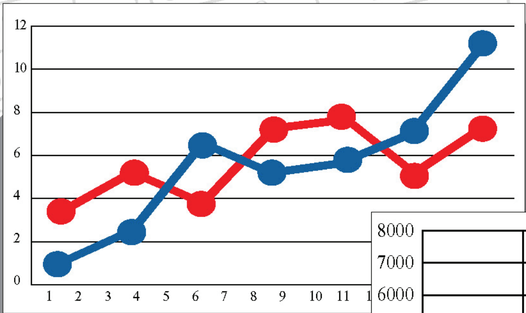
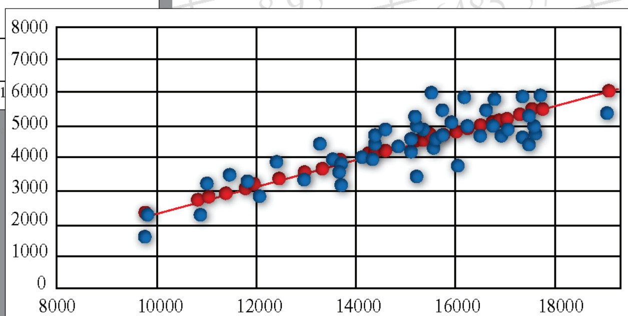
Министерство науки и высшего образования Российской Федерации
Тольяттинский государственный университет
Институт математики, физики и информационных технологий

О.А. Кузнецова, С.Ш. Палферова

ЭКОНОМЕТРИКА
(ПРОДВИНУТЫЙ УРОВЕНЬ)

Электронное учебно-методическое пособие





ISBN 978-5-8259-1525-8

УДК 330.43
ББК 65в631.8

Рецензенты:
д-р экон. наук, канд. пед. наук, профессор
Поволжского государственного университета сервиса *Л.В. Глухова*;
д-р техн. наук, профессор, профессор кафедры
«Высшая математика и математическое образование»
Тольяттинского государственного университета *П.Ф. Зибров*.

Кузнецова, О.А. Эконометрика (продвинутый уровень) : электронное учебно-методическое пособие / О.А. Кузнецова, С.Ш. Палферова. – Тольятти : Изд-во ТГУ, 2020. – 1 оптический диск. – ISBN 978-5-8259-1525-8.

Учебно-методическое пособие по изучению дисциплины «Эконометрика (продвинутый уровень)» включает учебно-методические материалы по дисциплине, материалы для контроля знаний, задания к контрольной работе с методическими рекомендациями для их выполнения.

Предназначено для студентов, обучающихся по направлению подготовки 38.04.01 «Экономика», очной и заочной форм обучения, а также может быть полезно для студентов дистанционной формы обучения.

Текстовое электронное издание.

Рекомендовано к изданию научно-методическим советом Тольяттинского государственного университета.

Минимальные системные требования: IBM PC-совместимый компьютер: Windows XP/Vista/7/8; PIII 500 МГц или эквивалент; 128 Мб ОЗУ; SVGA; CD-ROM; Adobe Acrobat Reader.

Редактор *Е.В. Пилясова*
Корректор *О.В. Горбань*
Технический редактор *Н.П. Крюкова*
Компьютерная верстка: *Л.В. Сызганцева*
Художественное оформление,
компьютерное проектирование: *И.И. Шишкина*

Дата подписания к использованию 10.07.2020.
Объем издания 5 Мб.
Комплектация издания: компакт-диск,
первичная упаковка.
Заказ № 1-78-19.

Издательство Тольяттинского государственного университета
445020, г. Тольятти, ул. Белорусская, 14,
тел. 8 (8482) 53-91-47, www.tltsu.ru

Содержание

ВВЕДЕНИЕ

В соответствии с федеральным государственным образовательным стандартом высшего образования по направлению подготовки 38.04.01 «Экономика» учебный курс «Эконометрика (продвинутый уровень)» включен в учебные планы подготовки магистров экономики как обязательная дисциплина, является базовым и преподается во всех ведущих университетах мира.

Учебно-методическое пособие предназначено для изучения дисциплины «Эконометрика (продвинутый уровень)», целью которой является «обучение студентов методологии и методике построения и применения эконометрических моделей для анализа состояния и оценки перспектив развития экономических и социальных систем в условиях взаимосвязей между их внутренними и внешними факторами.

Задачи

1. Формирование у студентов навыков анализа связей между экономическими факторами и показателями на основе статистических данных с использованием аппарата теории вероятностей и математической статистики.
2. Обучение студентов практическому применению методов экономической теории, экономической статистики, экономических измерений и математико-статистического инструментария.
3. Развитие навыков прогнозирования социально-экономических показателей, характеризующих состояние и развитие анализируемой системы» [21].

Данный учебный курс базируется на освоении таких дисциплин, как «Линейная алгебра», «Высшая математика», «Теория вероятностей и математическая статистика», «Экономическая теория», «Статистика» и другие, а также на владении основами современных компьютерных технологий.

В результате изучения дисциплины «Эконометрика (продвинутый уровень)» у студентов должны быть сформированы следующие компетенции:

– способность к абстрактному мышлению, анализу, синтезу;
– готовность действовать в нестандартных ситуациях, нести социальную и этическую ответственность за принятые решения;

– готовность к саморазвитию, самореализации, использованию творческого потенциала;
– готовность к коммуникации в устной и письменной формах на русском и иностранном языках для решения задач профессиональной деятельности;
– способность составлять прогноз основных социально-экономических показателей деятельности предприятия, отрасли, региона и экономики в целом [22].

В процессе изучения дисциплины «Эконометрика (продвинутый уровень)» обучающиеся должны:

• *знать* основные понятия линейной алгебры, теории вероятностей и математической статистики, микро- и макроэкономики, общей теории статистики, теоретические основы эконометрики, эконометрические методы исследования и прогнозирования основных социально-экономических и финансовых показателей деятельности предприятия, отрасли, региона и экономики в целом;

• *уметь* применять эконометрические методы исследования и прогнозирования основных социально-экономических и финансовых показателей деятельности предприятия, отрасли, региона и экономики в целом;

• *владеть* «навыками математической обработки экономических данных с широким использованием современных компьютерных вычислительных технологий, а также визуализацией результатов на всех этапах эконометрического моделирования» [22].

Учебный курс «Эконометрика (продвинутый уровень)» включает семинарские занятия и выполнение четырех практических работ, а также сдачу промежуточных тестов по модулям дисциплины. По лабораторным и практическим работам выставляется оценка «зачтено», если студент выполнил не менее 70 % от общего объема заданий, в противном случае выставляется оценка «не зачтено». По каждому модулю дисциплины предусмотрен промежуточный тест, который включает 24 задания по теоретическому и практическому материалу. Каждый тест оценивается по пятибалльной системе: оценка «отлично» выставляется студенту, если правильно выполнены 80–100 % заданий, «хорошо» – 60–79 %, «удовлетворительно» – 40–59 %, «неудовлетворительно» – 0–39 %.

В качестве итоговой аттестации по учебному курсу предусмотрен экзамен, который проводится в устной форме по билетам, содержащим два теоретических вопроса и задачу. К экзамену допускаются студенты, получившие зачет по лабораторным и практическим работам и положительные оценки за промежуточные тесты. Оценка «отлично» выставляется студенту, если он ответил на теоретические вопросы билета и правильно решил задачу; оценка «хорошо» – если студент ответил на теоретические вопросы билета, но решил задачу с ошибками или недочетами; оценка «удовлетворительно» выставляется студенту, если он ответил только на один вопрос билета и правильно решил задачу; оценка «неудовлетворительно» – если студент не ответил на вопросы билета и не решил задачу.

Учебно-методическое пособие содержит три раздела курса. По каждой теме разделов приводятся краткие теоретические сведения, примеры решения типовых задач, задания для самостоятельной работы, рекомендации по выполнению заданий, список рекомендуемой литературы. В конце пособия приведен глоссарий основных терминов и определений, а также добавлены приложения статистических таблиц и распределений, необходимых при выполнении заданий.

Каждый студент должен последовательно изучить учебный материал разделов, акцентировать внимание на основных понятиях и терминах, ответить на экзаменационные вопросы, приведенные в конце пособия. Необходимо выполнить контрольные задания по каждой теме в соответствии с вариантом, согласованным с преподавателем, после чего ответить на вопросы теста. В результате студент должен научиться решать задачи, аналогичные разобранным в учебно-методическом пособии, выполнить контрольную работу и тест итогового контроля.

Модуль 1. МОДЕЛЬ ПАРНОЙ РЕГРЕССИИ

Тема 1.1. Виды эконометрических моделей. Введение в регрессионный анализ

Форма проведения занятия: семинар.

Вопросы для обсуждения

1. Предмет и задачи эконометрики.
2. Характеристики случайных величин.
3. Типы эконометрических моделей.
4. Типы данных при эконометрическом моделировании.
5. Основные положения регрессионного анализа.

Методические указания по проведению занятия

Семинар начинается с уточнения базовых понятий. Все студенты готовят сообщение по всем вопросам семинара. Каждый студент готовит реферат по отдельным вопросам семинара. Один студент из группы готовит доклад творческого характера, содержащий элементы исследовательского характера.

1. Предмет и задачи дисциплины

«*Эконометрика* – это совокупность методов анализа количественных связей между экономическими факторами и показателями на основании реальных статистических данных с использованием аппарата теории вероятностей и математической статистики.

Анализируя характер имеющихся статистических данных, используя методы эконометрики, исследователь может сделать определенные заключения о возможной форме подходящей теоретической экономической модели. Статистические данные указывают на то, в каком направлении нужно искать теоретические модели. Построение окончательной модели производится с учетом представлений экономической теории и с учетом информации, содержащейся в эмпирических данных» [16].

Обобщенный вид эконометрической модели:

$$y = f(x_1, x_2, ..., x_n) + \varepsilon, \tag{1.1}$$

где y – объясняемая величина (искомое значение зависимой переменной); $x_1, x_2, ..., x_n$ – независимые (объясняющие) переменные; $f(x_1, x_2, ..., x_n)$ – зависящая от значений $x_1, x_2, ..., x_n$ объясненная часть; ε – случайная компонента.

В качестве примера исследуем связь между годовым располагаемым доходом (обозначим как переменную x) и расходами на личное потребление за год (y) домашних хозяйств (табл. 1.1). Количество исходных данных $n = 20$.

Известный ученый Кейнс, исследовавший психологические мотивы экономического поведения, в своих трудах вывел в качестве основополагающего закона «склонность людей (как правило и в среднем) увеличивать расходы на личное потребление по мере возрастания их доходов, но не в той степени, в какой возрастает доход» [1].

Опираясь на этот закон, можно ввести функциональную зависимость

$$y = f(x),$$

где переменные x и y должны быть представлены в одних единицах измерения. Сама же функция $f(x)$ согласно закону Кейнса – возрастающая, скорость ее изменения меньше единицы.

Таблица 1.1

i	x	y	i	x	y
1	2508	2406	11	2435	2311
2	2572	2464	12	2354	2278
3	2408	2336	13	2404	2240
4	2522	2281	14	2381	2183
5	2700	2641	15	2581	2408
6	2531	2385	16	2529	2379
7	2390	2297	17	2562	2378
8	2595	2416	18	2624	2554
9	2524	2460	19	2407	2232
10	2685	2549	20	2448	2356

Для выяснения формы функциональной связи в декартовой системе координат наносят точки (x, y), получая таким образом диаграмму рассеяния. Другое название полученного графика – поле корреляции. Для рассматриваемого случая график представлен на рис. 1.1.

Для простоты обычно предполагают наиболее часто встречающийся и доступный для понимания линейный тренд. Это простейшая модель связи двух переменных, вводимая для наблюдений различных зависимостей, которая имеет вид

$$y = \alpha + \beta x + \varepsilon, \tag{1.2}$$

где β – некоторая постоянная величина, изменяющаяся в пределах от нуля до единицы и характеризующая в исследуемом множестве представленных хозяйств их предрасположенность к тратам согласно выработанным привычкам; α – постоянное потребление; $\varepsilon = y - (\alpha + \beta x)$ – это отклонение реальных трат y_i от значения $y = a + \beta x$, которое было найдено по составленной линейной модели связи для i-го исследуемого объекта (хозяйства).

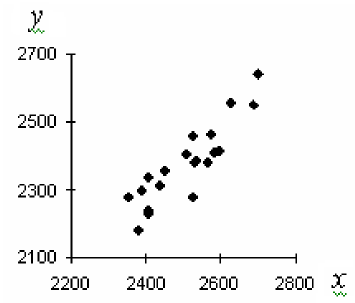


Рис. 1.1

Как видно из рис. 1.1, построенные по исходным данным точки не лежат на одной прямой. Они образуют так называемое «облако рассеяния». Это может быть объяснено наличием случайной составляющей ε.

Величины параметров α и β следует оценить при помощи составленной теоретической модели, созданной на основе исходных с учетом указанного отклонения. По полученным результатам согласно известным критериям выносится вердикт о возможности применения выбранной модели для прогноза.

2. Характеристики случайных величин

Обозначив через x_1, x_2, ..., x_n располагаемые доходы домашних хозяйств, рассмотренные в предыдущем примере, а через y_1, y_2, ..., y_n – расходы, мы таким образом введем понятие наблюдаемых значений двух переменных. В данном примере имеются $n = 20$ пар наблюдаемых значений переменных x и y: $(x_1; y_1)$, $(x_2; y_2)$, ..., $(x_{20}; y_{20})$.

Среди множества показателей, характеризующих последовательности дискретных величин x_1, x_2, ..., x_n и y_1, y_2, ..., y_n, наиболее простыми являются *средние значения*.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i. \qquad (1.3)$$

Средние значения для данных в примере показателей

$$\bar{x} = \frac{2508 + 2572 + \ldots + 2448}{20} = 2508,$$

$$\bar{y} = \frac{2406 + 2572 + \ldots + 2356}{20} = 2377{,}7.$$

«Математическое ожидание дискретных случайных величин – это сумма произведений всех значений дискретной величины на их вероятности, оно приближенно равно их средним значениям» [16].

$$M(X) = x_1p_1 + x_2p_2 + \cdots + x_np_n = \sum_{i=1}^{n} x_i\, p_i,$$
$$M(X) = y_1p_1 + y_2p_2 + \cdots + y_ip_n = \sum_{i=1}^{n} y_i\, p_i. \qquad (1.4)$$

Выборочные дисперсии рассчитывают согласно формулам (1.5). Они характеризуют отклонение значений дискретной величины от среднего значения.

$$\operatorname{var}(x) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2, \ \operatorname{var}(y) = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2. \qquad (1.5)$$

Другая характеристика – стандартное отклонение – измеряется в тех же единицах, что и сама исследуемая величина. Поэтому при оценке рассеяния дискретной случайной величины предпочтение отдают именно стандартному отклонению, а не вариации.

$$S(x) = \sqrt{\mathrm{var}(x)}, \quad S(y) = \sqrt{\mathrm{var}(y)}. \tag{1.6}$$

Кроме аналитических характеристик, имеется графическое средство анализа исходных данных. Им является диаграмма рассеяния (поле корреляции). Для ее построения в декартовой системе координат располагают точки с координатами (x_i, y_i), $i = 1, 2, ..., n$. Диаграмма должна иметь форму квадрата, что в обязательном порядке учитывают при выборе масштаба построения. Некоторые точки на «поле» обязательно должны быть расположены вдоль сторон квадрата. При соблюдении этих нехитрых правил полученный график позволяет судить о виде зависимости между величинами.

По формуле (1.7) рассчитывают *выборочный коэффициент корреляции*, который характеризует степень выраженности линейной связи между переменными x и y. Также для расчета r_{xy} можно применять формулу (1.9), в которой не используют округленные значения, получаемые при расчете средних значений, как в формуле (1.7), а подставляют непосредственно исходные данные.

Для характеристики степени зависимости двух случайных величин и степени их рассеяния используют *выборочную ковариацию* cov(x, y), *расчет которой производят по формуле* (1.8).

$$r_{xy} = \frac{\mathrm{cov}(x,y)}{S(x)S(y)} = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{S_x S_y}, \tag{1.7}$$

$$\mathrm{cov}(x,y) = \frac{1}{n}\sum\nolimits_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \overline{xy} - \bar{x} \cdot \bar{y}, \tag{1.8}$$

$$r_{xy} = \frac{n\sum_{i=1}^{n} x_i y_i - (\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} y_i)}{\sqrt{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}\sqrt{n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2}}. \tag{1.9}$$

Рассмотрим основные свойства выборочного коэффициента корреляции в случае наличия достаточно большого количества исходных статистических данных.

1-е свойство. Как было уже сказано, выборочный коэффициент корреляции показывает степень выраженности линейной связи между произвольными переменными x и y. Поэтому его называют коэффициентом линейной корреляции. По модулю его значение не больше единицы. Как показано на рис. 1.2, чем ближе абсолютная величина коэффициента к единице, тем теснее линейная связь.

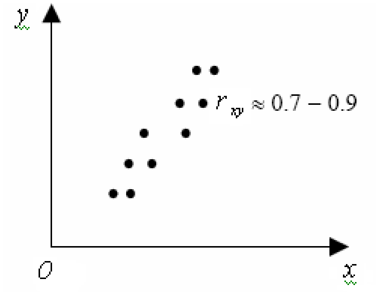


Рис. 1.2

2-е свойство. При коэффициенте корреляции, равном единице или минус единице, графиком будет являться прямая линия, на которой располагаются наблюдаемые значения, то есть корреляционная связь является линейной зависимостью. Данное свойство иллюстрируется рис. 1.3 и 1.4.

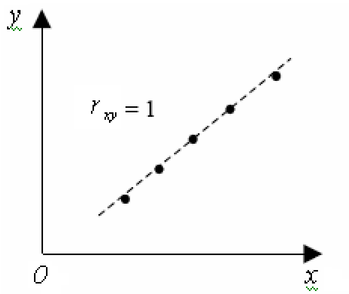


Рис. 1.3

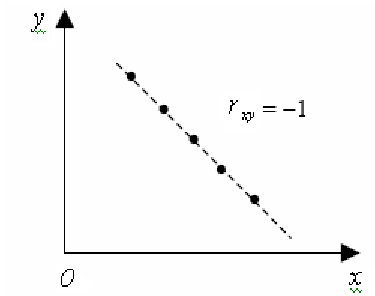


Рис. 1.4

3-е свойство. Как показано на рис. 1.5, 1.6, 1.7, при нулевом значении коэффициента корреляции линейная связь отсутствует. Возможно, в этом случае облако рассеивания представляет собой хаотично расположенные точки. Тогда между переменными нет никакой зависимости. Или зависимость является нелинейной, как на рис. 1.7.

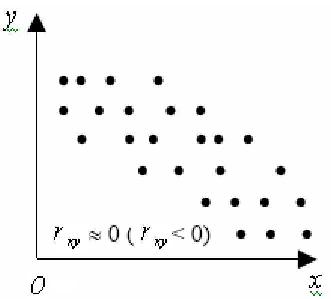


Рис. 1.5

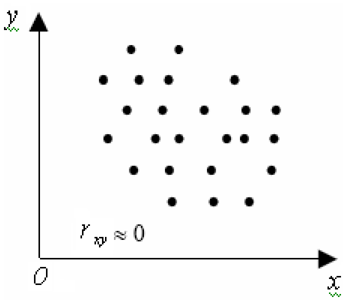


Рис. 1.6

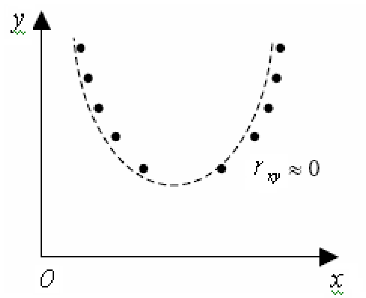


Рис. 1.7

3. Типы эконометрических моделей

Один из типов эконометрических моделей – это модели временны́х рядов.

«Модели, построенные по данным, характеризующим один объект за ряд определенных последовательных периодов, называются моделями временны́х рядов» [14].

К ним относятся модели тренда, описываемые формулой (1.10). В них входят временно́й тренд $T(t)$ и случайная, или стохастическая, компонента ε_t. Временно́й тренд, или тенденция, имеет заданный параметрический вид, например линейный. Эта компонента отражает влияние долговременных факторов, например рост населения, экономическое развитие, изменение структуры потребления.

$$y(t) = T(t) + \varepsilon_t. \qquad (1.10)$$

Формула (1.11) показывает модели сезонности. Они включают периодическую или сезонную компоненту $S(t)$ и случайную составляющую ε_t. Периодичность показывает повторяемость процессов за непродолжительный период времени. Примером может служить объем продаж билетов на поезда в южном направлении в разное время года.

$$y(t) = S(t) + \varepsilon_t. \qquad (1.11)$$

Модели тренда и сезонности бывают аддитивными, как показывает формула (1.12), или мультипликативными, соответствующими формуле (1.13). Такое деление осуществляется в зависимости от способа вхождения в модель трех составляющих: тренда $T(t)$, периодической $S(t)$ и случайной ε_t компонент.

$$y(t) = T(t) + S(t) + \varepsilon_t \text{ (аддитивная)}, \qquad (1.12)$$

$$y(t) = T(t) \cdot S(t) \cdot \varepsilon_t \text{ (мультипликативная)}. \qquad (1.13)$$

Модели адаптивного прогноза, модели авторегрессии и скользящего среднего являются более сложными моделями временных рядов. Последующее поведение временно́го ряда в них зависит только от предыдущих значений ряда.

Регрессионные модели с одним уравнением

Зависимость называется статистической, если из изменения одной из величин следует изменение распределения другой. Если изменение одной из величин влечет за собой изменение среднего значения другой случайной величины, то подобного рода статистическая зависимость называется корреляционной. Регрессионная зависимость – это зависимость математического ожидания случайной величины или ее среднего значения от одной или нескольких других независимых переменных, также являющихся случайными величинами.

Термин «регрессия к посредственности» был использован Фрэнсисом Гальтоном при описании наследственных признаков. Им было обнаружено, что дети обычно не наследуют выдающийся рост (низкий или высокий) родителей. Сначала термин «регрессия» использовался только в биологическом смысле, но работы Карла Пирсона позволили использовать этот термин также в статистике.

В регрессионных моделях зависимая или объясняемая переменная y представляется в виде функции, соответствующей формуле (1.14). В функцию входит условное математическое ожидание зависимой величины M_y, полученное при заданном наборе $(x_1, ..., x_p)$ значений независимых переменных, или так называемая функция регрессии, а также случайная составляющая ε.

$$y = M_y(x_1, ..., x_p) + \varepsilon. \quad (1.14)$$

Регрессионные модели классифицируют на линейные и нелинейные исходя из вида функций, их представляющих. Например, можно исследовать величину зарплаты в зависимости от наличия у сотрудника опыта работы (стажа) в данной области, от наличия образования по специальности и его уровня, от пола сотрудника, его возраста, наличия и количества детей и т. п.

По области применения такие модели более распространены, чем модели временных рядов.

Системы одновременных уравнений

Самые сложные из эконометрических моделей – это системы одновременных уравнений.

«Система одновременных уравнений представляет собой совокупность эконометрических уравнений, определяющих взаимозависимость экономических переменных. Важным отличительным признаком системы одновременных уравнений от прочих систем уравнений являетя наличие одних и тех же переменных в правых и левых частях разных уравнений системы.

Системы одновременных уравнений используются в макроэкономике. Они требуют использования более сложного математического аппарата. Это обусловлено наличием в них и тождеств, и регрессионных уравнений, которые в свою очередь включают в себя кроме объясняющих переменных объясняемые переменные из других уравнений системы.

Рассмотрим, например, модель спроса и предложения.

$$Q_t^S = \alpha_1 + \alpha_2 + \alpha_3 P_t + \alpha_3 P_{t-1} + \varepsilon_t \text{ (предложение)},$$
$$Q_t^D = \beta_1 + \beta_2 + \beta_3 P_t + \beta_3 Y_t + u_t \text{ (спрос)},$$
$$Q_t^S = Q_t^D \text{ (равновесие)},$$

где Q_t^S – предложение товара в момент времени t; Q_t^D – спрос на товар в момент времени t; P_t – цена товара в момент времени t; Y_t – доход в момент времени t.

Цена на товар P_t и спрос на товар Q_t^S (равный предложению Q_t^D) – эндогенные переменные. Они определяются из уравнений модели. Значение P_{t-1} и Y_t – экзогенные переменные» [20].

В системе присутствуют различные виды переменных. Взаимосвязанные переменные, описывающие экономический объект и формирующиеся внутри функционирования объекта, называются эндогенными. Переменные, задаваемые извне, – экзогенными.

Взятые в предыдущий момент времени переменные называют лаговыми.

4. Типы данных при эконометрическом моделировании

В эконометрике выделяют три типа данных. Первый тип – пространственный срез (данные в определенный момент времени), например набор сведений (объем производства или количество работников по разным предприятиям) в один и тот же момент времени.

Второй тип данных – временны́е ряды (данные, упорядоченные по времени). Это данные через определенные отрезки времени, например котировки акций за ряд лет.

Кроме наличия порядка по времени, наблюдения (данные в эконометрическом моделировании, отнесенные ко второму типу) часто бывают зависимыми. Такую зависимость в эконометрике называют автокорреляционной.

Третий тип данных – панельные данные – занимают промежуточное положение. Они отражают наблюдения по большому количеству объектов и показателей за несколько моментов времени, например: финансовые показатели работы нескольких крупных паевых инвестиционных фондов за несколько месяцев; суммы налогов, уплаченных нефтяными компаниями за последние несколько лет.

Собранные данные могут быть представлены в виде таблицы, диаграммы, графика.

5. Основные положения регрессионного анализа

В регрессионном анализе объясняемая переменная y записывается в виде функции, представленной формулой (1.15). Она состоит из суммы двух составляющих.

$$y = M_y(x_1, x_2, \ldots, x_n) + \varepsilon. \tag{1.15}$$

Первое слагаемое $M_y(x_1, x_2, ..., x_n)$ – это функция регрессии, значение которой является условным математическим ожиданием величины y, получается при данном наборе значений объясняющих переменных.

Функция регрессии, значение которой зависит от переменных $x_1, x_2, ..., x_n$, называется объясненной частью, зависящей от значений объясняющих переменных.

Переменную y называют наблюдаемым значением зависимой или объясняемой переменной, то есть результативным значением наблюдаемого экономического фактора.

$x_1, x_2, ..., x_n$ называются независимыми или объясняющими переменными, то есть значениями тех экономических факторов, которые влияют на объясняемую переменную.

