОЛЬНЫЕ ВОПРОСЫ.....	74

Навчальне видання

ЕКОНОМЕТРИКА

Методичні рекомендації та матеріали для самостійної роботи
для студентів факультету «Бізнес-управління»,
які навчаються за напрямом підготовки
6.030504 Економіка та підприємництво

(російською мовою)

Автори-упорядники: *С. Б. Данилевич, О. В. Дьячкова*

Комп'ютерна верстка *О. В. Дьячкова*

Підписано до друку 5.05.2017. Формат 60×84/16.
Папір офсетний. Гарнітура «Таймс».
Ум. друк. арк. 4,43. Обл.-вид. арк. 3,89.
Тираж 100 пр. Зам. №

План 2015/16 навч. р., поз. № 7 в переліку робіт кафедри

Видавництво
Народної української академії
Свідоцтво № 1153 від 16.12.2002.

Надруковано у видавництві
Народної української академії

Україна, 61000, Харків, МСП, вул. Лермонтовська, 27.