Второе слагаемое ε является случайной, или стохастической, составляющей, также называется ошибкой наблюдения.

В случае парной регрессии значение функции регрессии зависит от одной объясняющей переменной, как показывает формула (1.16).

$$y = M_y(x) + \varepsilon. \tag{1.16}$$

Для линейной парной регрессии эта функция имеет линейный вид, соответствующий формуле (1.17). Ее графиком будет являться прямая линия.

$$y_i = \alpha + \beta x_i + \varepsilon. \tag{1.17}$$

Приведем пример парной регрессионной зависимости.

Как уже отмечалось, известна «склонность людей (как правило и в среднем) увеличивать расходы на личное потребление по мере возрастания их доходов» [24]. Зависимость расходов на личное потребление семейных хозяйств от их среднемесячных доходов исследуется при помощи модели парной регрессии.

Стандартные предположения регрессионного анализа

Стандартные предположения регрессионного анализа – это предположения о свойствах статистических данных, имеющихся у исследователя.

Данные допущения называют также предпосылками метода наименьших квадратов. Их четыре.

Во-первых, в модели не все значения объясняющих, или независимых, переменных $x_1, x_2, ..., x_n$ совпадают между собой. Это предположение называют условием идентифицируемости.

Тогда можно вычислить сумму квадратов отклонений значений независимых переменных от среднего значения, входящую в формулы числовых характеристик величины x:

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 \neq 0.$$

Иначе говоря, если значения переменной x будут одинаковы, тогда среднее значение совпадет с этими значениями. Отклонение значений независимых переменных от среднего значения будет равно нулю, и невозможно будет вычислить числовые характеристики.

Во-вторых, $y_1, y_2, ..., y_n$ получаются наложением на значения $(\alpha + \beta x_i)$ случайных ошибок ε_i. Значения $y_1, y_2, ..., y_n$ носят случайный характер, определяемый случайным характером ε_i. Это условие предполагает отсутствие автокорреляционной зависимости остатков от номера наблюдения. Автокорреляция – статистическая взаимосвязь между случайными величинами из одного ряда, взятыми со сдвигом.

В-третьих, ошибки $\varepsilon_1, \varepsilon_2, ..., \varepsilon_n$ – независимые случайные величины, имеющие нормальное распределение.

Соответственно, и $y_1, y_2, ..., y_n$ – независимые случайные величины. Математическое ожидание, или среднее значение ошибок, равно нулю: $M\varepsilon_i = 0$.

Четвертое предположение о характере имеющихся статистических данных следующее. Дисперсия ошибок ε_i (соответственно, и величины y_i) постоянна для любого i.

$$\mathrm{var}(\varepsilon_i) = \sigma^2(\varepsilon_i) = \mathrm{const}, \quad \mathrm{var}(y_i) = \sigma^2(y_i) = \mathrm{const}. \qquad (1.18)$$

Это предположение называют условием равноизменчивости (гомоскедастичности) ошибок ε_i и соответственно объясняемой переменной y_i. Под гетероскедастичностью ошибок ε_i и объясняемой переменной y_i понимают условие

$$\mathrm{var}(\varepsilon_i) = \sigma^2(\varepsilon_i) \neq \mathrm{const}, \quad \mathrm{var}(y_i) = \sigma^2(y_i) \neq \mathrm{const}. \qquad (1.19)$$

Рекомендуемая литература

1. Домбровский В.В. Эконометрика : учебник. – М. : Новый учебник, 2004. – 342 с.
2. Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика : начальный курс : учебник для вузов. – М. : Дело, 2004. – 575 с.
3. Эконометрика : учебник для вузов / под ред. И.И. Елисеевой. – М. : Проспект, 2010. – 288 с.

Тема 1.2. Парная линейная регрессия. Метод наименьших квадратов

Форма проведения занятия: семинар.

Вопросы для обсуждения

1. Модель парной линейной регрессии. Оценка параметров по методу наименьших квадратов (МНК).
2. Использование оцененной модели для прогноза.
3. Интервальные оценки функции регрессии и ее параметров.
4. Оценка значимости уравнения линейной регрессии (адекватности имеющимся статистическим данным).
5. Проверка выполнения стандартных предположений об ошибках в линейной модели наблюдений графическим методом.

Методические указания по проведению занятия

Семинар начинается с уточнения базовых понятий. Все студенты готовят сообщение по всем вопросам семинара. Каждый студент готовит реферат по отдельным вопросам семинара. Один студент из группы готовит доклад творческого характера, содержащий элементы исследовательского характера.

1. Модель парной линейной регрессии. Оценка параметров по методу наименьших квадратов (МНК)

Мы уже отмечали ранее, что если между переменными x и y существует теоретическая линейная связь, то наблюдаемые значения этих переменных связаны линейной моделью наблюдений, как показано в формуле (1.20).

$$y = \alpha + \beta x. \tag{1.20}$$

Если α и β – истинные значения параметров линейной модели связи, то величина ε_i представляет собой ошибку в i-м наблюдении. Формула (1.21) показывает, что ошибка в i-м наблюдении равна разности между наблюдаемым значением объясняемой переменной y и ее значением, рассчитанным по теоретической линейной модели.

$$y_i = (\alpha + \beta x_i) + \varepsilon_i, \quad i = 1, 2, ..., n. \tag{1.21}$$

Поиск коэффициентов α и β осуществляется таким образом, чтобы величина ошибки стремилась к минимуму, а в идеале – к нулю. Если ошибка равна нулю, то все точки лежат на одной прямой. В результате получают подобранную модель линейной связи, соответствующую формуле (1.22):

$$\hat{y}_i = a + bx_i. \tag{1.22}$$

В этой модели наблюдаемое значение переменной x сопоставляется с подобранным значением переменной y. Значения подобранное и реальное наблюдаемое y обычно отличаются, как показано в формуле (1.23). Разность называется остатком в i-м наблюдении.

$$e_i = y_i - \hat{y}_i = y_i - (a + bx_i). \tag{1.23}$$

В реальной жизни при расчете исходных данных остатки либо положительны, либо отрицательны, редкий случай, когда остатки равны нулю. При этом необходимо соблюдение *принципа наименьших квадратов*:

$$\begin{gathered} \sum_{i=1}^{n\Sigma} e_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \to \min, \\ \sum_{i-1}^{n} (y_i - a - bx_i)^2 \to \min. \end{gathered} \tag{1.24}$$

Получаемые оценки параметров a и b носят название по применяемому при расчете принципу: оценки наименьших квадратов. Свойством оценок наименьших квадратов является то, что соответствующая им прямая проходит через точку $(\bar{x}, \bar{y})$. Поиск пары чисел a и b с помощью метода наименьших квадратов (МНК) дает формулы для расчета коэффициентов в подобранной модели:

$$b = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\overline{x^2} - \bar{x}^2} = \frac{\mathrm{cov}(x, y)}{\mathrm{var}(x)}, \tag{1.25}$$

$$a = \bar{y} - b\bar{x}. \tag{1.26}$$

где $\mathrm{cov}(x, y) = \dfrac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n} = \overline{xy} - \bar{x} \cdot \bar{y}$, $\mathrm{var}(x) = \dfrac{\sum_{i=1}^{n}(x - \bar{x})^2}{n}$.

При подстановке в формулу (1.22) выражения (1.26) получаем оценку уравнения парной линейной регрессии (функция регрессии):

$$\hat{y}_i = \bar{y} - b\bar{x} + bx_i. \tag{1.27}$$

МНК-оценки параметров уравнения регрессии при выполнении стандартных предположений регрессионного анализа будут обладать следующими *статистическими свойствами*:

1) *несмещенность*. Несмещенность статистической оценки характеризует равенство ее математического ожидания истинному значению этого параметра. В случае парной линейной регрессии: $M(a) = \alpha$, $M(b) = \beta$;
2) *состоятельность*. При неограниченном возрастании объема выборки значение оценки должно стремиться по вероятности к истинному значению параметра, а дисперсии оценок параметров должны уменьшаться и в пределе стремиться к 0: $\mathrm{var}(a) \to 0$, $\mathrm{var}(b) \to 0$ при $n \to \infty$;
3) *эффективность*. При наличии минимальной дисперсии по сравнению с другими оценками заданного класса оценка считается эффективной.

2. Использование оцененной модели для прогноза

Пусть модель зависимости имеет вид:

$$y_i = \alpha + \beta x_i + \varepsilon_i, \quad i = 1, 2, ..., n. \tag{1.28}$$

Подобранная модель после расчета методом наименьших квадратов коэффициентов α и β определена формулой

$$\hat{y} = a + bx, \tag{1.29}$$

где $\hat{y}$ – прогнозное значение объясняемой переменной.

Однако при выборе для y_0 значения $\hat{y}_0 = a + bx_0$ допускается ошибка прогноза.

$$\hat{y}_0 - y_0 = (a + bx_0) - (\alpha + \beta x_0 + \varepsilon_0) = (a - \alpha) + (b - \beta)x_0 + \varepsilon_0. \tag{1.30}$$

«Эта ошибка является следствием:

– неопределенности, связанной с отклонением вычисленных значений случайных величин а и b от истинных значений параметров α и β;
– неопределенности, связанной со случайной ошибкой ε_0 в $(n + 1)$-м наблюдении» [14].

Ошибка прогноза – случайная величина с математическим ожиданием, рассчитываемым согласно формуле (1.31).

$$M(\hat{y}_0 - y_0) = M(a - \alpha) + x_0 M(b - \beta) + M(\varepsilon_0) = 0; \quad [M(a) = \alpha;\ M(b) = \beta;\ M(\varepsilon_0) = 0]. \tag{1.31}$$

Теоретическая точность прогноза определяется дисперсией ошибки прогноза по формуле (1.32).

$$\mathrm{var}(\hat{y}_0 - y_0) = \mathrm{var}(\varepsilon_0) = 0. \tag{1.32}$$

Далее мы будем рассматривать оценку этой дисперсии $S\varepsilon_i$, то есть оценку дисперсии ошибки прогноза (дисперсию остатков):

$$S^2{}_{\varepsilon_i} = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n - m} = \frac{\sum_{i=1}^{n} e_i^2}{n - m}, \tag{1.33}$$

где n – количество наблюдений; m – количество параметров уравнения регрессии.

3. Интервальные оценки функции регрессии и ее параметров

Доверительный интервал для среднего значения y, полученного по функции регрессии $M\hat{y}(x)$, строится с использованием t-статистики (распределения Стьюдента).

Задается надежность (доверительная вероятность)

$$\gamma = 1 - \alpha\ (0{,}95 - 0{,}99),$$

с которой значение, полученное по уравнению регрессии, должно находиться в доверительном интервале.

Уравнение регрессии

$$\hat{y} = a + bx, \quad \hat{y} = \bar{y} - b\bar{x} + bx, \tag{1.34}$$

следовательно

$$\hat{y} = \bar{y} + b(x - \bar{x}).$$

Остаток в i-м наблюдении

$$e_i = y_i - \hat{y}_i = y_i - \bar{y} - b(x - \bar{x}) = y_i - (a + bx_i). \tag{1.35}$$

На рис. 1.8 линия регрессии изображена графически. Для произвольного наблюдения значения y_i отмечены: среднее значение – $\bar{y}$, приращение $b(x_i - \bar{x})$, составляющие расчетное значение $\hat{y}_i$, остаток регрессии e_i, $\mathrm{tg}\,\alpha = b$.

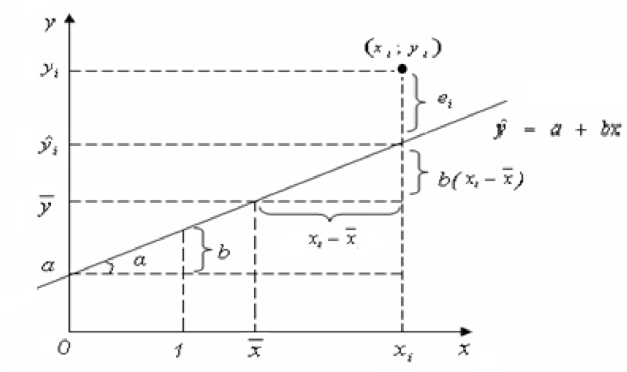


Рис. 1.8

Вариация (дисперсия) группового среднего значения $\hat{y}$ равна сумме дисперсий двух слагаемых выражения

$$\hat{y} = \bar{y} + b(x - \bar{x}), \tag{1.36}$$

$$\mathrm{var}(y) = \mathrm{var}(\bar{y}) + \mathrm{var}[b(x - \bar{x})]. \tag{1.37}$$

Запишем выражение (1.37) через стандартное отклонение и вынесем неслучайную величину $(x - \bar{x})$ за знак дисперсии, возведя ее в квадрат:

$$\sigma_{\hat{y}}^2 = \sigma_{\bar{y}}^2 + \sigma_b^2 (x - x_i)^2. \tag{1.38}$$

Следовательно, оценка дисперсии значения $\hat{y}$, найденного по уравнению регрессии, складывается из дисперсии среднего значения $\bar{y}$ и дисперсии b.

$$S_{\hat{y}}^2 = S_{\varepsilon_i}^2 \left(\frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \right). \tag{1.39}$$

По формуле (1.40) находят доверительный интервал для математического ожидания величины $\hat{y}$.

$$\hat{y} - t_{1-\alpha;k} S_{\hat{y}} \le M_{\hat{y}}(x) \le \hat{y} + t_{1-\alpha;k} S_{\hat{y}}, \tag{1.40}$$

где $t = \frac{\hat{y} - M_{\hat{y}}(x)}{S_{\hat{y}}}$ – статистика, имеющая t-распределение Стьюдента с $k = n - 2$ степенями свободы; n – объем выборки.

Величина доверительного интервала зависит от объясняющей переменой x (рис. 1.9). При $x = \bar{x}$ она стремится к минимуму, по мере удаления x от $\bar{x}$ увеличивается.

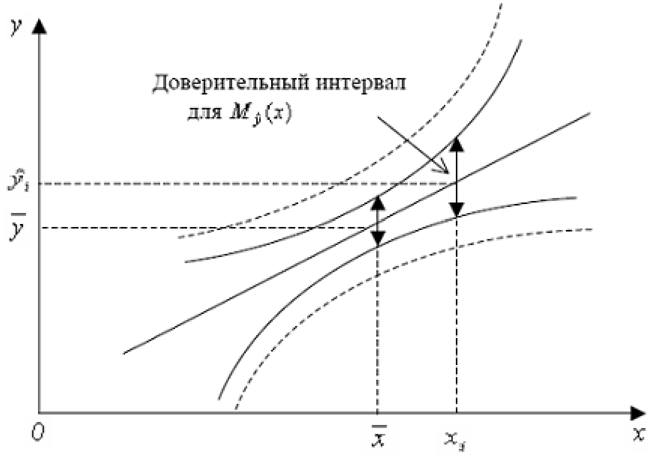


Рис. 1.9

Прогноз значений объясняемой переменной y по уравнению регрессии считается оправданным при условии, если значение переменной x не покидает диапазона наблюдаемых значений x_0.

Доверительный интервал для индивидуальных значений зависимой переменной. Доверительная область для $M\hat{y}(x)$ (рис. 1.9) определяет местоположение условного математического ожидания или среднего значения зависимой переменной (модельной линии регрессии), но не определяет расположение отдельных возможных значений зависимой переменной. Поэтому в оценку суммарной дисперсии $S^2_{\hat{y}}$ включают величину $S^2_{\varepsilon_i}$. В результате для оценки дисперсии индивидуальных значений y_0 при $x = x_0$ применяют формулу (1.41).

$$S^2_{\hat{y}_0} = S^2_{\varepsilon_i}\left[1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right]. \qquad (1.41)$$

По формуле (1.42) рассчитывают соответствующий доверительный интервал для прогнозов индивидуальных значений y_0.

$$\hat{y}_0 - t_{1-\alpha;n-2}S_{\hat{y}_0} \le y_0 \le \hat{y}_0 + t_{1-\alpha;n-2}S_{\hat{y}_0}. \qquad (1.42)$$

На рис. 1.9 данный доверительный интервал показан пунктиром.

Доверительные интервалы для параметров регрессионной модели

Так как статистика $t = \frac{b-\beta}{S}\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}$ имеет t-распределение Стьюдента с $k = n - 2$ степенями свободы, то для интервальной оценки параметра β на уровне значимости α применяют формулу (1.43).

$$b - t_{1-\alpha;n-2}\frac{S_{\varepsilon_i}}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}} \le \beta \le b + t_{1-\alpha;n-2}\frac{S_{\varepsilon_i}}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}}. \quad (1.43)$$

При $b > 0$ делается вывод о росте результативного признака при увеличении признака-фактора, при $b < 0$ – об уменьшении результативного признака при увеличении признака-фактора, при $b = 0$ – вывод о независимости от независимой переменной, что наглядно продемонстрировано на рис. 1.10. Таким образом, границы доверительного интервала для коэффициента b не должны содержать одновременно положительных и отрицательных значений, как, например, $-1{,}1 \le b \le 0{,}6$.

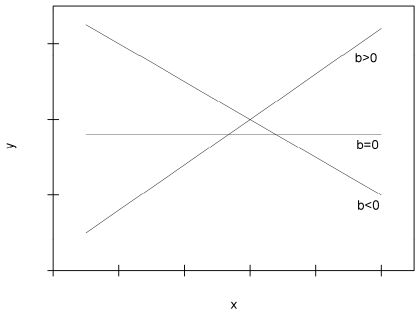


Рис. 1.10

Формула расчета интервальной оценки параметра α на уровне значимости α *представлена под номером* (1.44).

$$\alpha - t_{1-\alpha;n-2}S_{\varepsilon_i}\sqrt{\frac{\sum_{i=1}^{n}(x_i)^2}{n\sum_{i=1}^{n}(x_i-\bar{x})^2}} \le \alpha \le \alpha + t_{1-\alpha;n-2}S_{\varepsilon_i}\sqrt{\frac{\sum_{i=1}^{n}(x_i)^2}{n\sum_{i=1}^{n}(x_i-\bar{x})^2}}. \quad (1.44)$$

Так как статистика $\frac{nS^2}{\delta^2}$ имеет χ^2-распределение с $k = n - 2$ степенями свободы, то для интервальной оценки δ^2 на уровне значимости α *применяют формулу* (1.45):

$$\frac{nS_{\varepsilon_i}^2}{\chi^2_{\frac{\alpha}{2};n-2}} \le \delta^2 \le \frac{nS_{\varepsilon_i}^2}{\chi^2_{\frac{1-\alpha}{2};n-2}}. \qquad (1.45)$$

Доверительный интервал должен соответствовать требованию, чтобы для вероятности выполнялось условие (1.46).

$$P\left(\chi^2 < \chi^2_{\frac{1-\alpha}{2};n-2}\right) = P\left(\chi^2 > \chi^2_{\frac{1-\alpha}{2};n-2}\right) = \frac{\alpha}{2}. \qquad (1.46)$$

4. Оценка значимости уравнения регрессии (адекватности имеющимся статистическим данным)

Обозначим основную идею дисперсионного анализа. Для проверки значимости уравнения регрессии нужно установить, соответствует ли математическая модель экспериментальным данным. Также следует проверить, достаточно ли включенных в уравнение объясняющих переменных.

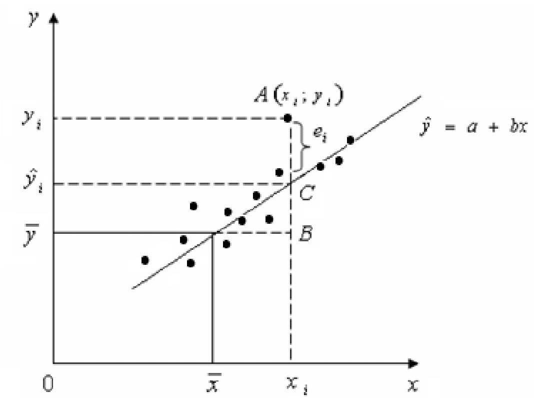


Рис. 1.11

На рис. 1.11 изображены наблюдаемые значения переменных x_i, y_i, соответствующая этим значениям линия регрессии, обозначены составляющие регрессии:

$$AB = y_i - \bar{y}; \quad AC = y_i - \hat{y}; \quad CB = \hat{y}_i - \bar{y}.$$

По рисунку

$$AB = CB + AC \quad \text{или} \quad y_i - \bar{y} = (y_i - \hat{y}) + (\hat{y}_i - \bar{y}),$$

где $\hat{y}_i = a + bx_i$ – ордината точки прямой, соответствующей уравнению регрессии, имеющей абсциссу x_i.

Возведя обе части в квадрат и просуммировав выражение для каждого i-го случая, имеем

$$\sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 + \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 + 2\sum_{i=1}^{n}(y_i - \hat{y}_i)(\hat{y}_i - \bar{y}) \approx \approx \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 + \sum_{i=1}^{n}(y_i - \hat{y}_i)^2, \qquad (1.47)$$

где $Q = \sum_{i=1}^{n}(y_i - \bar{y})^2$ — полная сумма квадратов; $Q_R = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$ — сумма квадратов, объясненная моделью; $Q_e = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} e_i^2$ — остаточная сумма квадратов.

Если поделить выражение на n, то получим

$$\mathrm{var}(y) = \mathrm{var}(\hat{y}) + \mathrm{var}(e), \qquad (1.48)$$

то есть дисперсия переменной y частично объясняется изменчивостью $\hat{y}$, а частично — изменчивостью остатка регрессии.

Оценка этой изменчивости

$$S_R^2 = \frac{Q_R}{m-1}. \qquad (1.49)$$

S_R^2 обусловлена уравнением регрессии или независимой переменной.

$$S_{\varepsilon_i}^2 = \frac{Q_e}{n-m}. \qquad (1.50)$$

$S_{\varepsilon_i}^2$ обусловлена воздействием неучтенных случайных факторов.

Процедура проверки значимости линейной связи между переменными. Эта процедура будет иметь смысл при соблюдении стандартных предположений о модели.

По формуле (1.51) определяют среднюю ошибку аппроксимации, которая позволяет судить «о качестве модели из относительных отклонений по каждому наблюдению» [27] и которая должна быть не более 8–10 %.

$$\bar{A} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \cdot 100\ \% = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{e_i}{y_i}\right| \cdot 100\ \%. \qquad (1.51)$$

На основе F-критерия Фишера производится оценка значимости уравнения регрессии в целом. Ранее проводится дисперсионный анализ, согласно которому «общая сумма квадратов отклонений переменной y от среднего значения $\bar{y}$ раскладывается на две части — «объясненную» (S_R^2) и «необъясненную» ($S_{\varepsilon_i}^2$)» [14].

Величины $S_R^2 = \frac{Q_R}{m-1}$ и $S_{\varepsilon_i}^2 = \frac{Q_e}{n-m}$ имеют χ^2-распределение (с $(m-1)$ и $(n-m)$ степенями свободы соответственно), и поэтому уравнение регрессии значимо на уровне α, если фактически наблюдаемое значение статистики удовлетворяет неравенству (1.52).

$$F = \frac{Q_R(n-m)}{Q_e(m-1)} = \frac{S_R^2}{S_{\varepsilon_i}^2} > F_{\alpha;k_1;k_2}, \qquad (1.52)$$

где $Q_R = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$ – сумма квадратов, объясненная моделью; $Q_e = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} e_i^2$ – остаточная сумма квадратов; m – число оцениваемых параметров уравнения регрессии; n – число наблюдений; $F_{\alpha;k_1;k_2}$ – табличное значение критерия.

При $m = 2$ получаем линейную парную регрессию, уравнение которой значимо на уровне значимости α, если выполняется неравенство (1.53).

$$F = \frac{Q_R(n-2)}{Q_e} > F_{\alpha;1;n-2}. \qquad (1.53)$$

Коэффициент детерминации – эффективная оценка значимости уравнения регрессии, характеризующая степень выраженности связи между переменными:

$$R^2 = \frac{Q_R}{Q} = 1 - \frac{Q_e}{Q}. \qquad (1.54)$$

«Величина R^2 показывает, какая доля вариации зависящей переменной обусловлена вариацией объясняющей переменной. Так как $0 \le Q_R \le Q$, то $0 \le R^2 \le 1$» [16].

Свойства коэффициента детерминации

1. При $R^2 \to 1$ регрессия хорошо отражает эмпирические данные, наблюдения примыкают к линии регрессии.

2. При $R^2 = 1$ точки $(x_i; y_i)$ лежат на линии регрессии и между переменными существует линейная функциональная зависимость.

3. При $R^2 = 0$ вариация зависимой переменной полностью обусловлена воздействием неучтенных в модели переменных и линия регрессии параллельна оси OX» [2].

При известном коэффициенте детерминации R^2 критерий значимости уравнения регрессии можно найти по формуле

$$F = \frac{R^2(n-m)}{(1-R^2)(m-1)} > F_{\alpha;k_1;k_2}. \qquad (1.55)$$

Для линейной регрессии применяют формулу

$$R^2 = r_{xy}^2. \tag{1.56}$$

Оценка статистической значимости коэффициентов парной линейной регрессии и корреляции. В парной линейной регрессии оценивается значимость как уравнения в целом, так и отдельных его параметров.

Стандартная ошибка коэффициента регрессии определяется по формуле

$$S_b = \frac{S_{\varepsilon_i}}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}}, \tag{1.57}$$

где ${S^2}_{\varepsilon_i} = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n-m} = \frac{\sum_{i=1}^{n} e_i^2}{n-m}$ – остаточная дисперсия.

Фактическое значение *t*-критерия Стьюдента $t_b = \frac{b}{S_b}$ сравнивается с табличным значением при заданном уровне значимости α и числе степеней свободы (n – 2). Доверительный интервал для коэффициента регрессии: $b \pm t_{1-\alpha;\, n-2} \cdot S_b$.

Стандартную ошибку параметра α рассчитывают по формуле

$$S_a = S_{\varepsilon_i}\sqrt{\frac{\sum_{i=1}^{n}(x_i)^2}{n\sum_{i=1}^{n}(x_i - \bar{x})^2}}. \tag{1.58}$$

Значим ли линейный коэффициент корреляции, проверяют по формуле

$$S_r = \sqrt{\frac{1 - r^2}{n - 2}}. \tag{1.59}$$

Фактическое значение *t*-критерия Стьюдента определяется отношением $t_r = \frac{r}{S_r}$.

Связь между *t*-критерием Стьюдента и *F*-критерием Фишера:

$$t_b = t_r = \sqrt{F}. \tag{1.60}$$

5. Проверка выполнения стандартных предположений об ошибках в линейной модели наблюдений графическим методом

Построенная модель проверяется при помощи дисперсионного анализа с использованием графиков стандартизированных остатков:

$$C_i = \frac{y_i - \hat{y}_i}{S_{\varepsilon_i}} = \frac{e_i}{S_{\varepsilon_i}}\,(i = 1, 2, \dots, n), \tag{1.61}$$

где $S_{\varepsilon_i}^2$ – оценка дисперсии остатков:

$$S_{\varepsilon_i}^2 = \frac{Q_e}{n-m} = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n-m}. \qquad (1.62)$$

По характеру поведения стандартизированных остатков можно судить о типичных отклонениях от стандартных предположений, так как они, по сути, имитируют поведение ошибок ε_i при большом количестве наблюдений.

Нарушения стандартных предположений о модели наблюдений легко видеть на построенном графике зависимости C_i от $\hat{y}_i$.

«1. Наблюдения со слишком большими по величине остатками (рис. 1.12). *Выделяющиеся наблюдения* – наличие наблюдений, для которых либо $M(\varepsilon_i) \neq 0$, либо $var(\varepsilon_i)$ существенно превышает величину дисперсий остальных ошибок.

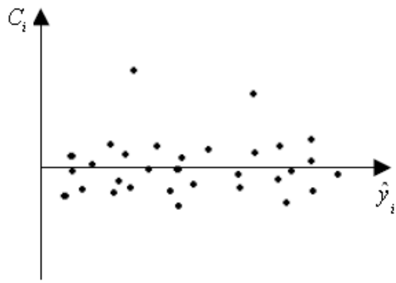


Рис. 1.12

2. Рост дисперсий ошибок с ростом значений $\hat{y}_i$ представлен на рис. 1.13: *неоднородность дисперсии* (гетероскедастичность), например в функциональной зависимости $var(\varepsilon_i)$ от величины $\hat{y}_i$.

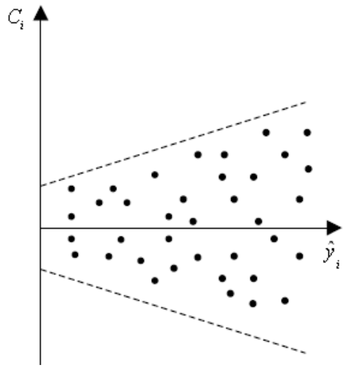


Рис. 1.13

3. На рис. 1.14 представлена *неправильная спецификация модели* в отношении множества объясняющих переменных.

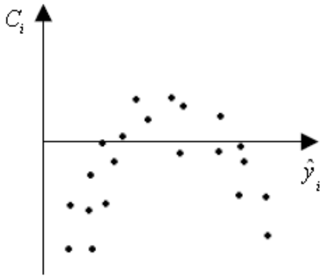


Рис. 1.14

График зависимости c_i от значений x_{ij} $(j = 1, 2, ..., n)$ j-й объясняющей переменной помогает выявить нелинейную зависимость y от j-й объясняющей переменной в случае множественной регрессионной модели (рис. 1.15, 1.16).

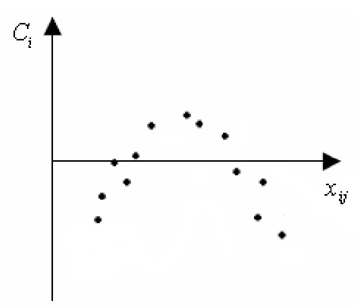


Рис. 1.15

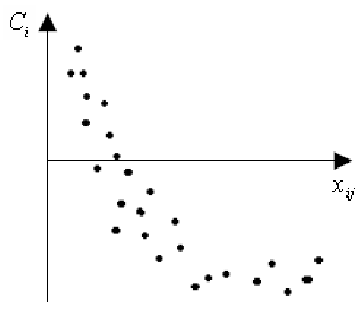


Рис. 1.16

График зависимости остатков от номера наблюдения полезен в случае, если наблюдения производятся через равные интервалы времени и по нему можно увидеть:

1) изменение дисперсии ошибок var(ε_i) с течением времени (рис. 1.17);

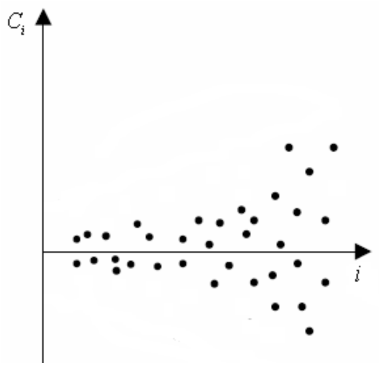


Рис. 1.17

2) невключение в модель переменных, зависящих от времени и существенно влияющих на объясняемую переменную (рис. 1.18);

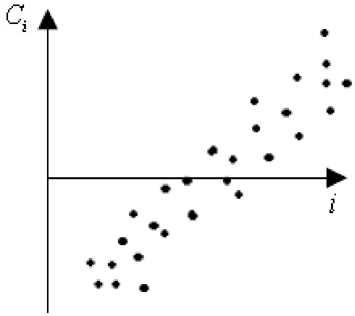


Рис. 1.18

3) невыполнение условия независимости в совокупности случайных ошибок ε_i в форме их автокоррелированности» [17]. График остатков в случае положительной автокоррелированности приведен на рис. 1.19, в случае отрицательной – на рис. 1.20.

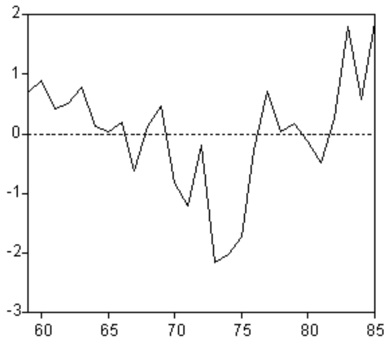


Рис. 1.19

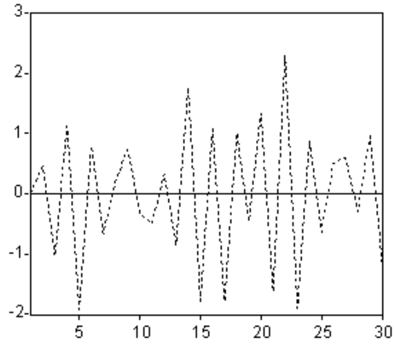


Рис. 1.20

Рекомендуемая литература

1. Айвазян С.А. Прикладная статистика. Основы эконометрики : учебник для экономических специальностей вузов. В 2 т. Т. 2. Основы эконометрики. – М. : ЮНИТИ-ДАНА, 2001. – 432 с.
2. Воскобойников Ю.Е. Эконометрика в Excel: парные и множественные регрессионные модели : учеб. пособие. – 2-е изд., стер. – СПб. : Лань, 2018. – 260 с.
3. Дорохина Е.Ю., Преснякова Л.Ф., Тихомиров Н.П. Сборник задач по эконометрике : учеб. пособие для студентов экономических вузов. – М. : Экзамен, 2003. – 224 с.

Тема 1.3. Нелинейные регрессионные модели

Форма проведения занятия: практическое занятие.

Вопросы для обсуждения

1. Изучение различных форм нелинейных моделей.
2. Изучение и применение метода линеаризации нелинейных уравнений регрессии.
3. Изучение и применение итерационных методов подбора нелинейных моделей.
4. Изучение и применение методов проверки качества моделей.

Методические указания по проведению занятия

1. Выполнить задания работы согласно варианту. Расчеты производить в Excel «Анализ данных», меню – Сервис.
2. Оформить отчет по проделанной работе, используя текстовый редактор MS Word, для ввода формул использовать надстройку Equation.
3. Исследования построенной модели произвести по порядку в соответствии с перечисленными в заданиях пунктами.

Используемые средства и материалы

1. Персональный компьютер со встроенным пакетом MS Office.
2. Принтер, бумага, необходимые для оформления отчета по практическому занятию.

Задания

Исследуется зависимость расходов на приобретение некоторого товара (группы товаров) семейными хозяйствами от располагаемого дохода.

В течение года i-я семья, имеющая располагаемый доход x_i, затратила на приобретение этого товара V_i руб.

Номер варианта находится по таблице по первой букве фамилии студента.

Буква	А	Б	В	Г	Д	Е	Ж
№ вар.	1	2	3	4	5	6	7
Буква	З	И	К	Л	М	Н	О
№ вар.	8	9	10	11	12	13	14
Буква	П	Р	С	Т	У	Ф	Х
№ вар.	15	16	17	18	19	20	21
Буква	Ц	Ч	Ш	Щ	Э	Ю	Я
№ вар.	22	23	24	25	26	27	28

Варианты заданий

Варианты							
с 1 по 7	**1**	**2**	**3**	**4**	**5**	**6**	**7**
x_i	V_i						
150537,1	3736,022	6107,689	4513,006	12360,85	6492,436	3304,019	3453,137
136570,9	3155,929	4962,273	3666,651	10441,58	5378,59	2818,311	2778,363
151518,1	4091,394	6706,055	4955,142	13536,61	7119,241	3615,949	3793,902
110318,6	3037,814	4385,618	3240,558	10050,79	4960,9	2771,366	2403,633
155144,1	3603,569	5962,619	4405,812	11922,62	6300,129	3177,29	3381,295
129398,2	3025,638	4655,843	3440,229	10010,5	5101,198	2716,574	2592,768
118036	3041,839	4511,824	3333,812	10064,1	5035,106	2756,336	2489,58
153232,6	3907,761	6433,963	4754,091	12929,05	6815,03	3449,771	3644,066
174761,2	3961,124	6873,953	5079,202	13105,61	7092,135	3451,21	3944,787
158744,2	4072,685	6800,956	5025,265	13474,71	7153,027	3582,684	3865,559
151702,4	3991,685	6545,808	4836,735	13206,72	6947,431	3527,398	3703,694
143872,3	3692,751	5928,584	4380,664	12217,68	6359,383	3280,574	3336,731
166110,4	4227,418	7188,594	5311,692	13986,66	7492,453	3701,97	4104,461
164493	3783,255	6408,18	4735,04	12517,12	6692,135	3316,259	3655,291
114337,7	3174,847	4649,526	3435,561	10504,17	5221,92	2886,033	2557,409
136811,3	3265,973	5138,914	3797,173	10805,66	5568,093	2916,069	2877,77
135744,2	3359,623	5269,74	3893,84	11115,51	5718,793	3002,036	2948,722
120100,7	2737,437	4088,58	3021,075	9056,97	4546,978	2476,208	2259,954
169115,2	3801,232	6510,397	4810,569	12576,6	6761,303	3322,795	3723,902
156830,3	3828,464	6362,19	4701,058	12666,7	6707,8	3371,934	3611,787
173678,5	3652,607	6322,829	4671,974	12084,86	6531,631	3184,386	3626,257
98372,26	2432,41	3354,253	2478,476	8047,768	3882,226	2244,641	1817,421
174902,9	3823,984	6638,118	4904,943	12651,87	6847,705	3331,455	3809,756
173312	4465,222	7722,974	5706,549	14773,45	7981,386	3893,657	4428,33
156933	4326,829	7192,263	5314,404	14315,56	7581,971	3810,621	4083,283

Варианты							
с 1 по 7	**1**	**2**	**3**	**4**	**5**	**6**	**7**
x_i	V_i						
140565,3	3504,023	5573,503	4118,293	11593,26	6006,37	3120,158	3129,598
176069,6	4505,115	7841,329	5794,002	14905,44	8078,157	3922,247	4503,297
161690,9	4461,26	7504,848	5545,374	14760,34	7864,375	3917,297	4273,493
172933,5	3684,829	6367,649	4705,091	12191,47	6583,589	3213,859	3650,392
155816,2	3777,938	6261,956	4626,994	12499,53	6610,692	3329,593	3552,58
142207,7	3406,962	5444,357	4022,866	11272,13	5853,578	3030,208	3060,634
145502,6	3913,684	6311,667	4663,726	12948,65	6755,063	3472,931	3556,343
98055,92	1978,545	2724,868	2013,42	6546,13	3155,805	1826,4	1475,929
151223,7	3533,754	5787,544	4276,448	11691,63	6146,527	3123,718	3273,624
136893,9	2937,451	4623,11	3416,042	9718,726	5008,606	2622,586	2589,078
168809,8	3807,876	6517,064	4815,495	12598,58	6770,674	3329,205	3727,042
148475	3600,899	5854,4	4325,849	11913,78	6240,382	3188,916	3305,371
132941,9	3695,458	5748,349	4247,487	12226,64	6264,266	3309,021	3209,829
166977,6	3968,5	6762,383	4996,763	13130,01	7040,89	3473,426	3863,119
154991,7	4603,27	7613,772	5625,859	15230,19	8046,326	4059,131	4317,21
159979,8	3234,597	5418,216	4003,55	10701,85	5689,875	2843,224	3082,019
169942,9	3819,756	6554,914	4843,463	12637,89	6800,891	3337,358	3751,197
174351,5	4093,711	7097,371	5244,287	13544,28	7326,082	3567,566	4072,045
151347,1	3071,714	5032,462	3718,514	10162,94	5343,737	2715,069	2846,757
190010,7	4094,785	7347,716	5429,269	13547,83	7455,147	3537,943	4252,092
167075,4	4366,154	7441,735	5498,739	14445,68	7747,313	3821,249	4251,459
161465,3	3937,949	6620,822	4892,163	13028,93	6939,937	3458,276	3769,575
109115,4	2428,233	3490,237	2578,956	8033,951	3956,737	2217,682	1910,803
143582,7	3847,491	6172,036	4560,552	12729,65	6623,195	3418,731	3473,051
124368,6	3400,901	5150,962	3806,075	11252,08	5688,604	3065,635	2857,143

Варианты							
с 8 по 14	**8**	**9**	**10**	**11**	**12**	**13**	**14**
x_i	V_i						
150537,1	3304,019	7059,309	6243,028	7178,972	6243,028	7982,318	5103,085
136570,9	2818,311	5848,209	5222,574	5889,718	5222,574	6548,793	4105,896
151518,1	3615,949	7740,841	6841,309	7877,173	6841,309	8758,649	5606,671
110318,6	2771,366	5394,05	4920,935	5317,599	4920,935	5912,653	3552,116
155144,1	3177,29	6850,21	6039,874	6987,361	6039,874	7769,265	4996,916
129398,2	2716,574	5546,597	4980,022	5555,91	4980,022	6177,631	3831,622
118036	2756,336	5474,735	4960,885	5433,758	4960,885	6041,81	3679,129
153232,6	3449,771	7410,068	6541,609	7549,064	6541,609	8393,824	5385,241
174761,2	3451,21	7711,368	6718,687	7959,975	6718,687	8850,717	5829,65
158744,2	3582,684	7777,577	6841,826	7951,515	6841,826	8841,31	5712,566
151702,4	3527,398	7554,031	6675,395	7688,006	6675,395	8548,315	5473,36
143872,3	3280,574	6914,638	6142,84	7000,079	6142,84	7783,406	4931,058
166110,4	3701,97	8146,64	7134,053	8366,695	7134,053	9302,951	6065,619
164493	3316,259	7276,443	6378,255	7465,685	6378,255	8301,114	5401,83
114337,7	2886,033	5677,861	5161,349	5617,453	5161,349	6246,061	3779,368
136811,3	2916,069	6054,258	5405,629	6098,301	5405,629	6780,717	4252,801
135744,2	3002,036	6218,116	5556,281	6258,449	5556,281	6958,786	4357,654
120100,7	2476,208	4943,987	4472,19	4915,499	4472,19	5465,557	3339,786
169115,2	3322,795	7351,651	6426,346	7563,78	6426,346	8410,187	5503,224
156830,3	3371,934	7293,477	6423,757	7447,549	6423,757	8280,949	5337,54
173678,5	3184,386	7101,926	6191,545	7326,331	6191,545	8146,166	5358,923
98372,26	2244,641	4221,193	3895,341	4113,943	3895,341	4574,304	2685,806
174902,9	3331,455	7445,597	6486,603	7686,258	6486,603	8546,371	5630,101
173312	3893,657	8678,262	7567,412	8950,584	7567,412	9952,178	6544,235
156933	3810,621	8243,974	7260,435	8418,676	7260,435	9360,748	6034,322

Варианты							
с 8 по 14	**8**	**9**	**10**	**11**	**12**	**13**	**14**
x_i	V_i						
140565,3	3120,158	6530,802	5815,355	6596,144	5815,355	7334,27	4624,956
176069,6	3922,247	8783,483	7647,083	9073,418	7647,083	10088,76	6655,024
161690,9	3917,297	8551,035	7508,404	8758,364	7508,404	9738,449	6315,416
172933,5	3213,859	7158,42	6243,48	7381,436	6243,48	8207,438	5394,591
155816,2	3329,593	7187,89	6334,869	7334,971	6334,869	8155,774	5250,042
142207,7	3030,208	6364,67	5660,842	6435,821	5660,842	7156,006	4523,04
145502,6	3472,931	7344,866	6517,699	7444,006	6517,699	8277,009	5255,604
98055,92	1826,4	3431,347	3167,486	3343,088	3167,486	3717,188	2181,144
151223,7	3123,718	6683,197	5907,718	6799,578	5907,718	7560,469	4837,798
136893,9	2622,586	5445,921	4862,173	5485,87	4862,173	6099,753	3826,169
168809,8	3329,205	7361,84	6436,416	7572,894	6436,416	8420,321	5507,864
148475	3188,916	6785,246	6008,939	6890,754	6008,939	7661,847	4884,715
132941,9	3309,021	6811,216	6098,962	6841,11	6098,962	7606,649	4743,521
166977,6	3473,426	7655,65	6700,6	7866,538	6700,6	8746,824	5708,961
154991,7	4059,131	8748,873	7714,695	8923,16	7714,695	9921,686	6380,022
159979,8	2843,224	6186,673	5438,112	6329,938	5438,112	7038,275	4554,642
169942,9	3337,358	7394,696	6460,817	7611,782	6460,817	8463,561	5543,56
174351,5	3567,566	7965,743	6941,945	8220,62	6941,945	9140,53	6017,714
151347,1	2715,069	5810,314	5135,7	5911,977	5135,7	6573,543	4206,97
190010,7	3537,943	8106,077	7003,746	8437,703	7003,746	9381,905	6283,79
167075,4	3821,249	8423,752	7372,45	8656,306	7372,45	9624,97	6282,853
161465,3	3458,276	7545,882	6626,734	7727,761	6626,734	8592,518	5570,72
109115,4	2217,682	4302,21	3929,167	4236,585	3929,167	4710,671	2823,807
143582,7	3418,731	7201,484	6398,958	7289	6398,958	8104,658	5132,514
124368,6	3065,635	6185,292	5575,535	6171,163	5575,535	6861,732	4222,318

Варианты							
с 15 по 21	**15**	**16**	**17**	**18**	**19**	**20**	**21**
x_i	V_i						
144354,6	3932,599	6322,109	4671,442	13011,23	6776,966	3492,481	3559,406
149545,6	3886,132	6336,321	4681,943	12857,49	6744,377	3439,044	3580,033
170410,3	3352,526	5759,444	4255,685	11092,03	5972,289	2928,33	3296,876
153039,5	3295,72	5423,528	4007,475	10904,08	5746,198	2909,829	3071,388
159254,2	3900,406	6521,631	4818,87	12904,72	6854,843	3430,032	3707,984
149145,2	3777,833	6153,137	4546,588	12499,18	6552,91	3344,101	3475,602
160664,6	4602,151	7722,165	5705,951	15226,48	8102,415	4043,583	4394,441
116919,8	3422,525	5057,22	3736,808	11323,62	5654,494	3104,24	2787,874
144054,5	3700,39	5943,856	4391,948	12242,95	6374,151	3286,944	3345,75
168181,5	3845,074	6570,918	4855,288	12721,65	6831,718	3362,981	3756,44
139014,7	3869,305	6127,275	4527,478	12801,82	6617,816	3449,247	3436,735
144764	3561,304	5731,7	4235,185	11782,78	6140,599	3161,844	3227,915
158904,1	3595,007	6005,703	4437,648	11894,29	6315,333	3162,159	3413,893
136085,9	3385,927	5316,342	3928,275	11202,54	5766,466	3024,78	2975,547
126789,5	3124,223	4768,541	3523,502	10336,67	5246	2810,808	2650,125
142027,4	3055,907	4880,89	3606,517	10110,65	5249,091	2718,319	2743,524
148287,3	3590,945	5835,264	4311,709	11880,85	6221,558	3180,504	3294,15
151869,5	3719,061	6101,429	4508,38	12304,73	6474,361	3286,122	3452,638
170769,1	3972,955	6831,049	5047,501	13144,75	7080,518	3469,526	3911,117
112560,6	2699,635	3928,889	2903,078	8931,9	4426,413	2457,899	2157,649
158135,5	4338,21	7233,232	5344,675	14353,22	7613,528	3817,729	4109,678
147938,8	3711,616	6025,678	4452,408	12280,1	6427,602	3288,155	3400,844
180925,9	3467,633	6101,596	4508,504	11472,86	6251,767	3010,79	3513,711
171602,9	3238,695	5579,432	4122,673	10715,41	5777,561	2826,93	3196,06
144688,6	3481,719	5602,445	4139,678	11519,47	6002,748	3091,347	3154,958

Варианты							
с 15 по 21	**15**	**16**	**17**	**18**	**19**	**20**	**21**
x_i	V_i						
125323,1	3476,83	5282,092	3902,967	11503,29	5824,507	3131,683	2932,119
178958,8	3883,503	6803,54	5027,174	12848,8	6986,245	3375,56	3913,657
113467,8	2908,816	4246,932	3138,082	9623,988	4777,056	2646,224	2334,184
128714,3	2995,496	4599,7	3398,744	9910,772	5045,028	2690,937	2560,146
145694,9	3762,642	6071,287	4486,108	12448,92	6496,079	3338,458	3421,352
130069,7	3945,544	6083,974	4495,483	13054,06	6659,04	3540,68	3389,827
147528,6	3937,871	6385,899	4718,577	13028,68	6815,634	3489,566	3603,149
126329,9	3174,962	4838,95	3575,527	10504,55	5327,327	2857,495	2688,279
145643	3586,361	5786,019	4275,322	11865,68	6191,294	3182,164	3260,479
138883,1	3569,578	5650,498	4175,184	11810,16	6104,024	3182,361	3169,014
128861,8	2984,161	4584,395	3387,435	9873,269	5027,089	2680,447	2551,919
193164,5	4359,968	7875,25	5819,067	14425,21	7964,13	3760,868	4564,883
144182,9	3298,282	5299,847	3916,087	10912,56	5682,508	2929,503	2983,508
163252,1	5028,73	8492,038	6274,814	16637,85	8881,772	4411,334	4840,278
155851,7	3293,124	5458,87	4033,59	10895,49	5762,619	2902,248	3097,037
170377,9	4275,861	7345,122	5427,352	14146,93	7616,856	3734,906	4204,485
143300,6	3499,246	5608,978	4144,505	11577,46	6021,346	3109,905	3155,593
178801,3	4350,788	7619,496	5630,088	14394,84	7825,491	3782,06	4382,64
120644,4	3071,097	4595,222	3395,435	10160,9	5105,808	2776,773	2541,147
106478,5	2658,765	3784,381	2796,3	8796,678	4311,237	2434,172	2066,777
130654,1	3383,745	5227,053	3862,299	11195,32	5715,994	3035,169	2913,679
140898,9	3245,975	5167,95	3818,628	10739,5	5566,68	2889,695	2902,563
137216,7	3046,9	4799,888	3546,664	10080,85	5197,676	2719,662	2688,712
152373,6	3325,377	5462,795	4036,49	11002,2	5792,852	2937,294	3092,277
174615,3	3845,7	6671,421	4929,55	12723,72	6884,325	3350,924	3828,24

Варианты							
с 22 по 28	**22**	**23**	**24**	**25**	**26**	**27**	**28**
x_i	V_i						
144354,6	3492,481	7368,681	6544,014	7462,229	6544,014	8297,273	5260,131
149545,6	3439,044	7333,247	6489,578	7452,628	6489,578	8286,597	5290,613
170410,3	2928,33	6493,746	5672,091	6686,22	5672,091	7434,425	4872,161
153039,5	2909,829	6247,914	5516,355	6364,308	5516,355	7076,491	4538,932
159254,2	3430,032	7453,358	6554,512	7622,489	6554,512	8475,466	5479,7
149145,2	3344,101	7125,062	6307,034	7239,113	6307,034	8049,189	5136,284
160664,6	4043,583	8809,859	7740,596	9017,72	7740,596	10026,83	6494,154
116919,8	3104,24	6148,203	5576,438	6096,39	5576,438	6778,592	4119,952
144054,5	3286,944	6930,695	6156,326	7017,223	6156,326	7802,468	4944,387
168181,5	3362,981	7428,213	6496,868	7638,322	6496,868	8493,07	5551,309
139014,7	3449,247	7195,635	6414,466	7259,571	6414,466	8071,936	5078,845
144764	3161,844	6676,751	5927,842	6763,431	5927,842	7520,276	4770,248
158904,1	3162,159	6866,742	6039,969	7021,017	6039,969	7806,687	5045,089
136085,9	3024,78	6269,952	5601,191	6312,208	5601,191	7018,561	4397,296
126789,5	2810,808	5704,043	5131,826	5701,995	5131,826	6340,064	3916,385
142027,4	2718,319	5707,403	5076,903	5770,475	5076,903	6416,206	4054,411
148287,3	3180,504	6764,779	5991,571	6869,099	5991,571	7637,769	4868,132
151869,5	3286,122	7039,655	6220,164	7165,296	6220,164	7967,112	5102,348
170769,1	3469,526	7698,737	6723,2	7928,594	6723,2	8815,824	5779,893
112560,6	2457,899	4812,895	4381,928	4754,236	4381,928	5286,248	3188,598
158135,5	3817,729	8278,286	7285,091	8460,171	7285,091	9406,886	6073,329
147938,8	3288,155	6988,813	6191,455	7094,918	6191,455	7888,859	5025,805
180925,9	3010,79	6797,626	5902,074	7041,141	5902,074	7829,064	5192,602
171602,9	2826,93	6282,015	5483,325	6472,726	5483,325	7197,041	4723,174
144688,6	3091,347	6526,865	5795,07	6611,254	5795,07	7351,07	4662,432

Варианты							
с 22 по 28	**22**	**23**	**24**	**25**	**26**	**27**	**28**
x_i	V_i						
125323,1	3131,683	6333,061	5704,375	6323,427	5704,375	7031,036	4333,118
178958,8	3375,56	7596,233	6602,683	7859,76	6602,683	8739,288	5783,646
113467,8	2646,224	5194,154	4725,253	5134,969	4725,253	5709,585	3449,484
128714,3	2690,937	5485,523	4927,798	5491,823	4927,798	6106,372	3783,412
145694,9	3338,458	7063,269	6266,987	7159,554	6266,987	7960,727	5056,112
130069,7	3540,68	7240,458	6497,495	7256,37	6497,495	8068,377	5009,524
147528,6	3489,566	7410,726	6567,054	7521,148	6567,054	8362,784	5324,774
126329,9	2857,495	5792,47	5213,275	5788,289	5213,275	6436,013	3972,768
145643	3182,164	6731,873	5973,164	6823,396	5973,164	7586,953	4818,372
138883,1	3182,361	6636,983	5917,023	6695,321	5917,023	7444,546	4683,204
128861,8	2680,447	5466,018	4909,713	5472,922	4909,713	6085,356	3771,255
193164,5	3760,868	8659,5	7469,604	9028,618	7469,604	10038,94	6746,035
144182,9	2929,503	6178,663	5487,83	6256,359	5487,83	6956,462	4409,062
163252,1	4411,334	9657,263	8471,605	9900,924	8471,605	11008,86	7153,017
155851,7	2902,248	6265,769	5522,055	6394,128	5522,055	7109,648	4576,837
170377,9	3734,906	8281,905	7234,131	8527,217	7234,131	9481,436	6213,436
143300,6	3109,905	6547,086	5818,63	6625,346	5818,63	7366,74	4663,371
178801,3	3782,06	8508,756	7396,504	8803,166	7396,504	9788,263	6476,715
120644,4	2776,773	5551,611	5019,561	5522,115	5019,561	6140,054	3755,335
106478,5	2434,172	4687,663	4291,683	4604,88	4291,683	5120,178	3054,306
130654,1	3035,169	6215,073	5574,828	6231,524	5574,828	6928,849	4305,868
140898,9	2889,695	6052,722	5388,371	6114,73	5388,371	6798,984	4289,44
137216,7	2719,662	5651,499	5044,527	5694,297	5044,527	6331,504	3973,409
152373,6	2937,294	6298,641	5563,568	6413,183	5563,568	7130,835	4569,802
174615,3	3350,924	7485,414	6522,365	7726,091	6522,365	8590,661	5657,415

Порядок выполнения задания

1. Подберите модель зависимости, в которой эластичность потребления рассматриваемого товара по отношению к располагаемому доходу не зависит от размера располагаемого дохода.

   Замечание. Постоянство эластичности предполагает оценивание модели, линейной в логарифмах уровней.
2. Постройте график подбора значений регрессии.
3. Рассчитайте среднюю ошибку аппроксимации. Сделайте выводы.
4. Проверьте значимость подобранной модели на уровне $\alpha = 0{,}05$.

   Замечание. Используйте коэффициент детерминации и критерий Фишера.
5. Оцените значение объясняемой переменной при $X = 153000$.
6. Найдите 95%-е доверительные интервалы для среднего и индивидуального значений объясняемой переменной при том же значении X.
7. Найдите с надежностью 0,95 интервальные оценки параметров уравнения регрессии α и β, дисперсии ошибок $\mathrm{var}(\varepsilon_i)$. Сделайте выводы.
8. С помощью графического метода оцените соответствие используемых для построения модели статистических данных стандартным предположениям регрессионного анализа.
9. В рамках подобранной модели проверьте гипотезы о том, что потребление данного товара эластично по отношению к располагаемому доходу.

   Замечание. Эластичное потребление соответствует значению эластичности, превышающему единицу по абсолютной величине ($|\eta| = |\beta| > 1$).
10. В рамках подобранной модели проверьте гипотезы о том, что потребление данного товара неэластично по отношению к располагаемому доходу ($|\eta| = |\beta| < 1$).

Критерии оценки:

– «зачтено» – выполнено не менее 70 % всех заданий;

– «не зачтено» – выполнено менее 70 % всех заданий.

Примеры решения задач

Исходные данные для исследования представлены в табл. 1.2.

Таблица 1.2

Номер наблюдения, i	Располагаемый доход семейного хозяйства, X (руб.)	Расходы семейного хозяйства на приобретение некоторого товара, V (руб.)
1	150527,1	6233,028
2	136571,9	5222,564
3	151519,1	6831,309
4	110318,5	4920,925
5	155144,2	6039,864
6	129399,2	4980,021
7	118036,1	4960,875
8	153231,6	6541,619
9	174762,2	6718,688
10	158744,1	6841,726
11	151702,3	6675,385
12	143873,3	6142,84
13	166110,4	7134,053
14	164493	6378,255
15	114337,7	5161,349
16	136811,3	5405,629
17	135744,2	5556,281
18	120100,7	4472,19
19	169115,2	6426,346
20	156830,3	6423,757

1. Для подтверждения предположения об эластичности потребления построим поле корреляции (рис. 1.21).

По форме облака рассеяния видно, что предположение о степенной модели наблюдений подтверждается. Линеаризуем степенную модель логарифмированием:

$$V = \alpha \cdot x^{\beta} \to \ln V = \ln \alpha x^{\beta},$$

$$\ln V = \ln \alpha + \ln x^{\beta},$$

$$\ln V = \ln \alpha + \beta \ln x.$$

Пусть $\ln V = V'$, $\ln \alpha = \alpha'$, $\ln x = x'$, откуда $V' = \alpha' + \beta' \cdot x'$.

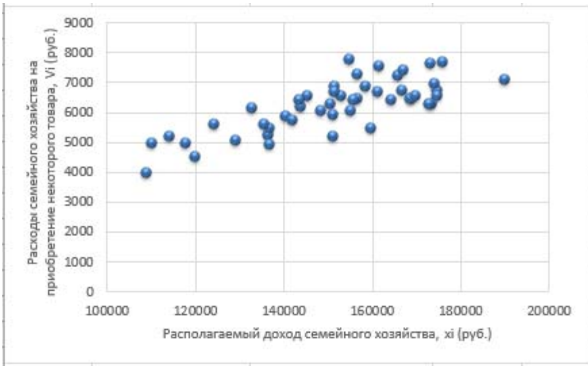


Рис. 1.21. Поле корреляции

Построим линейную модель регрессии с помощью надстройки MS Excel «Данные / Анализ данных / Регрессия» (рис. 1.22).

	A	B	C	D	E	F	G
71	ВЫВОД ИТОГОВ						
72							
73	*Регрессионная статистика*			R2=1-Ql/Q	0,694148	0,305852	
74	Множеств	0,861478699		F	108,9389		
75	R-квадрат	0,742145548		Fкр	4,042652		
76	Нормиров	0,73677358					
77	Стандартн	0,093687409					
78	Наблюдени	50					
79							
80	Дисперсионный анализ						
81		*df*	*SS*	*MS*	*F*	*Значимость F*	
82	Регрессия	1	1,212601628	1,212602	138,1515	9,86E-16	
83	Остаток	48	0,421311869	0,008777			
84	Итого	49	1,633913497				
85							
86		*Коэффициенты*	*Стандартная ошибка*	*t-статисти*	*P-Значение*	*Нижние 95%*	*Верхние 95%*
87	Y-пересеч	-3,417988676	1,030866298	-3,315647	0,001747	-5,490684	-1,345293
88	X'=lnxi	1,017735505	0,086587875	11,75379	9,86E-16	0,843639	1,191832

Рис. 1.22. Итоги линейной регрессии

Получили уравнение $\hat{V}' = 1{,}018x' - 3{,}418$, связывающее логарифмы уровней. Перейдем к исходной форме модели $V = \alpha \cdot x^{\beta}$. Для этого рассчитаем коэффициент $a = e^{\alpha'} = e^{-3{,}418} = 0{,}033$. Таким образом, подобранная модель с постоянной эластичностью $\eta = b = 1{,}018$ имеет вид: $\hat{V} = 0{,}033 \cdot x^{1{,}018}$.

2. Построим график подбора значений регрессии. Для этого вычислим значения $\hat{V}$, подставляя в найденную модель наблюдаемые значения X. Эти и дальнейшие вычисления отразим в табл. 1.3.

Таблица 1.3

i	x_i, руб.	V_i, руб.	$X'_i = \ln x_i$	$V' = \ln V_i$	$\hat{V}$	$V - \hat{V}$	$(V - \hat{V})^2$	A_i, %	$(V - \bar{V})^2$	C_i
1	150537,1	6243,028	11,92	8,74	6096,18	146,84	21563,24	2,35	26603,80	0,26
2	136570,9	5222,574	11,82	8,56	5521,06	-298,49	89095,83	5,72	735044,39	-0,53
3	151518,1	6841,309	11,93	8,83	6136,62	704,69	496590,20	10,30	579711,23	1,26
4	110318,6	4920,935	11,61	8,50	4442,93	478,01	228492,96	9,71	1343249,24	0,85
5	155144,1	6039,874	11,95	8,71	6286,11	-246,24	60632,02	4,08	1603,79	-0,44
6	129398,2	4980,022	11,77	8,51	5226,09	-246,07	60551,59	4,94	1209778,47	-0,44
7	118036	4960,885	11,68	8,51	4759,44	201,45	40581,21	4,06	1252242,24	0,36
8	153232,6	6541,609	11,94	8,79	6207,29	334,31	111765,98	5,11	213155,53	0,60
9	174761,2	6718,687	12,07	8,81	7095,92	-377,23	142305,94	5,61	408021,62	-0,67
10	158744,2	6841,826	11,98	8,83	6434,60	407,23	165836,70	5,95	580498,77	0,73
11	151702,4	6675,395	11,93	8,81	6144,21	531,18	282153,02	7,96	354588,93	0,95
12	143872,3	6142,84	11,88	8,72	5821,61	321,23	103190,53	5,23	3958,76	0,57
13	166110,4	7134,053	12,02	8,87	6738,60	395,45	156384,30	5,54	1111193,64	0,71
14	164493	6378,255	12,01	8,76	6671,83	−293,57	86184,82	4,60	89003,00	−0,53
15	114337,7	5161,349	11,65	8,55	4607,71	553,64	306513,34	10,73	843775,07	0,99
16	136811,3	5405,629	11,83	8,60	5530,95	−125,33	15706,42	2,32	454670,11	−0,22
17	135744,2	5556,281	11,82	8,62	5487,05	69,23	4792,67	1,25	274199,16	0,12
18	120100,7	4472,19	11,70	8,41	4844,18	−371,99	138376,16	8,32	2584799,93	−0,67
19	169115,2	6426,346	12,04	8,77	6862,68	−436,33	190383,69	6,79	120010,07	−0,78
20	156830,3	6423,757	11,96	8,77	6355,65	68,11	4638,65	1,06	118222,99	0,12
21	173678,5	6191,545	12,06	8,73	7051,18	−859,64	738977,37	13,88	12459,85	−1,54
22	98372,26	3895,341	11,50	8,27	3953,76	−58,42	3412,61	1,50	4772391,09	−0,10
23	174902,9	6486,603	12,07	8,78	7101,78	−615,17	378439,29	9,48	165390,01	−1,10
24	173312	7567,412	12,06	8,93	7036,04	531,37	282356,45	7,02	2212628,58	0,95
25	156933	7260,435	11,96	8,89	6359,89	900,55	810990,02	12,40	1393612,60	1,61
26	140565,3	5815,355	11,85	8,67	5685,45	129,91	16875,73	2,23	69995,33	0,23

Окончание табл. 1.3

i	x_i, руб.	V_i, руб.	$X'_i = \ln x_i$	$V' = \ln V_i$	$\hat{V}$	$V - \hat{V}$	$(V - \hat{V})^2$	A_i, %	$(V - \bar{V})^2$	C_i
27	176069,6	7647,083	12,08	8,94	7149,99	497,09	247098,43	6,50	2455995,79	0,89
28	161690,9	7508,404	11,99	8,92	6556,18	952,23	906737,48	12,68	2040562,82	1,70
29	172933,5	6243,48	12,06	8,74	7020,40	−776,92	603606,82	12,44	26751,45	−1,39
30	155816,2	6334,869	11,96	8,75	6313,83	21,04	442,82	0,33	64998,33	0,04
31	142207,7	5660,842	11,87	8,64	5753,06	−92,22	8504,85	1,63	175627,46	−0,16
32	145502,6	6517,699	11,89	8,78	5888,75	628,95	395574,52	9,65	191649,31	1,12
33	98055,92	3167,486	11,49	8,06	3940,82	−773,33	598044,20	24,41	8482279,38	−1,38
34	151223,7	5907,718	11,93	8,68	6124,48	−216,76	46986,94	3,67	29653,98	−0,39
35	136893,9	4862,173	11,83	8,49	5534,35	−672,18	451825,71	13,82	1482910,92	−1,20
36	168809,8	6436,416	12,04	8,77	6850,06	−413,65	171103,93	6,43	127088,47	−0,74
37	148475	6008,939	11,91	8,70	6011,21	−2,27	5,14	0,04	5038,49	0,00
38	132941,9	6098,962	11,80	8,72	5371,79	727,17	528779,73	11,92	362,55	1,30
39	166977,6	6700,6	12,03	8,81	6774,40	−73,80	5446,99	1,10	385242,05	−0,13
40	154991,7	7714,695	11,95	8,95	6279,83	1434,87	2058851,02	18,60	2672485,05	2,57
41	159979,8	5438,112	11,98	8,60	6485,57	−1047,46	1097171,45	19,26	411919,18	−1,87
42	169942,9	6460,817	12,04	8,77	6896,86	−436,04	190134,33	6,75	145081,53	−0,78
43	174351,5	6941,945	12,07	8,85	7078,99	−137,05	18781,78	1,97	743084,86	−0,25
44	151347,1	5135,7	11,93	8,54	6129,57	−993,87	987775,65	19,35	891553,86	−1,78
45	190010,7	7003,746	12,15	8,85	7726,56	−722,81	522461,39	10,32	853452,08	−1,29
46	167075,4	7372,45	12,03	8,91	6778,44	594,01	352845,58	8,06	1670630,44	1,06
47	161465,3	6626,734	11,99	8,80	6546,87	79,87	6378,79	1,21	299004,13	0,14
48	109115,4	3929,167	11,60	8,28	4393,61	−464,45	215710,85	11,82	4625744,06	−0,83
49	143582,7	6398,958	11,87	8,76	5809,68	589,28	347247,08	9,21	101784,42	1,05
50	124368,6	5575,54	11,73	8,63	5019,43	556,11	309252,88	9,97	254405,54	0,99
Сумма	7478971,88	303996,07					15007559,08	379,31	49068114,34	0,00
Среднее	149579,44	6079,92						7,59		
σ(X)	21517,71									

На одном графике изобразим поле корреляции и подобранную по модели кривую (рис. 1.23).

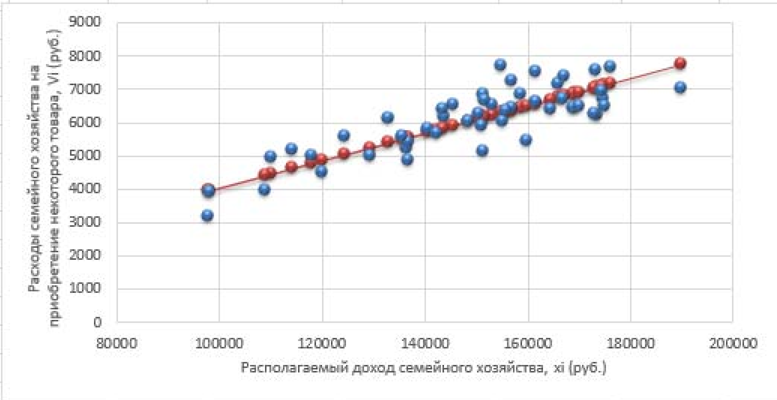


Рис. 1.23. График подбора

По графику видно, что подобранная модель хорошо аппроксимирует исходные данные.

3. Рассчитаем среднюю ошибку аппроксимации:

$$A_i = \left|\frac{V_i - \hat{V}_{x_i}}{V_i}\right| \cdot 100\ \%, \quad \bar{A} = \frac{1}{n}\sum A_i = 7{,}59\ \%.$$

В среднем расчетные значения отличаются от фактических на 7,59 %, что свидетельствует о хорошем подборе модели к исходным данным.

4. Проверим значимость подобранной модели на уровне $\alpha = 0{,}05$.

Коэффициент детерминации:

$$R^2 = \frac{Q_R}{Q} = 1 - \frac{Q_e}{Q} = 1 - \frac{15007559{,}08}{49068114{,}34} = 0{,}694.$$

Найденное значение $R^2 = 0{,}694$ показывает, что уравнением регрессии объясняется 69,4 % дисперсии результативного признака, а на долю прочих факторов приходится 30,6 %. То есть 69,4 % вариации расходов на приобретение некоторого товара (V) объясняется вариацией фактора X – дохода семейного хозяйства.

Оценим качество уравнения регрессии в целом с помощью F-критерия Фишера. Рассчитаем фактическое значение F-критерия:

$$F = \frac{R^2(n-m)}{(1-R^2)(m-1)} = \frac{0{,}694}{1-0{,}694} \cdot 48 = 108{,}94.$$

Найдем табличное значение F-критерия по прил. 1:

$$(k_1 = m - 1 = 2 - 1 = 1,\ k_2 = n - 2 = 50 - 2 = 48,\ \alpha = 0{,}05) : F_{\text{табл}} = 4{,}04.$$

Так как $F_{\text{факт}} > F_{\text{табл}}$, то признается статистическая значимость уравнения в целом.

5. Оценим значение объясняемой переменной при $X = 153000$.

$$\hat{V}(153000) = 0{,}018 \cdot 153000^{1{,}033} = 6197{,}705 \text{ (руб.)}.$$

Таким образом, если доход составит 153000 руб., то расходы на приобретение некоторого товара составят 6197,71 руб.

6. Найдем 95%-е доверительные интервалы для среднего и индивидуального значений объясняемой переменной при том же значении X.

Определим оценку дисперсии ошибки прогноза по формуле

$$S^2_{\varepsilon_i} = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n - m} = \frac{15007559{,}08}{50 - 2} = 312657{,}481.$$

Рассчитаем оценку дисперсии значения $\hat{y}$:

$$S^2_{\hat{y}} = S^2_{\varepsilon_i}\left(\frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right) =$$

$$= 312657{,}481\left(\frac{1}{50} + \frac{(153000 - 149579{,}44)^2}{22687573970}\right) = 10{,}895.$$

Доверительный интервал для математического ожидания (среднего значения) $\hat{y}$, найденного по уравнению регрессии, построим

$$\hat{y} - t_{1-\alpha;k}S_{\hat{y}} \le M_{\hat{y}}(x) \le \hat{y} + t_{1-\alpha;k}S_{\hat{y}},$$

$$6197{,}705 - 2{,}011 \cdot \sqrt{10{,}895} \le M_{\hat{y}}(x) \le 6797{,}705 + 2{,}011 \cdot \sqrt{10{,}895},$$

$$6191{,}069 \le M_{\hat{y}}(x) \le 6204{,}342.$$

Таким образом, с вероятностью 0,95 семьи с доходами 153000 руб. будут расходовать на некоторый товар в среднем от 6191,069 до 6204,342 руб.

Определим доверительный интервал для индивидуальных значений y_0 зависимой переменной:

$$S^2_{\hat{y}} = S^2_{\varepsilon_i}\left(\frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right) =$$

$$= 312657{,}481\left(1 + \frac{1}{50} + \frac{(153000 - 149579{,}44)^2}{22687573970}\right) = 11{,}895.$$

Найдем доверительный интервал прогноза:

$$\hat{y}_0 - t_{1-\alpha;n-2} S_{\hat{y}_0} \le y_0 \le \hat{y}_0 + t_{1-\alpha;n-2} S_{\hat{y}_0},$$

$$6197{,}705 - 2{,}011 \cdot \sqrt{11{,}895} \le y_0 \le 6197{,}705 + 2{,}011 \cdot \sqrt{11{,}895},$$

$$6190{,}771 \le y_0 \le 6204{,}64.$$

Таким образом, с вероятностью 0,95 семья с доходом 153000 руб. будет расходовать на некоторый товар от 6190,771 до 6204,64 руб.

7. Найдем с надежностью 0,95 интервальные оценки параметров уравнения регрессии α и β, дисперсии ошибок $\text{var}(\varepsilon_i)$. Сделаем выводы.

Случайные ошибки параметров линейной регрессии показаны на рис. 1.24, выпишем их:

$$S_\beta = \frac{S_{\varepsilon_i}}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}} = 0{,}004,$$

$$S_{\alpha'} = S_{\varepsilon_i} \sqrt{\frac{\sum_{i=1}^n (x_i)^2}{n \sum_{i=1}^n (x_i - \bar{x})^2}} = 1{,}089.$$

Рассчитаем доверительные интервалы для параметров регрессии a и b: $\alpha' \pm t_{1-a;\,n-2} \cdot S_{\alpha'}$ и $\beta \pm t_{1-a;\,n-2} \cdot S_\beta$. Получим, что $\alpha' \in [-6{,}915; -1{,}883]$ и $\beta \in [1{,}01; 1{,}025]$.

Интервалы для $\alpha \in [e^{-6{,}915}; e^{-1{,}883}] \Rightarrow \alpha \in [0{,}0009; 0{,}152]$.

Оцениваем доверительный интервал для дисперсии случайной с надежностью $\gamma = 0{,}95$:

$$\frac{(n-m) \cdot S^2}{{\chi_{\text{кр}_2}}^2} < \sigma^2 < \frac{(n-m) \cdot S^2}{{\chi_{\text{кр}_1}}^2}.$$

Находим критическую точку ${\chi_{\text{кр}_2}}^2 = \chi^2_{\frac{\alpha}{2};n-m}$ по таблице χ^2-распределения, где $\frac{\alpha}{2} = \frac{0{,}05}{2} = 0{,}025$, здесь $\alpha = 1 - \gamma = 0{,}975$; $n - m = 48$:

$${\chi_{\text{кр}_2}}^2 = \chi^2_{0{,}025;48} = 47{,}947.$$

Находим также критическую точку ${\chi_{\text{кр}_1}}^2 = \chi^2_{1-\frac{\alpha}{2};n-m}$;

$${\chi_{\text{кр}_1}}^2 = \chi^2_{0{,}975;48} = 46{,}728.$$

То есть интервальная оценка остаточной дисперсии $S^2 = 312657{,}481$ будет иметь вид:

$$\frac{48 \cdot 2312657{,}481}{47{,}947} < \sigma^2 < \frac{48 \cdot 312657{,}481}{46{,}728},$$

$$313\,003{,}09 < \sigma^2 < 321\,168{,}45.$$

8. С помощью графического метода оценим соответствие используемых для построения модели статистических данных стандартным предположениям регрессионного анализа (рис. 1.24).

Данные для построения графика зависимости стандартизированных остатков C_i (как ординат) от оцененных значений $\hat{V}$ (по оси абсцисс) показаны в табл. 1.3.

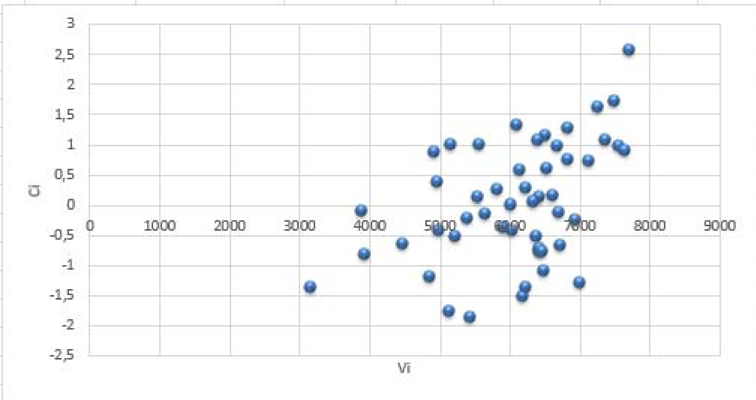


Рис. 1.24. График стандартизированных остатков

На графике не наблюдается функциональной зависимости $\mathrm{var}(\varepsilon_i)$ от величины $\hat{V}_i$, то есть дисперсия ошибок гомоскедастична. Судя по графику, условие $M(\varepsilon_i) = 0$ выполняется, то есть спецификация модели подобрана правильно.

Таким образом, используемые для построения модели статистические данные соответствуют стандартным предположениям регрессионного анализа.

9. Проверим гипотезу А о том, что потребление данного товара эластично по отношению к располагаемому доходу.

$$S_b = \frac{S\varepsilon_i}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}} = \frac{559{,}158}{\sqrt{22687573970}} = 0{,}0037.$$

Значение критерия Стьюдента: $t_{1-0,05;48} = 2,011$.

$$\left|\frac{1,072 - \beta_0}{0,0037}\right| > 2,011.$$

Данная гипотеза будет выполняться только при условии подстановки в неравенство чисел из интервала

$$b - t_{1-\alpha;\, n-2} \cdot S_b \leq \beta \leq b + t_{1-\alpha;\, n-2} \cdot S_b \text{ или } 1,065 \leq \beta \leq 1,072.$$

Таким образом, гипотеза А (потребление товара эластично по отношению к располагаемому доходу) с вероятностью 0,95 принимается для значений $1,065 \leq \beta \leq 1,072$ и отвергается с этой же вероятностью для других значений эластичности, больших 1.

10. В рамках подобранной модели проверим гипотезу о том, что потребление данного товара неэластично по отношению к располагаемому доходу ($|\eta| = |\beta| < 1$):

$$\left|\frac{1,072 - \beta_0}{0,0037}\right| > 2,011.$$

Нетрудно видеть, что при подстановке любого числа, меньшего 1 по абсолютной величине, данное неравенство выполняется. Так как наблюдаемое значение отношения больше табличного по абсолютной величине, такую гипотезу с вероятностью 0,95 следует отвергнуть. Это означает слишком большое отклонение оценки b от гипотетического значения β_0 параметра β в сравнении с оценкой S_b стандартного отклонения этого параметра.

Рекомендуемая литература

1. Доугерти К. Введение в эконометрику : [пер. с англ.]. – М. : ИНФРА-М, 1999. – 402 с.
2. Эконометрика: практикум для студентов экономических специальностей / П.Ф. Зибров [и др.] ; под ред. Ю.К. Черновой. – Тольятти : ТГУ, 2008. – 69 с.
3. Ивченко Ю.С. Эконометрика в MS EXCEL : лабораторный практикум. – Саратов : Ай Пи Эр Медиа, 2018. – 94 с.

Модуль 2. МОДЕЛЬ МНОЖЕСТВЕННОЙ РЕГРЕССИИ

Тема 2.1. Линейная модель множественной регрессии и корреляции. Оценка качества уравнений множественной регрессии

Форма проведения занятия: практическое занятие.

Вопросы для обсуждения

1. Построение и исследование модели множественной регрессии и корреляции.
2. Осуществление прогнозов по модели.

Методические указания по проведению занятия

1. Выполнить задания согласно варианту. Расчеты производить в Excel «Анализ данных», меню – Сервис.
2. Оформить отчет по проделанной работе, используя текстовый редактор MS Word, для ввода формул использовать надстройку Equation.
3. Исследования построенной модели произвести по порядку в соответствии с перечисленными в заданиях пунктами.

Отчет должен содержать:

1) расчетные формулы, графики;
2) порядок расчета значений параметров и характеристик, пояснения к расчетам, выводы по полученным данным;
3) анализ данных, произведенный в Excel, в качестве приложения и обоснования верности проведенных расчетов и выводов.

Используемые средства и материалы

1. Персональный компьютер со встроенным пакетом MS Office.
2. Принтер, бумага, необходимые для оформления отчета по практическому занятию.

Задания к практической работе

Номер варианта находится по таблице по первой букве имени студента.

Буква	А, Х	В, У	Д	Е	И	К	Л	Б	Н	С
№ вар.	1	2	3	4	5	6	7	8	9	10
Буква	О	Ш	Р, Щ	П	Г, Ж	Ф, Э	Ч, Ю	М, Я	Т	Ц, З
№ вар.	11	12	13	14	15	16	17	18	19	20

Варианты заданий

Вариант 1

Государство	y	x_1	x_2	x_3	x_4
Австрия	52	3,2	2,4	2,3	111
Азербайджан	34	2,4	2,3	2,4	100
Албания	55	2,7	2,6	2,7	102
Алжир	50	4,0	2,7	2,5	99
Бруней	51	2,7	3,0	2,2	97
Ботсвана	49	2,5	2,9	3,3	90
Бутан	59	5,3	1,7	2,0	81
Венгрия	52	2,9	2,3	1,8	82
Венесуэла	53	2,4	3,0	2,9	112
Нигерия	51	4,4	3,0	2,7	88
Вьетнам	49	5,2	2,9	2,8	65
Габон	60	3,9	3,2	2,9	111
Гаити	61	4,9	2,1	2,2	70
Гайана	58	6,3	3,2	2,6	96
Гамбия	63	6,6	3,3	4,1	50
Греция	60	8,0	2,9	2,9	75
Кения	49	5,1	3,3	2,7	112
Куба	58	7,2	3,0	3,1	93
Лаос	54	4,7	2,7	2,9	98
Молдавия	55	6,9	2,8	2,4	63

Для варианта 1 в таблице исходных данных введем обозначения:

y – средняя продолжительность жизни, ожидаемая при рождении ребенка, число лет;

x_1 – ВВП в соотношении к покупательской возможности;

x_2 – скорость прироста населения в соотношении с предыдущим годом, %;

x_3 – коэффициент роста задействованных на рабочих специальностях в соотношении с предыдущим годом, %;

x_4 – младенческая смертность, %.

Вариант 2

№ п/п	y	x_1	x_2	x_3	x_4	x_5	x_6	x_7
1	13,0	1	1	37,0	21,5	6,5	0	20
2	16,5	1	1	60,0	27,0	22,4	0	10
3	17,0	1	1	60,0	30,0	15,0	0	10
4	15,0	1	1	53,0	26,2	13,0	0	15
5	14,2	1	1	35,0	19,0	9,0	0	8
6	10,5	1	1	30,3	17,5	5,6	1	15
7	23,0	1	1	43,0	25,5	8,5	0	5
8	12,0	1	1	30,0	17,8	5,5	1	10
9	15,6	1	1	35,0	18,0	5,3	1	3
10	12,5	1	1	32,0	17,0	6,0	1	5
11	11,3	1	0	31,0	18,0	5,5	1	10
12	13,0	1	0	33,0	19,6	7,0	0	5
13	21,0	1	0	53,0	26,0	16,0	1	5
14	12,0	1	0	32,2	18,0	6,3	0	20
15	11,0	1	0	31,0	17,3	5,5	1	15
16	11,0	1	0	36,0	19,0	8,0	1	5

Для варианта 2 в таблице исходных данных приняты обозначения:

y – цены квартир в населенном пункте, тыс. долл.;

x_1 – число комнат в квартире;

x_2 – район города (1 – расположение близко к центру города, 0 – расположение далеко от центра);

x_3 – S общая (м^2);

x_4 – S жилая (м^2);

x_5 – S кухни (м^2);

x_6 – тип дома (1 – панельный, 0 – другой);

x_7 – расстояние от остановок общественного транспорта, минут пешком.

Вариант 3

Страны мира	y	x_1	x_2	x_3	x_4	x_5	x_6
Австралийский Союз	0,914	76,5	105	3113	23,2	56,2	74,2
Республика Ангола	0,912	75,8	102	3101	21,3	61,3	78,2
Белиз	0,763	78,4	74	3101	25,7	59,3	68,0
Республика Бенин	0,924	71,7	111	3243	17,8	63,3	77,2
Босния и Герцеговина	0,918	84,4	113	3237	15,9	64,1	77,2
Федеративная Республика Бразилия	0,916	75,9	111	3330	22,4	57,0	77,2
Буркина-Фасо	0,905	76,2	119	3818	20,6	50,7	75,7
Республика Вануату	0,546	67,5	146	2415	23,2	57,1	62,6
Республика Джибути	0,894	78,2	113	3295	20,7	62,0	78,0
Арабская Республика Египет	0,923	78,1	112	3504	16,5	61,6	78,2
Исламская Республика Иран	0,932	78,6	123	3056	19,7	58,6	79,0
Государство Катар	0,740	84,0	71	3007	18,5	71.7	67,6
Ливанская Республика	0,701	59,2	212	2844	42,4	48,0	69,8
Республика Парагвай	0,744	91,2	100	2861	24,0	63,9	68,4
Независимое Государство Папуа Новая Гвинея	0,922	72,8	118	3239	20,2	59,1	74,9

Для варианта 3 в таблице исходных данных приняты обозначения:

y – индекс развития гражданина;

x_1 – расходы на конечное потребление в текущих ценах, % к ВВП;

x_2 – ВВП, % к предыдущему десятилетию;

x_3 – калорийность питания в сутки для населения страны, ккал на душу населения;

x_4 – валовое накопление, % к ВВП;

x_5 – затраты семей, % к ВВП;

x_6 – средняя продолжительность жизни, ожидаемая при рождении ребенка, число лет.

Вариант 4

№ п/п	y	x_1	x_2	x_3
1	6,5	82,6	6,8	221,0
2	3,2	6,6	17,8	32,5
3	6,3	51,2	106,8	81,3
4	3,5	14,9	17,5	50,2
5	0,2	30,0	69,6	300,3
6	3,7	14,3	15,2	42,6
7	1,4	5,7	3,9	16,5
8	5,2	24,2	52,3	136,0
9	1,9	12,2	17,8	83,2
10	3,1	15,4	33,3	113,3
11	4,5	31,3	72,0	215,5
12	2,8	26,3	92,7	612,0
13	1,7	6,3	10,5	42,8
14	1,4	11,5	30,5	100,3
15	3,2	14,2	33,7	115,2
16	2,8	6,7	12,8	48,2
17	2,7	22,9	63,8	51,4
18	1,8	16,0	29,3	471,0
19	2,4	9,7	13,1	73,0
20	0,8	19,8	32,5	45,3

Для варианта 4 в таблице исходных данных приняты обозначения:
y – чистый доход, млрд долл. США;
x_1 – использованный капитал, млрд долл.;
x_2 – оборот капитала, млрд долл.;
x_3 – численность служащих, тыс. чел.

Вариант 5

Номер предприятия	y	x_1	x_2	x_3
1	7,0	10,0	3,9	36
2	7,0	14,0	3,9	37
3	7,0	15,0	3,7	42
4	7,0	16,0	4,0	39
5	7,0	17,0	3,8	41
6	7,0	19,0	4,8	38
7	8,0	19,0	5,4	46
8	8,0	20,0	4,4	41
9	8,0	20,0	5,3	40
10	10,0	20,0	6,8	45
11	9,0	21,0	6,0	45
12	11,0	22,0	6,4	46
13	9,0	22,0	6,8	49
14	11,0	25,0	7,2	44
15	12,0	28,0	8,0	45
16	12,0	29,0	8,2	46
17	12,0	30,0	8,1	48
18	12,0	31,0	8,5	48
19	14,0	32,0	9,6	51
20	14,0	36,0	9,0	50

Для варианта 5 в таблице исходных данных приняты обозначения:

y – выработка продукции на одного сотрудника предприятия, тыс. руб.;

x_1 – отношение количества рабочих высокой квалификации к общему числу рабочих предприятия, %;

x_2 – коэффициент отношения стоимости ввода новых основных фондов к стоимости фондов на конец года, %;

x_3 – внедрение рацпредложений, %.

Вариант 6

Государство	y	x_1	x_2	x_3
Республика Науру	1612	0,836	13,8	71,6
Королевство Нидерландов	7111	0,843	11,9	58,8
Республика Палау	6636	0,799	12,7	63,6
Румыния	6331	0,822	16,9	70,3
Словацкая Республика	6103	0,874	11,3	71,3
Республика Уганда	4200	0,698	11,0	64,7
Черногория	4000	0,513	33,8	72,3
Республика Эль-Сальвадор	3690	0,666	40,7	73,4
Тоголезская Республика	3655	0,718	22,6	71,3
Королевство Тонга	3180	0,713	20,5	69,8
Республика Панама	2683	0,667	16,8	64,3
Республика Никарагуа	2605	0,598	21,5	63,7
Монголия	2611	0,666	16,9	70,8
Республика Молдова	2221	0,523	16,7	72,6
Республика Мали	2155	0,455	47,8	69,9
Ливанская Республика	1372	0,338	42,3	68,7
Королевство Лесото	1288	0,389	41,5	65,7

Для варианта 6 в таблице исходных данных приняты обозначения:
y – среднедушевой доход, долл.;
x_1 – индекс развития человека;
x_2 – индекс бедности;
x_3 – средняя по стране продолжительность жизни, число лет.

Вариант 7

Номер фирмы	y	x_1	x_2	x_3	x_4
1	1,0	1,5	27	2,64	14,5
2	1,1	1,4	28	3,45	15,6
3	1,6	1,6	28	4,21	17,4
4	0,9	1,5	26	3,45	16,8
5	0,85	1,5	25	2,88	12,0
6	0,95	1,4	22	3,12	14,8
7	1,0	1,5	28	5,23	15,5
8	1,1	1,6	26	4,43	12,9
9	1,2	1,4	25	3,34	16,2
10	1,4	1,3	25	3,32	16,5
11	1,3	1,3	28	4,1	15,6
12	1,1	1,6	31	5,15	16,0
13	1,0	1,4	32	4,3	14,4
14	1,5	1,5	29	2,9	15,2
15	1,2	1,7	27	3,48	13,6

Для варианта 7 в таблице исходных данных приняты обозначения:
y – себестоимость единицы продукции, тыс. руб.;
x_1 – оптовая цена за 1 т энергоносителя, млн руб.;
x_2 – доля прибыли, изымаемая государством, %;
x_3 – объем производства, млн руб.;
x_4 – трудоемкость единицы продукции, человеко-часов.

Вариант 8

Номер района	y	x_1	x_2	x_3	x_4
1	33	27	0,5	1223	10
2	37	28	0,1	1348	15
3	25	25	0,1	1541	13
4	18	27	0,2	1720	8
5	34	30	0,1	2678	17
6	39	33	0,2	1738	16
7	56	39	0,2	1363	26
8	43	31	0,3	1458	22
9	44	44	0,1	1289	27
10	48	41	0,1	1946	23
11	51	43	0,2	1352	18
12	50	40	0,3	1425	15
13	51	40	0,3	1382	21
14	55	39	0,1	1623	31
15	56	43	0,1	1296	30
16	47	39	0,1	1362	24
17	43	40	0,1	1634	13
18	78	68	0,2	1295	32
19	71	66	0,1	1362	33
20	59	48	0,1	1348	25

Для варианта 8 в таблице исходных данных приняты обозначения:

y – число абитуриентов, поступивших в вузы, %;

x_1 – число абитуриентов, имеющих аттестаты с положительными оценками, %;

x_2 – число абитуриентов, окончивших школу с медалью, %;

x_3 – среднедушевой доход, руб.;

x_4 – количество школ с углубленным изучением отдельных дисциплин, %.

Вариант 9

Номер предприятия	y	x_1	x_2	x_3
1	9,0	350	18	108
2	9,3	305	20	113
3	7,4	300	19	71
4	7,0	284	42	210
5	7,4	286	23	94
6	9,2	326	20	118
7	9,3	335	25	130
8	8,0	334	22	127
9	7,4	270	23	61
10	9,3	364	18	117
11	7,2	275	20	46
12	9,1	292	17	107
13	9,2	355	17	110
14	8,3	318	30	99
15	9,1	328	20	101
16	9,2	316	14	105
17	8,9	330	21	113
18	5,5	241	25	146
19	9,1	381	21	119
20	9,1	333	22	110

Для варианта 9 в таблице исходных данных приняты обозначения:
y – инвестиции, млн руб.;
x_1 – совокупный доход, млн руб.;
x_2 – запасы капитала, млн руб.;
x_3 – налоги, млн руб.

Вариант 10

Номер области	y	x_1	x_2	x_3
1	60	43	61,5	38
2	65	50	62,0	49
3	64	48	53,9	47
4	64	45	58,3	27
5	67	43	62,1	31
6	66	48	49,2	60
7	72	45	39,3	58
8	73	48	35,5	57
9	71	50	33,3	51
10	75	48	44,5	34
11	81	51	42,5	29
12	80	53	39,3	32
13	80	52	35,4	52
14	82	52	36,7	49
15	85	54	33,9	48
16	87	53	36,4	42
17	89	58	39,7	51
18	91	58	44,1	53
19	90	57	43,0	42
20	92	56	48,0	59
21	93	58	53,6	56
22	90	61	59,4	54
23	91	62	61,7	39
24	92	63	46,5	47
25	95	59	57,9	48

Для варианта 10 в таблице исходных данных приняты обозначения:

y – душевой доход в день, руб.;
x_1 – среднедневная заработная плата рабочего, руб.;
x_2 – возраст безработных в среднем по области, лет;
x_3 – неработающие пенсионеры, %.

Вариант 11

Государство	y	x_1	x_2	x_3
Республика Маршалловы Острова	0,914	77,4	3351	10
Мальдивская Республика	0,823	77,3	3111	13
Республика Кот-д'Ивуар	0,837	71,9	3236	12
Демократическая Республика Конго	0,743	65,1	3122	16
Лаосская Народно-Демократическая Республика	0,823	78,3	3544	7
Корейская Народно-Демократическая Республика (Северная Корея)	0,799	64,9	2940	40
Соединенное Королевство Великобритании и Северной Ирландии	0,917	73,2	3136	8
Республика Ботсвана	0,797	70,1	3411	20
Габонская Республика	0,916	71,9	3333	23
Кооперативная Республика Гайана	0,869	77,1	3576	27
Боливарианская Республика Венесуэла	0,915	75,9	3708	8
Исламская Республика Афганистан	0,617	66,4	3288	57
Аргентинская Республика	0,888	77,3	3273	10
Антигуа и Барбуда	0,546	62,3	2414	50
Княжество Андорра	0,895	78,1	3245	20
Демократическая Республика Восточный Тимор	0,901	78,4	3503	21
Королевство Нидерландов	0,933	79,1	3052	15
Федеративная Республика Сомали	0,733	64,5	3111	30
Республика Таджикистан	0,743	68,5	2132	33
Япония	0,745	64,4	2322	18

Для варианта 11 в таблице исходных данных приняты обозначения:
y – индекс развития человека;
x_1 – продолжительность жизни, ожидаемая при рождении, число лет;
x_2 – среднесуточная калорийность питания граждан страны, ккал;
x_3 – младенческая смертность, %.

Вариант 12

Номер области	y	x_1	x_2	x_3
1	68	7,4	3,1	46
2	59	7,4	2,8	73
3	47	4,9	3,1	124
4	60	8,3	2,9	90
5	51	5,7	2,5	96
6	57	7,5	2,4	55
7	67	7,0	3,0	45
8	69	10,8	1,1	34
9	57	7,8	2,9	56
10	51	7,6	2,9	90
11	72	12,1	1,3	16
12	63	14,2	2,0	56
13	64	14,1	1,6	51
14	66	10,6	2,2	39
15	65	12,4	2,0	55
16	57	9,0	2,3	64
17	66	12,4	2,9	44
18	69	15,6	2,2	36
19	71	14,3	1,9	37
20	74	13,1	1,0	13

Для варианта 12 в таблице исходных данных приняты обозначения:
y – % населения, способного к труду;
x_1 – ежегодный прирост населения, %;

x_2 – ежегодный прирост трудоспособного населения, %;
x_3 – число действующих фирм и учреждений, использующих рабочую силу.

Вариант 13

Номер области	y	x_1	x_2	x_3
1	12,4	2,0	55	65
2	9,0	2,3	64	57
3	12,4	2,9	44	66
4	15,6	2,2	36	69
5	14,3	1,9	37	71
6	13,1	1,0	13	74
7	19,6	2,2	34	70
8	9,7	2,2	36	67
9	13,5	2,7	41	68
10	18,5	1,9	39	69
11	15,6	0,2	13	70
12	14,0	2,0	47	66
13	28,0	0,9	35	69
14	22,2	1,7	23	73
15	20,7	1,7	48	67
16	20,0	0,3	14	70
17	13,4	0,3	11	72
18	29,3	2,3	23	71
19	18,6	2,2	50	64
20	23,7	1,9	33	72

Для варианта 13 в таблице исходных данных приняты обозначения:
y – % семей, имеющих двух и более детей;
x_1 – прирост населения за год, %;
x_2 – % семей, не нуждающихся в жилье;
x_3 – средняя продолжительность жизни, число лет.

Вариант 14

Номер завода	y	x_1	x_2	x_3
1	1,8	49,0	16	71
2	1,6	20,0	44	67
3	1,8	31,9	13	72
4	2,7	33,4	12	71
5	2,1	35,3	12	72
6	1,0	24,6	18	73
7	2,0	30,8	22	73
8	0,9	43,4	9	78
9	1,9	42,4	10	72
10	1,0	53,8	7	77
11	1,5	60,6	7	76
12	1,7	58,1	6	77
13	3,5	61,1	8	77
14	1,4	70,2	6	77
15	0,4	73,7	7	78
16	1,0	78,3	6	78
17	0,1	65,8	5	76
18	1,3	85,1	5	79
19	0,3	68,7	4	79
20	0,6	73,9	6	78

Для варианта 14 в таблице исходных данных приняты обозначения:

y – темп обновления основных фондов предприятия;
x_1 – средний возраст рабочего;
x_2 – степень травматизма, чел/год;
x_3 – % работников, являющихся членами профсоюза.

Вариант 15

Подразделение	y	x_1	x_2	x_3	x_4
1	6	1,1	70,2	7,7	1,4
2	7	0,2	73,7	7,8	0,4
3	6	1,3	78,3	7,8	1,0
4	5	0,5	65,8	7,6	0,1
5	5	1,6	85,1	7,9	1,3
6	4	0,6	68,7	7,9	0,3
7	6	0,7	73,9	7,8	0,6
8	8	0,4	80,3	7,7	0,5
9	6	0,5	78,0	7,8	0,8
10	4	2,0	84,4	7,6	1,7
11	6	0,8	78,7	7,7	0,5
12	8	1,0	100,0	7,7	1,1
13	6	0,3	78,7	7,5	0,1
14	4	0,3	82,0	8,0	0,6
15	6	1,0	95,9	7,8	0,8

Для варианта 15 в таблице исходных данных приняты обозначения:

y – доля брака, %;

x_1 – степень обновления инструментария, %;

x_2 – оснащенность рабочего места, %;

x_3 – выполненная высококвалифицированная работа, %;

x_4 – ежегодное увеличение заработной платы, %.

Вариант 16

Номер предприятия	y	x_1	x_2	x_3	x_4
1	23,9	72	53	74	16
2	32,5	98	51	150	10
3	43,0	100	45	350	25
4	17,8	59	39	62	10
5	28,0	75	40	180	3
6	32,7	85	59	90	5
7	31,0	66	48	60	2
8	33,0	81	52	120	10
9	28,0	76	49	100	5
10	21,5	55	41	60	15
11	15,3	53	38	55	3
12	21,0	57	38	63	7
13	35,5	62	52	80	3
14	22,0	74	47	100	15
15	29,0	70	45	90	2

Для варианта 16 в таблице исходных данных приняты обозначения:
y – цена детали, у. е.;
x_1 – число изготовленных деталей;
x_2 – число деталей высшего качества;
x_3 – количество занятых на рабочих специальностях;
x_4 – количество менеджеров.

Вариант 17

Номер предприятия	y	x_1	x_2	x_3
1	8,0	29,1	15,5	924
2	6,4	29,3	14,2	695
3	7,9	31,6	14,6	923
4	7,9	32,8	15,0	914
5	7,4	31,8	13,8	833
6	7,8	35,5	14,4	918
7	7,7	29,2	14,1	913
8	6,9	27,5	13,9	721
9	6,9	35,6	13,3	728
10	7,7	36,4	12,8	927
11	7,0	29,4	14,6	752
12	6,7	27,0	13,6	747
13	7,2	33,3	12,6	852
14	7,3	33,4	15,0	802
15	7,8	33,5	15,1	927
16	7,8	32,6	13,4	921
17	6,8	28,6	13,9	744
18	7,0	28,4	14,3	701
19	6,8	30,0	14,7	740
20	7,9	30,6	14,0	932

Для варианта 17 в таблице исходных данных приняты обозначения:

y – доход, млн руб.;

x_1 – выработка изделий на одного рабочего, тыс. руб.;

x_2 – потребление, тыс. руб.;

x_3 – запас средств, тыс. руб.

Вариант 18

Номер факультета	y	x_1	x_2	x_3
1	10,0	80	110	4,3
2	9,0	62	92	3,0
3	8,5	76	115	3,7
4	8,2	51	75	2,6
5	9,0	46	74	2,3
6	9,0	62	88	3,4
7	9,0	74	106	4,6
8	6,5	49	74	2,3
9	33,0	110	176	6,5
10	9,5	73	92	2,4
11	9,0	69	96	3,8
12	15,0	129	176	7,5
13	10,0	66	93	2,7
14	9,5	76	107	4,3
15	16,5	81	116	3,5
16	8,0	62	90	4,6
17	12,0	75	114	3,6
18	25,6	74	115	3,1
19	12,5	56	87	4,1
20	11,8	66	102	2,7

Для варианта 18 в таблице исходных данных приняты обозначения:

y – студенты, успешно сдавшие сессию, %;

x_1 – количество студентов, занимающих бюджетные места;

x_2 – количество студентов факультета;

x_3 – число восстановленных после академического отпуска студентов, %.

Вариант 19

Номер района	y	x_1	x_2	x_3	x_4
1	9,70	0,14	0,32	0,25	1,59
2	9,95	0,66	0,77	0,26	0,46
3	8,40	0,46	0,59	0,29	0,28
4	8,78	0,64	0,55	0,27	1,13
5	10,5	0,82	0,76	0,24	0,64
6	11,2	0,89	0,99	0,31	0,59
7	12,0	0,20	0,64	0,27	0,73
8	10,1	0,35	0,38	0,27	1,34
9	9,67	0,42	0,48	0,28	1,06
10	9,55	0,51	0,62	0,26	1,33
11	10,3	0,44	0,81	0,23	0,95
12	11,6	0,37	0,75	0,28	0,86
13	12,4	0,28	0,46	0,27	0,97
14	11,3	0,19	0,42	0,26	1,05
15	12,1	0,30	0,57	0,26	1,38
16	11,7	0,38	0,76	0,25	1,25
17	10,4	0,47	0,79	0,28	0,77
18	9,93	0,40	0,77	0,26	0,94
19	9,48	0,63	0,86	0,25	0,82
20	8,85	0,52	0,88	0,26	0,78

Для варианта 19 в таблице исходных данных приняты обозначения:

y – урожайность зерновых культур, ц/га;

x_1 – масса химических средств защиты на 1 га, ц/га;

x_2 – масса удобрений, расходуемых на 1 га, т/га;

x_3 – количество зерноуборочных комбайнов на 100 га;

x_4 – количество тракторов приведенной мощности на 100 га.

Вариант 20

Номер предприятия	y	x_1	x_2	x_3
1	243	100	118	7
2	411	88	125	10
3	310	116	244	10
4	356	124	122	5
5	460	90	81	6
6	351	115	165	10
7	428	107	97	10
8	270	93	101	16
9	750	175	150	10
10	382	96	90	8
11	235	93	96	10
12	650	176	333	21
13	231	75	65	15
14	444	116	90	11
15	340	89	93	3
16	232	74	93	10
17	265	75	81	12
18	370	111	87	5
19	303	90	92	15
20	423	100	110	7

Для варианта 20 в таблице исходных данных приняты обозначения:

y – число рабочих мест;

x_1 – руководство, чел.;

x_2 – труд высококвалифицированных специалистов, чел.;

x_3 – количество вакантных мест работы.

Порядок выполнения задания

1. Постройте линейную модель множественной регрессии.
2. Запишите стандартизованное уравнение множественной регрессии. На основе стандартизованных коэффициентов регрессии и средних коэффициентов эластичности ранжируйте факторы по степени их влияния на результат.
3. Найдите коэффициенты парной, частной и множественной корреляции. Проанализируйте их.
4. Найдите скорректированный коэффициент множественной детерминации. Сравните его с нескорректированным (общим) коэффициентом детерминации.
5. С помощью F-критерия Фишера оцените статистическую надежность уравнения регрессии и коэффициента детерминации $R^2_{yx_1x_2}$.
6. С помощью частных F-критериев Фишера оцените целесообразность включения в уравнение множественной регрессии фактора x_1 после x_2 и фактора x_2 после x_1.
7. Составьте уравнение линейной парной регрессии, оставив лишь один значащий фактор.

Критерии оценки:

– «зачтено» – выполнено не менее 70 % всех заданий;

– «не зачтено» – выполнено менее 70 % всех заданий.

Примеры решения задач

Исходные данные приведены в таблице.

Страна	y	x_1 – средняя ожидаемая продолжительность жизни при рождении, число лет	x_2 – ВВП в паритетах покупательной способности	x_3 – темпы прироста рабочей силы по сравнению с предыдущим годом, %	x_4 – коэффициент младенческой смертности, %
Мозамбик	47	3,0	2,6	2,4	113
Бурунди	49	2,3	2,6	2,7	98
Чад	48	2,6	2,5	2,5	117
Непал	55	4,3	2,5	2,4	91
Буркина- Фасо	49	2,9	2,8	2,1	99
Мадагаскар	52	2,4	3,1	3,1	89
Бангладеш	58	5,1	1,6	2,1	79

Страна	y	x_1 – средняя ожидаемая продолжительность жизни при рождении, число лет	x_2 – ВВП в паритетах покупательной способности	x_3 – темпы прироста рабочей силы по сравнению с предыдущим годом, %	x_4 – коэффициент младенческой смертности, %
Гаити	57	3,4	2,0	1,7	72
Мали	50	2,0	2,9	2,7	123
Нигерия	53	4,5	2,9	2,8	80
Кения	58	5,1	2,7	2,7	58
Того	56	4,2	3,0	2,8	88
Индия	62	5,2	1,8	2,0	68
Бенин	50	6,5	2,9	2,5	95
Никарагуа	68	7,4	3,1	4,0	46
Гана	59	7,4	2,8	2,7	73
Ангола	47	4,9	3,1	2,8	124
Пакистан	60	8,3	2,9	3,3	90
Мавритания	51	5,7	2,5	2,7	96
Зимбабве	57	7,5	2,4	2,2	55

Решение

1. Построим линейную модель множественной регрессии с помощью надстройки MS Excel «Данные / Анализ данных / Регрессия» (рис. 2.1).

	A	B	C	D	E	F	G
27	*Регрессионная статистика*						
28	Множественный R	0,919059					
29	R-квадрат	0,844669					
30	Нормированный R-квадрат	0,803248					
31	Стандартная ошибка	2,501345					
32	Наблюдения	20					
33							
34	Дисперсионный анализ						
35		*df*	*SS*	*MS*	*F*	*Значимость F*	
36	Регрессия	4	510,3491	127,5873	20,39201	6,31E-06	
37	Остаток	15	93,85093	6,256729			
38	Итого	19	604,2				
39							
40		*Коэффициен*	*Стандартна*	*t-статисти*	*P-Значение*	*Нижние 95%*	*Верхние 95%*
41	Y-пересечение	66,81415	4,790036	13,94857	5,39E-10	56,60443	77,02387
42	Переменная X 1	0,441497	0,380772	1,159478	0,264386	-0,3701	1,253094
43	Переменная X 2	-5,41033	2,224273	-2,43241	0,027992	-10,1513	-0,66941
44	Переменная X 3	5,25415	1,881468	2,79258	0,013663	1,243895	9,264404
45	Переменная X 4	-0,16034	0,034491	-4,64871	0,000315	-0,23386	-0,08682

Рис. 2.1. Итоги множественной линейной регрессии

Получили следующее уравнение линейной регрессии:

$$\hat{y} = 66{,}814 + 0{,}441x_1 - 5{,}41x_2 + 5{,}254x_3 - 0{,}16x_4.$$

2. Найдем стандартизованное уравнение множественной регрессии:

$$t_y = \beta_1 t_{x_1} + \beta_2 t_{x_2} + \beta_3 t_{x_3} + \beta_4 t_{x_4} + \varepsilon.$$

Найдем сначала средние квадратические отклонения признаков с помощью функции СТАНДОТКЛОН:

Показатель	y	x_1	x_2	x_3	x_4
σ	5,496	1,887	0,413	0,490	21,345

Коэффициенты β_i стандартизованного уравнения регрессии находим по формуле

$$\beta_i = b_i \cdot \frac{\sigma_{x_i}}{\sigma_y}.$$

$$\beta_1 = 0{,}441 \cdot \frac{1{,}887}{5{,}496} = 0{,}152;\ \beta_2 = -5{,}41 \cdot \frac{0{,}413}{5{,}496} = -0{,}406;$$

$$\beta_3 = 5{,}254 \cdot \frac{0{,}49}{5{,}496} = 0{,}468;\ \beta_4 = -0{,}16 \cdot \frac{21{,}345}{5{,}496} = -0{,}623.$$

То есть уравнение регрессии в стандартизованном масштабе будет выглядеть следующим образом:

$$t_y = 0{,}152z_{x_1} - 0{,}406z_{x_2} + 0{,}468z_{x_3} - 0{,}623z_{x_4}.$$

Так как стандартизованные коэффициенты регрессии можно сравнивать между собой, то по степени влияния на результат факторы можно расположить следующим образом: x_4, x_3, x_2, x_1.

3. Найдем коэффициенты парной корреляции с помощью надстройки MS Excel «Данные / Анализ данных / корреляция»:

	y	x_1	x_2	x_3	x_4
y	1				
x_1	0,639	1			
x_2	–0,207	0,040	1		
x_3	0,283	0,316	0,728	1	
x_4	–0,853	–0,570	0,238	–0,101	1

У фактора x_4 сильная обратная зависимость с результатом, фактор x_1 имеет среднюю прямую зависимость с y, фактор x_2 имеет слабую обратную, а фактор x_3 – слабую прямую зависимость с результатом. Также между факторами x_2 и x_3 существует колинеарность ($r > 0{,}7$).

Коэффициент множественной корреляции найден на рис. 2.1 (ячейка B28): R = 0,919, находится в диапазоне от 0,9 до 1, связь с факторами очень сильная.

Частный коэффициент корреляции характеризует тесноту линейной зависимости между результатом и соответствующим фактором при устранении влияния других факторов. Он вычисляется по формуле

$$r_{jk.1,2,\ldots,m} = -\frac{C_{jk}}{\sqrt{C_{jj} \cdot C_{kk}}},$$

где C_{jk}, C_{jj}, C_{kk} – алгебраические дополнения к соответствующим элементам матрицы парных корреляций, то есть фактически это элементы матрицы, обратной к матрице парных корреляций.

Вычисляем обратную матрицу парных корреляций с помощью функции МОБР:

	A	B	C	D	E	F	G
59		i/j	0	1	2	3	4
60		0	6,4379	-0,9759	2,6150	-3,0143	4,0087
61	r^* =	1	-0,9759	1,7984	-0,3019	-0,0462	0,2606
62		2	2,6150	-0,3019	3,7550	-3,2948	0,8332
63		3	-3,0143	-0,0462	-3,2948	4,1259	-1,3981
64		4	4,0087	0,2606	0,8332	-1,3981	4,228674

Построим матрицу частных коэффициентов корреляции:

	A	B	C	D	E	F	G
66		i/j	0	1	2	3	4
67		0	1				
68	ρ =	1	0,2868	1			
69		2	-0,5319	0,1162	1		
70		3	0,5849	0,0170	0,8371	1	
71		4	-0,7683	-0,0945	-0,2091	0,3347	1

Сравним парные и частные коэффициенты корреляции:

	r (парные)	ρ (частные)
$r_{y,x1()}$ =	0,639	0,287
$r_{y,x2()}$ =	–0,207	–0,532
$r_{y,x3()}$ =	0,283	0,585
$r_{y,x4()}$ =	0,853	–0768

Выводы:

– частный коэффициент корреляции между y и x_1 отличается от парного и не подтверждает среднюю прямую связь между этими показателями;
– частный коэффициент корреляции между y и x_2 отличается от парного и не подтверждает слабую обратную связь между этими показателями;
– частный коэффициент корреляции между y и x_3 отличается от парного и не подтверждает слабую обратную связь между этими показателями;
– частный коэффициент корреляции между y и x_4 хоть немного «заминает» результат, тем не менее подтверждает сильную обратную связь между этими показателями.

4. Найдем скорректированный коэффициент множественной детерминации (ячейка B30 на рис. 2.1) $R^2_{\text{скор}} = 0{,}844$. Нескорректированный коэффициент детерминации (ячейка B29 на рис. 2.1) $R^2 = 0{,}803$.

5. С помощью F-критерия Фишера оценим статистическую надежность уравнения регрессии и коэффициента детерминации R^2:

$$F_{\text{факт}} = \frac{R^2}{1 - R^2} \cdot \frac{n - m - 1}{m} = \frac{0{,}844}{1 - 0{,}844} \cdot \frac{20 - 3 - 1}{3} = 20{,}392.$$

Получили, что $F_{\text{факт}} > F_{\text{табл}} = 3{,}59$ (при $n = 20$ и уровне значимости $\alpha = 0{,}05$), а значит, подтверждается статистическая значимость всего уравнения и показателя тесноты связи.

6. С помощью частных F-критериев Фишера оценим целесообразность включения в уравнение множественной регрессии фактора x_1 после x_2 и фактора x_2 после x_1:

$$F_{\text{частн.}x_1} = \frac{R^2_{yx_1x_2} - R^2_{yx_2}}{1 - R^2_{yx_1x_2}} \cdot \frac{n - m - 1}{1}; \; F_{\text{частн.}x_2} = \frac{R^2_{yx_1x_2} - R^2_{yx_1}}{1 - R^2_{yx_1x_2}} \cdot \frac{n - m - 1}{1}.$$

Найдем $R^2_{yx_1x_2}$, $R^2_{yx_1}$ и $R^2_{yx_2}$.

$R^2_{yx_1x_2}$ найдем с помощью инструмента «Анализ данных / Регрессия»:

	A	B
75	*Регрессионная статистика*	
76	Множественный R	0,679798
77	R-квадрат	0,462125
78	Нормированный R-квадр	0,398846
79	Стандартная ошибка	4,372266
80	Наблюдения	20

$$R^2_{yx_1} = r^2_{yx_1} = 0{,}639^2 = 0{,}408;\ R^2_{yx_2} = r^2_{yx_2} = 0{,}207^2 = 0{,}043.$$

Тогда

$$F_{\text{частн.}x_1} = \frac{0{,}462 - 0{,}043}{1 - 0{,}462} \cdot \frac{20 - 2 - 1}{1} = 13{,}246;$$

$$F_{\text{частн.}x_2} = \frac{0{,}462 - 0{,}408}{1 - 0{,}462} \cdot \frac{20 - 2 - 1}{1} = 1{,}714.$$

$F_{\text{част.}x_1} > F_{\text{табл}} = 3{,}59$, то есть вероятность его случайного формирования меньше принятого стандарта $\alpha = 0{,}05$ (5 %). Следовательно, значение частного F-критерия для дополнительно включенного фактора x_1 неслучайно, является статистически значимым, надежным, достоверным: прирост факторной дисперсии за счет дополнительного фактора x_1 является существенным. Фактор x_1 должен присутствовать в уравнении, в том числе в варианте, когда он дополнительно включается после фактора x_2.

$F_{\text{частн.}x_2} < F_{\text{табл}} = 3{,}59$. Следовательно, включение в модель фактора x_2 после того, как в модель включен фактор x_1, статистически нецелесообразно: прирост факторной дисперсии за счет дополнительного признака x_2 оказывается незначительным; фактор x_2 включать в уравнение после фактора x_1 не следует.

7. Исключим из модели факторы x_1, x_2 и x_3 как менее информативные. Составим уравнение линейной парной регрессии с фактором x_4 (рис. 2.2).

Получили следующее уравнение линейной регрессии:

$$\hat{y} = 73{,}56 - 0{,}22\,x_4.$$

	A	B	C	D	E	F	G
96	*Регрессионная статистика*						
97	Множественный R	0,852881					
98	R-квадрат	0,727406					
99	Нормированный R-квадр	0,712262					
100	Стандартная ошибка	3,024907					
101	Наблюдения	20					
102							
103	Дисперсионный анализ						
104		*df*	*SS*	*MS*	*F*	*ачимость F*	
105	Регрессия	1	439,4989	439,4989	48,03233	1,78E-06	
106	Остаток	18	164,7011	9,150064			
107	Итого	19	604,2				
108							
109		*Коэффициен*	*дартная о*	*статисти*	*P-Значение*	*ижние 95%*	*ерхние 95%*
110	Y-пересечение	73,56047	2,860202	25,71863	1,21E-15	67,55141	79,56953
111	x4	-0,21962	0,031688	-6,93054	1,78E-06	-0,28619	-0,15304

Рис. 2.2. Итоги парной линейной регрессии

Рекомендуемая литература

1. Ивченко Ю.С. Эконометрика : курс лекций. – Саратов : Вузовское образование, 2018. – 121 с.
2. Кремер Н.Ш., Путко Б.А. Эконометрика : учебник / под ред. Н.Ш. Кремера. – 3-е изд., перераб. и доп. – М. : ЮНИТИ-ДАНА, 2017. – 328 с.
3. Кулинич Е.И. Эконометрия. – М. : Финансы и статистика, 2001. – 304 с.

Тема 2.2. Фиктивные переменные

Форма проведения занятия: семинар.

Вопросы для обсуждения

1. Необходимость использования фиктивных переменных.
2. Модели, содержащие только качественные объясняющие переменные.
3. Модели, в которых объясняющие переменные носят как количественный, так и качественный характер.

Методические указания по проведению занятия

Семинар начинается с уточнения базовых понятий. Все студенты готовят сообщение по всем вопросам семинара. Каждый студент готовит реферат по отдельным вопросам семинара. Один студент из группы готовит доклад творческого характера, содержащий элементы исследовательского характера.

1. Необходимость использования фиктивных переменных

Обычно в моделях влияние качественного фактора выражается в виде фиктивной или искусственной переменной, которая отражает два противоположных состояния качественного фактора.

Например, «фактор действует» – «фактор не действует», «курс валюты фиксированный» – «курс валюты плавающий», «сезон летний» – «сезон зимний» и так далее.

В этом случае фиктивная переменная может выражаться в двоичной форме.

Тогда значение переменной d будет равно нулю, если фактор не действует, или единице, если фактор действует.

Возвращаясь к рассмотренным выше примерам, можно определить следующую фиктивную переменную, отражающую состояние на валютном рынке.

Значение переменной d будет равно нулю, если курс валюты фиксированный, или единице, если курс валюты плавающий.

Возможным вариантом качественной переменной, учитывающей в регрессионной модели фактор сезонности, является такая переменная: d равна нулю, если сезон летний. А если сезон не летний, тогда значение d равно единице.

В зависимости от целей исследования используются различные фиктивные переменные.

Например:

$$d = \begin{cases} 0, & \textit{если в обществе имеются инфляционные ожидания}, \\ 1, & \textit{если инфляционных ожиданий нет}. \end{cases}$$

Приведем примеры фиктивных переменных, которые могут быть использованы при составлении регрессионных моделей, описывающих процессы в сфере образования.

$$d = \begin{cases} 0, & \textit{если экзамен не сдан с первой попытки}, \\ 1, & \textit{если экзамен сдан с первой попытки}. \end{cases}$$

Или

$$d = \begin{cases} 0, & \textit{если в процессе обучения не используются компьютеры}, \\ 1, & \textit{в противоположном случае}. \end{cases}$$

Итак, мы сформулировали определение и рассмотрели примеры качественных переменных.

Отметим, что используемая для оценки качественного признака переменная d имеет несколько названий. Это *фиктивная, искусственная, двоичная переменная, или индикатор*.

Таким образом, кроме моделей, содержащих только количественные объясняющие переменные, обозначаемые x_i, в регрессионном анализе рассматриваются также модели, содержащие лишь качественные переменные, обозначаемые d_i.

Встречаются модели, содержащие и те и другие переменные одновременно.

Искусственные переменные могут входить в состав регрессионных моделей в качестве как объясняющих, так и объясняемых переменных.

Способы оценки параметров и качества моделей, содержащих фиктивные переменные, аналогичны рассмотренным ранее методам для моделей, содержащих только количественные переменные.

2. Модели, содержащие только качественные объясняющие переменные

Регрессионные модели, содержащие лишь качественные объясняющие переменные, называются *моделями дисперсионного анализа* (*ANOVA-моделями*).

Например, пусть y – начальная заработная плата. А искусственная переменная d показывает наличие или отсутствие у претендента соответствующего образования, то есть переменная равна нулю, если претендент не имеет высшего образования, или единице, если претендент имеет высшее образование. Тогда

$$d = \begin{cases} 0, & \textit{если претендент не имеет высшего образования}, \\ 1, & \textit{если претендент имеет высшее образование}. \end{cases}$$

Зависимость начальной заработной платы от высшего образования можно выразить моделью парной регрессии

$$y = \alpha + \beta d + \varepsilon. \qquad (2.1)$$

Очевидно, справедливо соотношение

$$M_y(d = 0) = \alpha + \beta \cdot 0 = \alpha.$$

Значение заработной платы при $d = 0$, то есть при отсутствии у претендента высшего образования, равно α, то есть α определяет среднюю начальную заработную плату при отсутствии высшего образования.

Как показывает выражение $M_y(d = 1) = \alpha + \beta \cdot 1 = \alpha + \beta$, при $d = 1$, то есть при наличии образования, заработная плата будет равна $\alpha + \beta$.

Коэффициент β указывает, на какую величину отличаются средние начальные заработные платы при наличии или отсутствии высшего образования.

Можно проверить статистическую значимость коэффициента β с помощью t-статистики либо значимость коэффициента детерминации R^2 с помощью F-статистики. Тогда определяем, влияет наличие высшего образования на начальную заработную плату или нет.

ANOVA-модели представляют собой кусочно-постоянные функции. Но подобного рода модели на практике встречаются редко. Намного чаще используются модели, включающие и качественные, и количественные переменные.

3. Модели, в которых объясняющие переменные носят как количественный, так и качественный характер

Модели, в которых объясняющие переменные носят как количественный, так и качественный характер, *называются моделями ковариационного анализа (ANCOVA-моделями).*

Рассмотрим простейшую ANCOVA-модель с одной количественной и одной качественной переменной, имеющей два альтернативных состояния:

$$y = \alpha + \beta_1 x + \beta_2 d + \varepsilon. \qquad (2.2)$$

Пусть, например, y – заработная плата сотрудника фирмы, x – стаж сотрудника, d – пол сотрудника, то есть

$$d = \begin{cases} 0, & \text{если сотрудник — женщина,} \\ 1, & \text{если сотрудник — мужчина.} \end{cases}$$

Тогда ожидаемое значение заработной платы сотрудников при x годах трудового стажа будет:

– для женщины

$$M_y(x, d = 0) = \alpha + \beta_1 x + \beta_2 \cdot 0 = \alpha + \beta_1 x; \qquad (2.3)$$

– для мужчины

$$M_y(x, d = 1) = \alpha + \beta_1 x + \beta_2 \cdot 1 = (\alpha + \beta_2) + \beta_1 x. \qquad (2.4)$$

Заработная плата в данном случае является линейной функцией от стажа работы. Причем и для мужчин, и для женщин заработная плата меняется с одним и тем же коэффициентом пропорциональности β_1. А вот свободные члены в моделях (2.3), (2.4) отличаются на величину β_2. Проверив с помощью t-статистики статистические значимости коэффициентов α и $(\alpha + \beta_2)$, можно определить, имеет ли место в фирме дискриминация по половому признаку. Если эти коэффициенты окажутся статистически значимыми, то, очевидно, дискриминация есть. Более того, при $\beta_2 > 0$ она будет в пользу мужчин, при $\beta_2 < 0$ – в пользу женщин.

В данном случае пол сотрудников имеет два альтернативных значения, и в модели это отражается одной фиктивной переменной. Возникает вопрос, нельзя ли с помощью большего числа фиктивных переменных обрисовать более сложные комбинации. Например, пусть

$$y = \alpha + \beta_1 x + \beta_2 d_1 + \beta_3 d_2 + \varepsilon. \qquad (2.5)$$

$$d_1 = \begin{cases} 0, & \text{если сотрудник — мужчина,} \\ 1, & \text{если сотрудник — женщина.} \end{cases}$$

$$d_2 = \begin{cases} 0, & \text{если сотрудник — женщина,} \\ 1, & \text{если сотрудник — мужчина.} \end{cases}$$

Но в этой ситуации между переменными d_1 и d_2 существует строгая линейная зависимость: $d_2 = 1 - d_1$. Мы попадаем в ситуацию совершенной мультиколлинеарности, при которой коэффициенты β_2 и β_3 однозначно определены быть не могут. Простейшим способом преодоления данной проблемы является отбрасывание одной из фиктивных переменных и использование для рассматриваемой задачи модели (2.2). Применяя аналогичные выкладки, можно получить следующее общее правило.

Если качественная переменная имеет k альтернативных значений, то при моделировании используются только (k – 1) *фиктивных переменных.*

Если не следовать данному правилу, то при моделировании исследователь попадает в ситуацию совершенной мультиколлинеарности (так называемую *ловушку фиктивной переменной*).

Значения фиктивной переменной можно изменять на противоположные. Суть модели от этого не изменится. Например, в модели (2.2) можно положить, что:

$$d = \begin{cases} 0, & \text{если сотрудник — мужчина,} \\ 1, & \text{если сотрудник — женщина.} \end{cases}$$

Однако при этом знак коэффициента β_1 изменится на противоположный.

Значение качественной переменной, для которого принимается $d = 0$, называется *базовым* или *сравнительным*. Выбор базового значения обычно диктуется целями исследования, но может быть и произвольным.

Коэффициент β_1 в модели (2.2) иногда называется *дифференциальным коэффициентом свободного члена*, так как он показывает, на какую величину отличается свободный член модели при значении фиктивной переменной, равном единице, от свободного члена модели при базовом значении фиктивной переменной.

Пусть рассматривается модель с двумя объясняющими переменными, одна из которых количественная, а другая — качественная. Причем качественная переменная имеет три альтернативы. Например, ситуация, связанная с расходами на содержание ребенка, может быть связана с доходами домохозяйств и возрастом ребенка: дошкольный, младший школьный и старший школьный. Так как качественная переменная связана с тремя альтернативами, то по общему правилу моделирования необходимо использовать две качественные переменные. Таким образом, модель может быть представлена в виде

$$y = \alpha + \beta_1 x + \beta_2 d_1 + \beta_3 d_2 + \varepsilon, \tag{2.6}$$

где y — расходы; x — доходы домохозяйств.

$$d_1 = \begin{cases} 0, & \text{если в семье ребенок дошкольного возраста,} \\ 1, & \text{если в семье ребенок другого возраста;} \end{cases}$$

$$d_2 = \begin{cases} 0, & \text{если в семье младший школьник,} \\ 1, & \text{в противоположном случае.} \end{cases}$$

Таким образом, получаются следующие зависимости.

Средний расход на дошкольника:

$$M_y(x, d_1 = 0, d_2 = 0) = \alpha + \beta_1 x. \quad (2.7)$$

Средний расход на младшего школьника:

$$M_y(x, d_1 = 1, d_2 = 0) = \alpha + \beta_1 x + \beta_2 \cdot 1 + \beta_3 \cdot 0 = (\alpha + \beta_2) + \beta_1 x. \quad (2.8)$$

Средний расход на старшего школьника:

$$M_y(x, d_1 = 1, d_2 = 1) = \alpha + \beta_1 x + \beta_2 \cdot 1 + \beta_3 \cdot 1 = (\alpha + \beta_2 + \beta_3) + \beta_1 x. \quad (2.9)$$

Здесь β_2, β_3 – дифференциальные свободные члены. Базовым значением качественной переменной является значение «дошкольник». После определения коэффициентов регрессии (2.6) определяется статистическая значимость коэффициентов β_2, β_3 на основе t-статистики. Если коэффициенты оказываются статистически незначимыми, то можно сделать вывод, что возраст ребенка не оказывает существенного влияния на его содержание.

Рекомендуемая литература

1. Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика. Начальный курс : учебник. – М. : Дело, 2004. – 576 с.
2. Носко В.П. Эконометрика. Элементарные методы и введение в регрессионный анализ временных рядов. – М. : Институт экономики переходного периода, 2004. – 501 с.
3. Тихомиров Н.П., Дорохина Е.Ю. Эконометрика : учебник. – М. : Экзамен, 2003. – 512 с.

Модуль 3. СРАВНЕНИЕ РАЗЛИЧНЫХ ЭКОНОМЕТРИЧЕСКИХ МОДЕЛЕЙ. МОДЕЛИ ИНТЕГРИРОВАННОГО ТИПА

Тема 3.1. Временные ряды

Форма проведения занятия: практическое занятие.

Вопросы для обсуждения

1. Теоретические основы постановки, решения и анализа задач моделирования временных рядов.
2. Методологические принципы постановки, решения и анализа задач моделирования временных рядов.
3. Конкретные подходы постановки, решения и анализа задач моделирования временных рядов.

Методические указания по проведению занятия

1. Построить автокорреляционную функцию и сделать вывод о наличии сезонных колебаний.

2. Построить аддитивную модель временного ряда (для нечетных вариантов) или мультипликативную модель временного ряда (для четных вариантов).

3. Сделать прогноз на 2 квартала вперед.

Номер варианта проверяемого задания находится по таблице по первой букве отчества студента.

Буква	А	Б	В	Г	Д	Е	Ж
№ вар.	1	2	3	4	5	6	7
Буква	З	И	К	Л	М	Н	О
№ вар.	8	9	10	11	12	13	14
Буква	П	Р	С	Т	У	Ф	Х
№ вар.	15	16	17	18	19	20	21
Буква	Ц, Э	Ч, Ю	Ш, Я	Щ			
№ вар.	22	23	24	25			

Расчеты производить в меню сервиса «Анализ данных» табличного процессора Excel.

Отчет по проделанной работе оформить в Word, при вводе формул использовать Equation.

Исследования модели произвести по порядку, в соответствии с перечисленными в заданиях пунктами.

Варианты заданий

Вариант 1

t	y_t	t	y_t
1	5,8	9	7,9
2	4,5	10	5,5
3	5,1	11	6,3
4	9,1	12	10,8
5	7,0	13	9,0
6	5,0	14	6,5
7	6,0	15	7,0
8	10,1	16	11,1

Вариант 2

t	y_t	t	y_t
1	5,5	9	8,0
2	4,6	10	5,6
3	5,0	11	6,4
4	9,2	12	10,9
5	7,1	13	9,1
6	5,1	14	6,4
7	5,9	15	7,2
8	10,0	16	11,0

Вариант 3

t	y_t	t	y_t
1	5,3	9	8,2
2	4,7	10	5,5
3	5,2	11	6,5
4	9,1	12	11,0
5	7,0	13	8,9
6	5,0	14	6,5
7	6,0	15	7,3
8	10,1	16	11,2

Вариант 4

t	y_t	t	y_t
1	5,5	9	8,3
2	4,8	10	5,4
3	5,1	11	6,4
4	9,0	12	10,9
5	7,1	13	9,0
6	4,9	14	6,6
7	6,1	15	7,5
8	10,0	16	11,2

Вариант 5

t	y_t	t	y_t
1	5,6	9	8,2
2	4,7	10	5,6
3	5,2	11	6,4
4	9,1	12	10,8
5	7,0	13	9,1
6	5,1	14	6,7
7	6,0	15	7,5
8	10,2	16	11,3

Вариант 6

t	y_t	t	y_t
1	5,8	9	7,8
2	4,6	10	5,5
3	5,1	11	6,3
4	9,1	12	10,8
5	7,0	13	9,0
6	5,0	14	6,5
7	6,0	15	7,0
8	10,1	16	11,1

Вариант 7

t	y_t	t	y_t
1	5,4	9	8,0
2	4,6	10	5,6
3	5,0	11	6,4
4	9,2	12	10,9
5	7,1	13	9,1
6	5,1	14	6,4
7	5,9	15	7,2
8	10,0	16	11,0

Вариант 8

t	y_t	t	y_t
1	5,3	9	8,2
2	4,8	10	5,5
3	5,2	11	6,5
4	9,1	12	11,0
5	7,0	13	8,9
6	5,0	14	6,5
7	6,0	15	7,3
8	10,1	16	11,2

Вариант 9

t	y_t	t	y_t
1	5,5	9	8,3
2	4,8	10	5,4
3	5,2	11	6,4
4	9,0	12	10,9
5	7,1	13	9,0
6	4,9	14	6,6
7	6,1	15	7,5
8	10,0	16	11,2

Вариант 10

t	y_t	t	y_t
1	5,6	9	8,2
2	4,7	10	5,6
3	5,2	11	6,4
4	9,2	12	10,8
5	7,0	13	9,1
6	5,1	14	6,7
7	6,0	15	7,5
8	10,2	16	11,3

Вариант 11

t	y_t	t	y_t
1	5,8	9	7,9
2	4,5	10	5,5
3	5,1	11	6,3
4	9,1	12	10,8
5	7,1	13	9,0
6	5,0	14	6,5
7	6,0	15	7,0
8	10,1	16	11,1

Вариант 12

t	y_t	t	y_t
1	5,5	9	8,0
2	4,6	10	5,6
3	5,0	11	6,4
4	9,2	12	10,9
5	7,1	13	9,1
6	5,2	14	6,4
7	5,9	15	7,2
8	10,0	16	11,0

Вариант 13

t	y_t	t	y_t
1	5,3	9	8,2
2	4,7	10	5,5
3	5,2	11	6,5
4	9,1	12	11,0
5	7,0	13	8,9
6	5,0	14	6,5
7	6,1	15	7,3
8	10,1	16	11,2

Вариант 14

t	y_t	t	y_t
1	5,5	9	8,3
2	4,8	10	5,4
3	5,1	11	6,4
4	9,0	12	10,9
5	7,1	13	9,0
6	4,9	14	6,6
7	6,2	15	7,5
8	10,0	16	11,2

Вариант 15

t	y_t	t	y_t
1	5,6	9	8,2
2	4,7	10	5,6
3	5,2	11	6,4
4	9,1	12	10,8
5	7,0	13	9,1
6	5,1	14	6,7
7	6,0	15	7,5
8	10,1	16	11,3

Вариант 16

t	y_t	t	y_t
1	5,8	9	7,9
2	4,5	10	5,5
3	5,1	11	6,3
4	9,1	12	10,8
5	7,0	13	9,0
6	5,0	14	6,5
7	6,0	15	7,0
8	10,1	16	11,1

Вариант 17

t	y_t	t	y_t
1	5,5	9	8,1
2	4,6	10	5,6
3	5,0	11	6,4
4	9,2	12	10,9
5	7,1	13	9,1
6	5,1	14	6,4
7	5,9	15	7,2
8	10,0	16	11,0

Вариант 18

t	y_t	t	y_t
1	5,3	9	8,2
2	4,7	10	5,4
3	5,2	11	6,5
4	9,1	12	11,0
5	7,0	13	8,9
6	5,0	14	6,5
7	6,0	15	7,3
8	10,1	16	11,2

Вариант 19

t	y_t	t	y_t
1	5,5	9	8,3
2	4,8	10	5,4
3	5,1	11	6,3
4	9,0	12	10,9
5	7,1	13	9,0
6	4,9	14	6,6
7	6,1	15	7,5
8	10,0	16	11,2

Вариант 20

t	y_t	t	y_t
1	5,6	9	8,2
2	4,7	10	5,5
3	5,2	11	6,4
4	9,1	12	10,8
5	7,0	13	9,1
6	5,1	14	6,7
7	6,0	15	7,5
8	10,2	16	11,3

Вариант 21

t	y_t	t	y_t
1	5,8	9	7,9
2	4,5	10	5,5
3	5,1	11	6,3
4	9,1	12	10,8
5	7,0	13	9,0
6	5,0	14	6,4
7	6,0	15	7,0
8	10,1	16	11,1

Вариант 22

t	y_t	t	y_t
1	5,5	9	8,0
2	4,6	10	5,6
3	5,0	11	6,4
4	9,2	12	10,9
5	7,1	13	9,1
6	5,1	14	6,4
7	5,9	15	7,1
8	10,0	16	11,0

Вариант 23

t	y_t	t	y_t
1	5,3	9	8,2
2	4,7	10	5,5
3	5,2	11	6,5
4	9,1	12	11,0
5	7,0	13	8,8
6	5,0	14	6,5
7	6,0	15	7,3
8	10,1	16	11,2

Вариант 24

t	y_t	t	y_t
1	5,5	9	8,3
2	4,8	10	5,4
3	5,1	11	6,4
4	9,0	12	10,9
5	7,1	13	9,0
6	4,9	14	6,6
7	6,1	15	7,4
8	10,0	16	11,2

Вариант 25

t	y_t	t	y_t
1	5,6	9	8,2
2	4,7	10	5,5
3	5,2	11	6,4
4	9,1	12	10,8
5	7,0	13	9,1
6	5,1	14	6,7
7	6,0	15	7,5
8	10,2	16	11,3

Примеры решения задач

Имеются условные данные об объемах потребления электроэнергии (y_t) жителями региона за 16 кварталов. Проведем эконометрическое исследование предложенных данных.

t	y_t	t	y_t
1	5,8	9	7,9
2	4,5	10	5,5
3	5,1	11	6,3
4	9,1	12	10,8
5	7	13	9
6	5	14	6,5
7	6	15	7
8	10,1	16	11,1

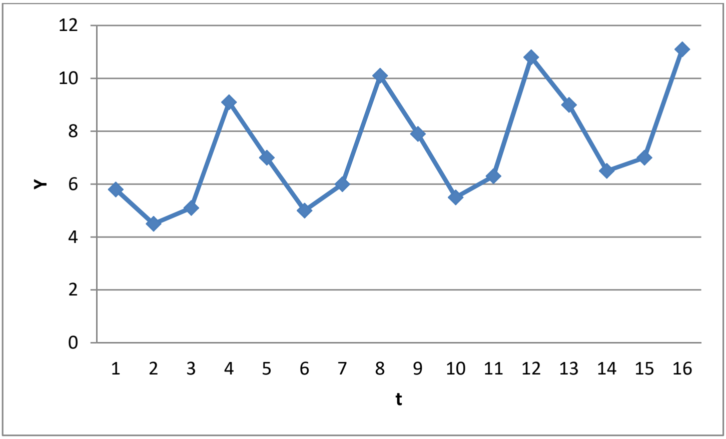
1. Коэффициент автокорреляции L-го порядка рассчитывается по формуле

$$r_{t,t-L} = \frac{\overline{y_t y_{t-L}} - \bar{y}_t \cdot \bar{y}_{t-L}}{\sigma_t \cdot \sigma_{t-L}}.$$

Рассчитаем несколько последовательных коэффициентов автокорреляции и представим их в виде автокорреляционной функции:

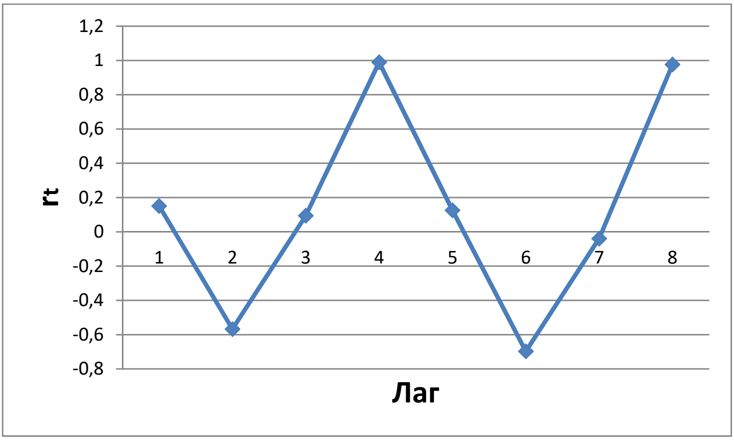
Лаг	Коэффициент автокорреляции уровней (автокорреляционная функция)
1	0,150809328
2	–0,567568016
3	0,094191144
4	0,989314661
5	0,125471542
6	–0,697323727
7	–0,039647984
8	0,976108227

2. Построим график исходных данных:



Значения *Y* образуют пилообразную фигуру, содержащую восходящую тенденцию.

Ниже приведем график автокорреляционной функции:



Анализ коррелограммы и графика исходных уровней временного ряда позволяет сделать вывод о наличии в изучаемом временном ряде сезонных колебаний периодичностью в четыре квартала, поскольку величина изучаемого признака в первый – второй кварталы ниже, чем в третий – четвертый.

3. Мультипликативная модель имеет вид $Y = T \cdot S \cdot E$.

Произведем выравнивание исходных уровней ряда методом скользящей средней:

Номер квартала, t	Объём потребления электроэнергии, y_t	Итого за четыре квартала	Скользящая средняя за четыре квартала	Центрированная скользящая средняя	Оценка сезонной компоненты
1	5,8	–	–	–	–
2	4,5	24,5	6,125	–	–
3	5,1	25,7	6,425	6,275	0,813
4	9,1	26,2	6,550	6,488	1,403
5	7	27,1	6,775	6,663	1,051
6	5	28,1	7,025	6,900	0,725
7	6	29,0	7,250	7,138	0,841
8	10,1	29,5	7,375	7,313	1,381
9	7,9	29,8	7,450	7,413	1,066
10	5,5	30,5	7,625	7,538	0,730
11	6,3	31,6	7,900	7,763	0,812
12	10,8	32,6	8,150	8,025	1,346
13	9	33,3	8,325	8,238	1,093
14	6,5	33,6	8,400	8,363	0,777
15	7	–	–	–	–
16	11,1	–	–	–	–

Найдем оценки сезонной компоненты как частное от деления фактических уровней ряда на центрированные скользящие средние. Для этого найдем средние за каждый квартал оценки сезонной компоненты S_i. Сезонные воздействия за период взаимопогашаются. В мультипликативной модели это выражается в том, что сумма значений сезонной компоненты по всем кварталам должна быть равна числу периодов в цикле. В нашем случае число периодов одного цикла равно 4.

Расчет средней сезонной компоненты S:

Показатели	Год	№ квартала, i			
		I	II	III	IV
	1	–	–	0,813	1,403
	2	1,051	0,725	0,841	1,381
	3	1,066	0,730	0,812	1,346
	4	1,093	0,777	–	–
Всего за i-й квартал		3,209	2,232	2,465	4,130
Средняя оценка сезонной компоненты для i-го квартала, S_i		1,070	0,744	0,822	1,377
Скорректированная сезонная компонента, S_i		1,067	0,742	0,819	1,373

Имеем

$$1{,}070 + 0{,}744 + 0{,}822 + 1{,}377 = 4{,}012.$$

Определяем корректирующий коэффициент:

$$k = 4/4{,}012 = 0{,}997.$$

Скорректированные значения сезонной компоненты S_i получаются при умножении ее средней оценки S_i на корректирующий коэффициент k.

Проверяем условие: равенство четырем суммы значений сезонной компоненты:

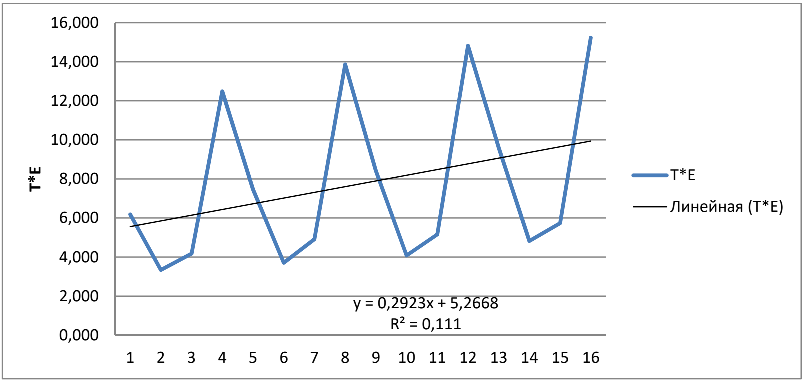
$$1{,}067 + 0{,}742 + 0{,}819 + 1{,}373 = 4.$$

Разделим каждый уровень исходного ряда на соответствующие значения сезонной компоненты. В результате получим величины $T \cdot E = YS$, которые содержат только тенденцию и случайную компоненту.

t	y_t	S_i	$T \cdot E$	$T_{расч}$	$T \cdot S$	E	$(y_t - T \cdot S)^2$	$(y_t - y)^2$
1	5,8	1,067	6,186	5,552	5,921	0,980	0,015	0,062
2	4,5	0,742	3,338	5,846	4,336	1,038	0,027	1,811
3	5,1	0,819	4,178	6,140	5,030	1,014	0,005	1,081
4	9,1	1,373	12,490	6,434	8,830	1,031	0,073	7,110
5	7	1,067	7,466	6,728	7,175	0,976	0,031	0,074

t	y_t	S_i	$T \cdot E$	$T_{расч}$	$T \cdot S$	E	$(y_t - T \cdot S)^2$	$(y_t - y)^2$
6	5	0,742	3,708	7,022	5,208	0,960	0,043	4,087
7	6	0,819	4,916	7,316	5,993	1,001	0,000	1,731
8	10,1	1,373	13,863	7,610	10,444	0,967	0,119	6,202
9	7,9	1,067	8,426	7,904	8,429	0,937	0,280	0,000
10	5,5	0,742	4,079	8,198	6,080	0,905	0,336	7,277
11	6,3	0,819	5,161	8,492	6,957	0,906	0,431	4,803
12	10,8	1,373	14,823	8,786	12,059	0,896	1,584	4,058
13	9	1,067	9,599	9,080	9,684	0,929	0,467	0,006
14	6,5	0,742	4,821	9,374	6,952	0,935	0,205	8,258
15	7	0,819	5,735	9,668	7,920	0,884	0,847	7,116
16	11,1	1,373	15,235	9,962	13,673	0,812	6,618	1,296
Сумма							11,081	54,971
Среднее	7,294							

Определим компоненту T в мультипликативной модели. Для этого рассчитаем параметры линейного тренда, используя уровни $T \cdot E$ (используем инструмент «Построение линии тренда» в Excel).

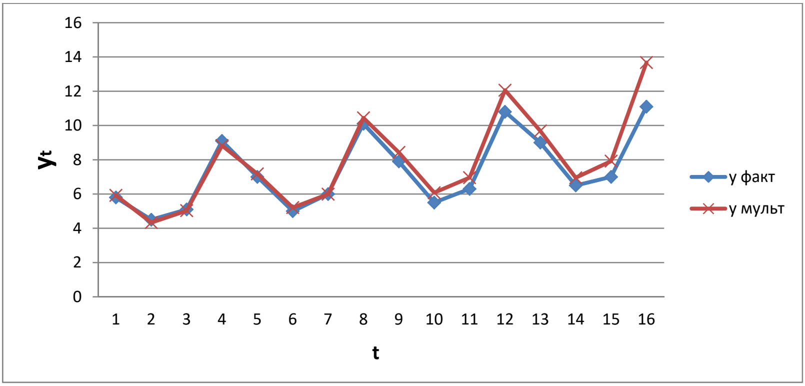


В результате получим уравнение тренда:

$$T = 5{,}2576 + 0{,}294 \cdot t.$$

Подставляя в это уравнение значения каждого момента времени $t = 1, 2, ..., 16$, найдем уровни T для каждого момента времени.

Найдем теоретические значения уровней ряда, умножив значения T на соответствующие значения сезонной компоненты. На одном графике откладываем фактические значения уровней временного ряда и теоретические, полученные по мультипликативной модели.



Расчет ошибки в мультипликативной модели производится по формуле

$$E = Y/(T \cdot S).$$

Для сравнения мультипликативной модели и других моделей временного ряда можно использовать сумму квадратов абсолютных ошибок $(y_t - T \cdot S)$:

$$R^2 = 1 - \frac{\sum(y_t - T \cdot S)^2}{\sum(y_t - y)^2} = 1 - \frac{11{,}081}{54{,}971} = 0{,}798.$$

Следовательно, мультипликативная модель объясняет 79,8 % общей вариации уровней временного ряда (потребления электроэнергии жителями региона).

4. Прогнозирование по мультипликативной модели. Прогнозное значение Ft – уровня временного ряда в мультипликативной модели – есть произведение трендовой и сезонной компонент. Для определения трендовой компоненты воспользуемся уравнением тренда:

$$T = 5{,}2576 + 0{,}294 \cdot t.$$

Получим

$$T_{17} = 5,2576 + 0,294 \cdot 17 = 10,256;$$

$$T_{18} = 5,2576 + 0,294 \cdot 18 = 10,550.$$

Значения сезонных компонент за соответствующие кварталы равны

$$S_1 = 1,067 \text{ и } S2 = 0,742.$$

Таким образом,

$$F_{17} = T_{17} \cdot S_1 = 10,256 \cdot 1,067 = 10,938;$$

$$F_{18} = T_{18} \cdot S_2 = 10,550 \cdot 0,742 = 7,825.$$

То есть в первые два квартала следующего года следует ожидать порядка 10,9 и 7,8 единицы объема потребления электроэнергии жителями региона.

Рекомендуемая литература

1. Практикум по эконометрике : учеб. пособие / под ред. И.И. Елисеевой. – М. : Финансы и статистика, 2003. – 192 с.
2. Тихомиров Н.П., Дорохина Е.Ю. Эконометрика : учебник. – М. : Экзамен, 2003. – 512 с.
3. Шалабанов А.К., Роганов Д.А. Эконометрика : учеб.-метод. пособие. – Казань : ТИСБИ, 2002. – 56 с.

Тема 3.2. Системы эконометрических уравнений

Форма проведения занятия: практическое занятие.

Вопросы для обсуждения

1. Теоретические основы постановки, решения и анализа задач моделирования систем эконометрических уравнений.
2. Методологические принципы постановки, решения и анализа задач моделирования систем эконометрических уравнений.
3. Конкретные подходы к постановке, решению и анализу задач моделирования систем эконометрических уравнений.

Методические указания по проведению занятия

Согласно варианту задания к практическому занятию выбрать систему уравнений и произвести эконометрическое исследование.

Номер варианта проверяемого задания находится по таблице согласно начальной букве фамилии студента:

Начальная буква	А, Х, О	В, У, Ш	Д, Р, Щ	Е, П	И, Г, Ж	К, Ф, Э	Л, Ч, Ю	Б, М, Я	Н, Т	С, Ц, З
№ вар.	1	2	3	4	5	6	7	8	9	10

Отчет оформлять в Word, для ввода формул использовать Equation.

Исследования модели произвести согласно представленному порядку исследования.

Варианты заданий

Вариант 1

$$\begin{cases} M_t = a_1 + b_{12}N_t + b_{13}S_t + b_{14}E_{t-1} + b_{15}M_{t-1} + \varepsilon_1, \\ N_t = a_2 + b_{21}M_t + b_{23}S_t + b_{26}Y_t + \varepsilon_2, \\ S_t = a_3 + b_{31}M_t + b_{32}N_t + b_{36}X_t + \varepsilon_3. \end{cases}$$

Вариант 2

$$\begin{cases} C_t = a_1 + b_{12}Y_t + b_{13}T_t + \varepsilon_1, \\ I_t = a_2 + b_{21}Y_t + b_{24}K_{t-1} + \varepsilon_2, \\ Y_t = C_t + I_t. \end{cases}$$

Вариант 3

$$\begin{cases} C_t = a_1 + b_{11}Y_t + b_{12}C_{t-1} + \varepsilon_1, \\ I_t = a_2 + b_{21}Y_t + b_{23}r_t + \varepsilon_2, \\ r_t = a_3 + b_{31}Y_t + b_{34}M_t + b_{35}r_{t-1} + \varepsilon_3, \\ Y_t = C_t + I_t + G_t. \end{cases}$$

Вариант 4

$$\begin{cases} C_t = a_1 + b_{11}Y_t + b_{12}Y_{t-1} + \varepsilon_1, \\ I_t = a_2 + b_{21}Y_t + \varepsilon_2, \\ Y_t = C_t + I_t + G_t. \end{cases}$$

Вариант 5

$$\begin{cases} R_t = a_1 + b_{12}Y_t + b_{14}M_t + \varepsilon_1, \\ Y_t = a_2 + b_{21}R_t + b_{23}I_t + b_{25}G_t + \varepsilon_2, \\ I_t = a_3 + b_{31}R_t + \varepsilon_3. \end{cases}$$

Вариант 6

$$\begin{cases} C_t = a_1 + b_{11}Y_t + \varepsilon_1, \\ I_t = a_2 + b_{21}Y_t + b_{22}Y_{t-1} + \varepsilon_2, \\ Y_t = C_t + I_t + G_t. \end{cases}$$

Вариант 7

$$\begin{cases} C_t = a_1 + b_{11}D_t + \varepsilon_1, \\ I_t = a_2 + b_{22}Y_t + b_{23}Y_{t-1} + \varepsilon_2, \\ Y_t = D_t + T_t, \\ D_t = C_t + I_t + G_t. \end{cases}$$

Вариант 8

$$\begin{cases} C_t = a_1 + b_{11}Y_t + b_{12}J_t + \varepsilon_1, \\ J_t = a_2 + b_{21}Y_{t-1} + \varepsilon_2, \\ T_t = a_3 + b_{31}Y_t + \varepsilon_3, \\ Y_t = C_t + J_t + G_t. \end{cases}$$

Вариант 9

$$\begin{cases} R_t = a_1 + b_{11}M_t + b_{12}Y_t + \varepsilon_1, \\ Y_t = a_2 + b_{21}R_t + b_{22}I_t + \varepsilon_2, \\ I_t = a_3 + b_{33}R_t + \varepsilon_3. \end{cases}$$

Вариант 10

$$\begin{cases} C_t = a_1 + b_{11}Y_t + b_{12}C_{t-1} + \varepsilon_1, \\ I_t = a_2 + b_{21}r_t + b_{22}I_{t-1} + \varepsilon_2, \\ r_t = a_3 + b_{31}Y_t + b_{32}M_t + \varepsilon_3, \\ Y_t = C_t + I_t + G_t. \end{cases}$$

Порядок выполнения задания

1. Для каждого уравнения модели определите его идентифицируемость, применив необходимое и достаточное условие идентификации.
2. Определите метод оценки параметров модели.
3. Запишите в общем виде приведенную форму модели.

Примеры решения задач

Исследуем модель протекционизма Сальватора:

$$\begin{cases} M_t = a_1 + b_{12}N_t + b_{13}S_t + b_{14}E_{t-1} + b_{15}M_{t-1} + \varepsilon_1, \\ N_t = a_2 + b_{21}M_t + b_{23}S_t + b_{26}Y_t + \varepsilon_2, \\ S_t = a_3 + b_{31}M_t + b_{32}N_t + b_{36}X_t + \varepsilon_3. \end{cases}$$

1. Проверим систему уравнений на идентифицируемость.

Необходимое условие

Число эндогенных переменных $H = 3$ (M_t, N_t, S_t).

Число экзогенных переменных $D = 4$ (M_{t-1}, E_{t-1}, Y_t, X_t).

D – число предопределенных переменных, отсутствующих в данном уравнении, но присутствующих в системе.

1-е уравнение: $D = 3$, $H = 2$, $D + 1 > H$ – уравнение сверхидентифицируемо,

2-е уравнение: $D = 3$, $H = 3$, $D + 1 > H$ – уравнение сверхидентифицируемо,

3-е уравнение: $D = 3$, $H = 3$, $D + 1 > H$ – уравнение сверхидентифицируемо.

Проверим достаточное условие идентификации.

Уравнение 1: отсутствуют переменные Y_t, X_t. Найдем определитель матрицы из коэффициентов при этих переменных во втором и третьем уравнениях:

$$\Delta A = \begin{vmatrix} b_{26} & 0 \\ 0 & b_{36} \end{vmatrix} = b_{26}b_{36} \neq 0.$$

Достаточное условие идентификации выполняется.

Уравнение 2: отсутствуют переменные E_{t-1}, M_{t-1}, X_t. Матрица из коэффициентов при этих переменных в первом и третьем уравнениях:

$$A = \begin{pmatrix} b_{14} & b_{15} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Так как минор второго порядка

$$\Delta_2 = \begin{vmatrix} b_{15} & 0 \\ 0 & 1 \end{vmatrix} = b_{15} \neq 0,$$

то достаточное условие идентификации *выполняется*.

Уравнение 3: отсутствуют переменные E_{t-1}, M_{t-1}, Y_t. Матрица из коэффициентов при этих переменных в первом и втором уравнениях:

$$A = \begin{pmatrix} b_{14} & b_{15} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Так как минор второго порядка

$$\Delta_2 = \begin{vmatrix} b_{15} & 0 \\ 0 & 1 \end{vmatrix} = b_{15} \neq 0,$$

то достаточное условие идентификации выполняется. Следовательно, достаточное условие идентифицируемости *выполняется*.

2. Все три уравнения – сверхидентифицируемые. Для решения сверхидентифицируемых систем уравнений применяется двухшаговый метод наименьших квадратов.

3. Приведенная форма модели:

$$\begin{cases} M_t = A_1 + \sigma_{11}E_{t-1} + \sigma_{12}M_{t-1} + \sigma_{13}X_t + \sigma_{14}Y_t + u_1, \\ N_t = A_1 + \sigma_{21}E_{t-1} + \sigma_{22}M_{t-1} + \sigma_{23}X_t + \sigma_{24}Y_t + u_2, \\ S_t = A_1 + \sigma_{31}E_{t-1} + \sigma_{32}M_{t-1} + \sigma_{33}X_t + \sigma_{34}Y_t + u_3. \end{cases}$$

Рекомендуемая литература

1. Эконометрика : лабораторный практикум / сост. Н.А. Чечерова. – Комсомольск-на-Амуре : АмГПГУ, 2010. – 176 с.
2. Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика : начальный курс : учебник для вузов. – М. : Дело, 2004. – 575 с.
3. Практикум по эконометрике : учеб. пособие / под ред. И.И. Елисеевой. – М. : Финансы и статистика, 2003. – 192 с.

МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПО ВЫПОЛНЕНИЮ САМОСТОЯТЕЛЬНОЙ РАБОТЫ

Для выполнения заданий следует изучить соответствующие разделы дисциплины по пособию и (или) по предложенным в списке рекомендуемой литературы источникам. В пособии даются некоторые начальные теоретические сведения и приводятся решения типовых примеров. Если студент испытывает затруднения в освоении теоретического или практического материала, то он может получить устную или письменную консультацию у ведущего преподавателя.

При выполнении лабораторных и практических работ соблюдайте условия:

1. Каждая работа должна быть выполнена в отдельном файле на листах формата А4, используется текстовый редактор MS Word, для ввода формул – надстройка Equation.
2. В заголовке работы на титульном листе должны быть написаны фамилия, имя и отчество студента, название дисциплины, тема работы; фамилия, имя, отчество преподавателя; здесь же следует указать название учебного заведения, дату выполнения работы и проставить личную подпись студента.
3. В работу должны быть включены все пункты исследования, указанные в задании, по порядку и в соответствии с вариантом студента.
4. В каждом пункте исследования должны быть соответствующие расчеты, выполненные с помощью программы MS Excel, а также интерпретация результатов и выводы.
5. В прорецензированной работе необходимо исправить отмеченные преподавателем ошибки и учесть его рекомендации и советы.

Экзамен проводится в традиционной форме в виде устного опроса студентов по билетам курса, которые включают два теоретических вопроса и задачу. **К экзамену допускаются студенты, получившие зачет по всем лабораторным и практическим работам.**

Критерии оценки:

– «*отлично*» выставляется студенту, если он ответил на теоретические вопросы билета и правильно решил задачу;
– «*хорошо*» выставляется студенту, если он ответил на теоретические вопросы билета, но решил задачу с ошибками или недочетами;
– «*удовлетворительно*» выставляется студенту, если он ответил только на один вопрос билета и правильно решил задачу;
– «*неудовлетворительно*» выставляется студенту, если он не ответил ни на один из теоретических вопросов и не решил задачу.

ЗАКЛЮЧЕНИЕ

Изучив материал курса «Эконометрика (продвинутый уровень)», выполнив задания по лабораторным и практическим работам, студент будет

✓ *иметь представление*: об эконометрике как науке, классификации эконометрических моделей и данных, основных положениях регрессионного анализа;

✓ *владеть понятиями*: эконометрическая модель, регрессионная модель, модель временных рядов, система одновременных уравнений, объясняемые и объясняющие переменные, поле корреляции, среднее значение, математическое ожидание, выборочная дисперсия, стандартное отклонение, выборочный коэффициент корреляции, выборочная ковариация случайных величин, стандартные предположения регрессионного анализа, гомоскедастичность и гетероскедастичность дисперсии ошибок; эндогенные и экзогенные переменные;

✓ *уметь* вычислять вероятностные характеристики случайных величин, строить поле корреляции, исходя из эмпирических данных;

✓ *владеть* навыками анализа статистических данных.

Для оценки результатов обучения по данному курсу предлагается ответить на вопросы самоконтроля, приведенные ниже.

Вопросы к экзамену

1. Классификация данных в эконометрике.
2. Понятия гомоскедастичности и гетероскедастичности дисперсии ошибок.
3. Коэффициент корреляции.
4. Свойства коэффициента корреляции при большом объеме выборки.
5. Задачи эконометрики, понятие «эконометрическая модель». Составляющие модели в эконометрике.
6. Характеристики случайных величин.
7. Выборочная дисперсия и стандартное отклонение случайной величины.
8. Поле корреляции.
9. Противоречия подходов в эконометрическом моделировании и подходы к их разрешению.
10. Среднее значение случайной величины.
11. Виды эконометрических моделей.
12. Понятие регрессионной модели.
13. Регрессионные модели с одним уравнением.
14. Математическое ожидание случайной величины.
15. Понятие «эконометрика», история ее возникновения как самостоятельной дисциплины.
16. Стандартные предположения регрессионного анализа.
17. Модель парной линейной регрессии.
18. Оценка статистической значимости коэффициентов парной линейной регрессии и корреляции.
19. Статистические свойства МНК-оценок параметров уравнения регрессии.
20. Выборочная ковариация.
21. Геометрический смысл регрессионной модели, составляющие дисперсии.
22. Доверительный интервал для параметра σ_2 регрессионной модели.
23. Основная идея дисперсионного анализа.
24. Использование модели парной линейной регрессии для прогноза.

25. Графический метод проверки стандартных предположений регрессионного анализа.
26. Метод наименьших квадратов оценки параметров парной регрессионной модели.
27. Доверительный интервал для индивидуальных значений зависимой переменной.
28. Процедура проверки значимости линейной связи между переменными, использование F-критерия (критерия Фишера – Снедекора).
29. Отбор факторов в модель множественной регрессии.
30. Нахождение доверительного интервала для функции регрессии (для $M_x(y)$).
31. Определение коэффициента R^2 и его свойства.
32. Предельная склонность потребления в модели «доход – потребление».
33. Способ приведения степенной модели к линейной форме модели. Оценка параметров модели и ее качества.
34. Понятие предельной склонности и эластичности функции. Условия постоянства предельной склонности и эластичности функции.
35. Доверительный интервал для параметра β регрессионной модели.
36. Определение и примеры моделей множественной линейной регрессии.
37. Модели с убывающей эластичностью, их линеаризация.
38. Итерационные методы подбора нелинейных моделей.
39. Обратно пропорциональная зависимость, линеаризация этой модели и ее эластичность.
40. Метод наименьших квадратов оценивания параметров множественной линейной регрессии.
41. Проверка статистических гипотез о значениях отдельных коэффициентов.
42. Отбор факторов в модель линейной множественной регрессии.
43. Методы построения уравнения множественной регрессии.
44. Нелинейные модели множественной регрессии, приводимые к линейной форме.

45. Проверка значимости уравнения множественной линейной регрессии с помощью критериев Фишера и Стьюдента.
46. Понятие частных коэффициентов эластичности.
47. Понятие средних коэффициентов эластичности.
48. Коэффициенты множественной корреляции и детерминации.
49. Частные и общий коэффициенты корреляции.
50. Уравнение множественной регрессии в стандартизированном масштабе.
51. Сущность метода взвешенных наименьших квадратов (обобщенного МНК).
52. Понятие и примеры фиктивных переменных.
53. Трудности идентификации структурных моделей.
54. Алгоритм выявления автокорреляции остатков на основе критерия Дарбина – Уотсона.
55. Необходимое и достаточное условия идентифицируемости.
56. Оценка параметров систем одновременных уравнений.
57. Коэффициент автокорреляции, его свойства.
58. Условия идентифицируемости структурных моделей.
59. Косвенный метод наименьших квадратов.
60. Модели с качественными объясняющими переменными.
61. Двухшаговый метод наименьших квадратов.
62. Виды моделей временных рядов.
63. Классификация систем регрессионных уравнений.
64. Стационарные и нестационарные временные ряды.
65. Моделирование тенденции временного ряда.
66. Эконометрические модели, в которых объясняющие переменные носят как количественный, так и качественный характер.
67. Составляющие временного ряда.
68. Методы оценки параметров структурной модели.
69. Автокорреляция в остатках. Критерий Дарбина – Уотсона.
70. Проблемы инверсии и идентификации структурных моделей.
71. Автокорреляционная функция, коррелограмма, их анализ.
72. Моделирование сезонных колебаний временного ряда.

Библиографический список

1. Айвазян, С.А. Прикладная статистика. Основы эконометрики : учебник для экономических специальностей вузов. В 2 т. Т. 1 Теория вероятностей и прикладная статистика / С.А. Айвазян, В.С. Мхитарян. – Москва : ЮНИТИ-ДАНА, 2001. – 656 с.
2. Айвазян, С.А. Прикладная статистика. Основы эконометрики : учебник для экономических специальностей вузов. В 2 т. Т. 2 Основы эконометрики / С.А. Айвазян. – Москва : ЮНИТИ-ДАНА, 2001. – 432 с.
3. Воскобойников, Ю.Е. Эконометрика в Excel: парные и множественные регрессионные модели : учеб. пособие / Ю.Е. Воскобойников. – 2-е изд., стер. – Санкт-Петербург : Лань, 2018. – 260 с.
4. Домбровский, В.В. Эконометрика : учебник / В.В. Домбровский. – Москва : Новый учебник, 2004. – 342 с.
5. Дорохина, Е.Ю. Сборник задач по эконометрике : учеб. пособие для студентов экономических вузов / Е.Ю. Дорохина, Л.Ф. Преснякова, Н.П. Тихомиров. – Москва : Экзамен, 2003. – 224 с.
6. Доугерти, К. Введение в эконометрику / К. Доугерти ; пер. с англ. – Москва : ИНФРА-М, 1999. – 402 с.
7. Эконометрика : практикум для студентов экономических специальностей / П.Ф. Зибров, Н.В. Колачева, С.Ш. Палферова, С.В. Пивнева ; под ред. Ю.К. Черновой. – Тольятти : ТГУ, 2008. – 69 с.
8. Ивченко, Ю.С. Эконометрика в MS EXCEL : лабораторный практикум / Ю.С. Ивченко. – Саратов : Ай Пи Эр Медиа, 2018. – 94 с.
9. Ивченко, Ю.С. Эконометрика : курс лекций / Ю.С. Ивченко. – Саратов : Вузовское образование, 2018. – 121 с.
10. Кремер, Н.Ш. Эконометрика : учебник / Н.Ш. Кремер, Б.А. Путко ; под ред. Н.Ш. Кремера. – 3-е изд., перераб. и доп. – Москва : ЮНИТИ-ДАНА, 2017. – 328 с.
11. Кулинич, Е.И. Эконометрия / Е.И. Кулинич. – Москва : Финансы и статистика, 2001. – 304 с.
12. Магнус, Я.Р. Эконометрика. Начальный курс : учебник / Я.Р. Магнус, П.К. Катышев, А.А. Пересецкий. – Москва : Дело, 2004. – 576 с.

13. Носко, В.П. Эконометрика. Элементарные методы и введение в регрессионный анализ временных рядов / В.П. Носко. – Москва : Институт экономики переходного периода, 2004. – 501 с.
14. Практикум по эконометрике : учеб. пособие / под ред. И.И. Елисеевой. – Москва : Финансы и статистика, 2003. – 192 с.
15. Тихомиров, Н.П. Эконометрика : учебник / Н.П. Тихомиров, Е.Ю. Дорохина. – Москва : Экзамен, 2003. – 512 с.
16. Шалабанов, А.К. Эконометрика : учеб.-метод. пособие / А.К. Шалабанов, Д.А. Роганов. – Казань : ТИСБИ, 2002. – 56 с.
17. Эконометрика : учебник / под ред. И.И. Елисеевой. – Москва : Проспект, 2010. – 288 с.
18. Эконометрика : лабораторный практикум / сост. Н.А. Чечерова. – Комсомольск-на-Амуре : АмГПГУ, 2010. – 176 с.
19. Электронная библиотека: библиотека диссертаций : сайт / Российская государственная библиотека. – Москва : РГБ, 2003. – URL: http://diss.rsl.ru/?lang=ru (дата обращения: 20.07.2018).
20. Elibrary.ru : научная электронная библиотека : сайт. – Москва, 2000. – URL: https://elibrary.ru (дата обращения: 09.01.2018).

Глоссарий

№ п/п	Понятие	Содержание (формула)
1	Среднее значение дискретной величины	наиболее простой показатель, характеризующий последовательность $x_1, x_2, \ldots, x_n$ ($y_1, y_2, \ldots, y_n$), вычисляется по формуле $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ $\left(\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i\right)$
2	Математическое ожидание дискретной случайной величины	$M(X) = x_1p_1 + x_2p_2 + \ldots + x_np_n = \sum_{i=1}^{n} x_ip_i$ $(M(Y) = y_1p_1 + y_2p_2 + \ldots + y_ip_n = \sum_{i=1}^{n} y_ip_i)$
3	Выборочная вариация	$\mathrm{var}(x) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$ $\left(\mathrm{var}(y) = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2\right)$
4	Стандартное отклонение	$S(x) = \sqrt{\mathrm{var}(x)}$ $\left(S(y) = \sqrt{\mathrm{var}(y)}\right)$
5	Диаграмма рассеяния (поле корреляции)	удобное графическое средство анализа данных, на нем в прямоугольной системе координат располагаются точки (x_i, y_i), $i = 1, 2, \ldots, n$
6	Выборочный коэффициент корреляции	$r_{xy} = \frac{\mathrm{cov}(x,y)}{S(x)S(y)} = \frac{\overline{xy} - \overline{x}\cdot\overline{y}}{S_xS_y}$
7	Выборочная ковариация	$\mathrm{cov}(x,y) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y}) = \overline{xy} - \overline{x}\cdot\overline{y}$
8	Регрессионная модель	модель, где зависимая (объясняемая) переменная y представляется в виде функции $y = M_y(x_1, \ldots, x_p) + \varepsilon$
9	Эндогенные переменные	формирующиеся внутри переменные экономического объекта
10	Экзогенные переменные	задаваемые извне переменные экономического объекта
11	Лаговые переменные	взятые в предыдущий момент времени переменные

№ п/п	Понятие	Содержание (формула)
12	Условие равноизменчивости (гомоскедастичности) ошибок ε_i и соответственно объясняемой переменной y_i	$\mathrm{var}(\varepsilon_i) = \delta^2(\varepsilon_i) = \mathrm{const}$, $\mathrm{var}(y_i) = \delta^2(y_i) = \mathrm{const}$
13	Условие гетероскедастичности ошибок ε_i и y_i	$\mathrm{var}(\varepsilon_i) = \delta^2(\varepsilon_i) \neq \mathrm{const}$, $\mathrm{var}(y_i) = \delta^2(y_i) \neq \mathrm{const}$
14	Принцип наименьших квадратов	поиск коэффициентов (параметров) уравнения регрессии осуществляется таким образом, чтобы величина остатка регрессии ε_i стремилась к минимуму (в идеале к нулю): $\sum_{i=1}^{n\Sigma} \varepsilon_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \to \min$. Если $\varepsilon_i = 0$, то все наблюдаемые значения (x_i, y_i) лежат на подобранной линии регрессии. В результате получают теоретическую модель регрессии
15	Метод наименьших квадратов (МНК)	коэффициенты (параметры) уравнения регрессии вычисляются по формулам, обеспечивающим принцип наименьших квадратов ошибок регрессии
16	Модель парной линейной регрессии	$y = (\alpha + \beta x) + \varepsilon$, где y – наблюдаемое значение зависимой переменной (объясняемая переменная); x – объясняющая переменная; $\hat{y} = \alpha + \beta x$ – объясненная часть, зависящая от значения объясняющих переменных; α и β – параметры модели (рассчитываются МНК); ε – случайная составляющая
17	Несмещенность МНК-оценки параметров уравнения регрессии	если математическое ожидание статистической оценки некоторого параметра равно истинному значению этого параметра, то она называется несмещенной
18	Состоятельность МНК-оценки параметров уравнения регрессии	при неограниченном возрастании объема выборки значение оценки должно стремиться по вероятности к истинному значению параметра, а дисперсии оценок параметров должны уменьшаться и в пределе стремиться к нулю

№ п/п	Понятие	Содержание (формула)
19	Доверительный интервал для прогноза среднего значения y, полученного по функции регрессии	$\hat{y} - t_{1-\alpha;k}S_{\hat{y}} \le M_{\hat{y}}(x) \le \hat{y} + t_{1-\alpha;k}S_{\hat{y}}$, где $S_{\hat{y}}^2 = S_{\varepsilon_i}^2\left(\frac{1}{n} + \frac{(x-\bar{x})^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\right)$ – оценка дисперсии значения $\hat{y}$, найденного по уравнению регрессии; $t = \frac{\hat{y}-M_{\hat{y}}(x)}{S_{\hat{y}}}$ – статистика, имеющая t-распределение Стьюдента с $k = n - 2$ степенями свободы (n – объем выборки)
20	Доверительный интервал для прогноза индивидуальных значений y_0 по функции регрессии	$\hat{y}_0 - t_{1-\alpha;n-2}S_{\hat{y}_0} \le y_0 \le \hat{y}_0 + t_{1-\alpha;n-2}S_{\hat{y}_0}$, где $S_{\hat{y}_0}^2 = S_{\varepsilon_i}^2\left[1 + \frac{1}{n} + \frac{(x_0-\bar{x})^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\right]$ – оценка дисперсии значения $\hat{y}_0$, найденного по уравнению регрессии; $t = \frac{\hat{y}_0-y_0}{S_{\hat{y}_0}}$ – статистика, имеющая t-распределение Стьюдента с $k = n - 2$ степенями свободы (n – объем выборки)
21	Доверительные интервалы для параметров регрессионной модели	доверительный интервал выбирается таким образом, чтобы вероятность $P\left(\chi^2 < \chi^2_{1-\frac{\alpha}{2};n-2}\right) = P\left(\chi^2 > \chi^2_{1-\frac{\alpha}{2};n-2}\right) = \frac{\alpha}{2}$
22	Средняя ошибка аппроксимации	$\bar{A} = \frac{1}{n}\sum_{i=1}^{n}\left\|\frac{y_i - \hat{y}_i}{y_i}\right\| \cdot 100\ \% = \sum_{i=1}^{n}\left\|\frac{e_i}{y_i}\right\|$ (менее 8–10 %)
23	Модель парной нелинейной регрессии	$y = f(x) + \varepsilon$, где $\hat{y} = f(x)$ – объясненная часть, зависящая от значения объясняющих переменных, представляет нелинейную функцию; ε – случайная составляющая
24	Линеаризация уравнения нелинейной регрессии	приведение нелинейного уравнения регрессии к линейной форме путем логарифмирования или замены переменной с целью оценки параметров модели
25	Степенная модель парной регрессии с мультипликативными ошибками	модель $y = \alpha \cdot x^{\beta} \cdot \varepsilon$, линеаризуется путем логарифмирования: $\ln y = \ln\alpha + \beta\ln x + \ln\varepsilon$

№ п/п	Понятие	Содержание (формула)
26	Предельная склонность к потреблению (норма потребления)	величина $C(x)$, которая для заданной величины располагаемого дохода x и затрат V на некоторый товар определяется формулой $C(x) = \lim_{\Delta x \to 0} \frac{V(x+\Delta x)-V(x)}{\Delta x} = \frac{dV}{dx} = V'(x)$
27	Предельная склонность величины y по отношению к величине x	если имеется связь между какими-то переменными экономическими факторами x и y в виде $Y = f(x)$, то функция $C(x) = \frac{dy}{dx} = f'(x) = \lim_{\Delta x \to 0} \frac{y(x+\Delta x) - y(x)}{\Delta x}$ является *предельной склонностью* величины y по отношению к величине x
28	Эластичность (функция эластичности)	величина $\eta(x) = \lim_{\Delta x \to 0} \frac{\frac{f(x+\Delta x)-f(x)}{f(x)} \cdot 100\,\%}{\frac{\Delta x}{x} \cdot 100\,\%} = \frac{x}{y} \cdot \frac{dy}{dx}$, $\eta(x) = \frac{x}{y} C(x)$
29	Степенная модель парной регрессии с аддитивными возмущениями	$y = \alpha \cdot x^{\beta} + \varepsilon$, то есть имеет аддитивные ошибки. Не преобразуется к линейной модели наблюдений
30	Итерационные методы подбора нелинейных моделей	методы, в процессе реализации которых сначала задаются некоторые «стартовые значения» оцениваемых параметров, а затем производится последовательное приближение значений оценок параметров к реальным величинам, минимизирующим квадраты ошибок $\sum_{i=1}^{n\Sigma} \varepsilon_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \to \min$
31	Модели множественной регрессии	регрессионные модели с числом переменных более двух
32	Модель множественной линейной регрессии	$\hat{y}_x = a + b_1 x_1 + b_2 x_2 + \ldots + b_m x_m$
33	Степенная модель множественной регрессии с мультипликативными возмущениями	$y = a \cdot x_1^{b_1} \cdot x_2^{b_2} \cdot \ldots \cdot x_n^{b_n} + \varepsilon$ линеаризуется с помощью логарифмирования $\ln y = \ln a + b_1 \ln x_1 + b_2 \ln x_2 + \ldots + b_n \ln x_n + \ln \varepsilon$

№ п/п	Понятие	Содержание (формула)
34	Логарифмическая, правая и левая полулогарифмические модели множественной регрессии	$\ln y = a + b_1 \ln x_1 + b_2 \ln x_2 + \ldots + b_n \ln x_n + \varepsilon$, $y = a + b_1 \ln x_1 + b_2 \ln x_2 + \ldots + b_n \ln x_n + \varepsilon$, $\ln y = a + b_1 x_1 + b_2 x_2 + \ldots + b_n x_n + \varepsilon$. Для линеаризации целесообразно заменить логарифмы уровней: $y' = a + b_1 x_1' + b_2 x_2' + \ldots + b_n x_n' + \varepsilon$, $y = a + b_1 x_1' + b_2 x_2' + \ldots + b_n x_n' + \varepsilon$, $y' = a + b_1 x_1 + b_2 x_2 + \ldots + b_n x_n + \varepsilon$
35	Экспоненциальная модель множественной регрессии	$y = e^{a + b_1 x_1 + b_2 x_2 + \ldots b_n x_n + \varepsilon}$ приводится к виду $\ln y = a + b_1 \ln x_1 + b_2 \ln x_2 + \ldots + b_n \ln x_n + \varepsilon$ с помощью логарифмирования
36	Обратная модель множественной регрессии	$\frac{1}{y} = a + b_1 x_1 + b_2 x_2 + \cdots + b_n x_n + \varepsilon$, линеаризуется заменой переменной $\frac{1}{y} = y'$ $y' = a + b_1 x_1 + b_2 x_2 + \ldots + b_n x_n + \varepsilon$
37	Полиномиальная модель множественной регрессии	$y = a + b_1 x_1 + b_2 x_2^2 + b_3 x_3^3 + \ldots b_n x_n^n + \varepsilon$. Применяется замена переменных: $x_2^2 = x_2'$; $x_3^3 = x_3'$; ...; $x_n^n = x_n'$, преобразуется к виду: $y = a + bx_1 + b_2 x_2' + b_3 x_3' + \ldots + b_n x_n' + \varepsilon$
38	Степенная модель множественной регрессии	$y = a + b_1 x + b_2 x^2 + b_3 x^3 + \ldots b_n x^n + \varepsilon$. Линеаризуют заменой $x_1' = x$; $x_2' = x^2$; ...; $x_n' = x^n$, получают: $y = a + b_1 x_1 + b_2 x_2' + b_3 x_3' + \ldots + b_n x_n' + \varepsilon$
39	Интерактивная модель множественной регрессии	$y = a + bx_1 + cx_2 + dx_1 x_2 + \ldots + \varepsilon$. После замены переменных $x_1 x_2 = x'$, ... получают линейную модель вида: $y = a + bx_1 + cx_2 + dx' + \ldots + \varepsilon$
40	Степенная модель множественной регрессии с аддитивными возмущениями	$y = a \cdot x_1^{b_1} \cdot x_2^{b_2} \cdot \ldots \cdot x_n^{b_n} + \varepsilon$. К линейной форме не приводится. Для оценки ее параметров используются итерационные методы подбора нелинейных моделей
41	Статистическая гипотеза	априорное предположение, например о значениях параметров регрессионной модели
42	Уровень значимости статистического критерия	выбранное при формулировании статистического критерия значение вероятности ошибки α
43	Система независимых уравнений	$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n + \varepsilon_1, \\ y_2 = a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n + \varepsilon_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_m = a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n + \varepsilon_m \end{cases}$

№ п/п	Понятие	Содержание (формула)
44	Приведенная форма модели	$\begin{cases} y_1 = \delta_{11}x_1 + \delta_{12}x_2 + \ldots + \delta_{1n}x_n + u_1, \\ y_2 = \delta_{21}x_1 + \delta_{22}x_2 + \ldots + \delta_{2n}x_n + u_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_m = \delta_{m1}x_1 + \delta_{m2}x_2 + \ldots + \delta_{mn}x_n + u_m \end{cases}$
45	Статистический критерий проверки гипотезы H_o	правило решения вопроса об отклонении или неотклонении статистической гипотезы H_o
46	Метод взвешенных наименьших квадратов (обобщенный МНК)	применяется, когда построенная модель множественной регрессии не отвечает требованию гомоскедастичности возмущений. Найденные параметры такой модели не обладают свойствами несмещенности и эффективности, о чем можно судить по гетероскедастичности остатков. Согласно обобщенному (взвешенному) МНК оценки параметров определяются из условия минимума суммы взвешенных квадратов ошибок: $\frac{\sum_{i=1}^{n} e_i^2}{S_{\varepsilon_i}} = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{S_{\varepsilon_i}} \to \min$
47	Модель сезонности	$y(t) = S(t) + \varepsilon_t$
48	Аддитивная модель тренда и сезонности	$y(t) = T(t) + S(t) + \varepsilon_t$
49	Мультипликативная модель тренда и сезонности	$y(t) = T(t) \cdot S(t) \cdot \varepsilon_t$
50	Доверительная вероятность (надежность)	вероятность осуществления статистической гипотезы $\gamma = (1 - \alpha)$

Приложение 1

Значения *F*-критерия Фишера (уровень значимости α = 0,05)

k_2	k_1									
	1	2	3	4	5	6	8	12	24	
1	161,5	199,5	215,7	224,6	230,2	233,9	238,9	243,9	249,0	254,3
2	18,51	19,00	19,16	19,25	19,30	19,33	19,37	19,41	19,45	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,84	8,74	8,64	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,04	5,91	5,77	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,82	4,68	4,53	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4,00	3,84	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,73	3,57	3,41	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,28	3,12	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,23	3,07	2,90	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,07	2,91	2,74	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	2,95	2,79	2,61	2,40
12	4,75	3,88	3,49	3,26	3,11	3,00	2,85	2,69	2,50	2,30
13	4,67	3,80	3,41	3,18	3,02	2,92	2,77	2,60	2,42	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,70	2,53	2,35	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,64	2,48	2,29	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,59	2,42	2,24	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,55	2,38	2,19	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,51	2,34	2,15	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,48	2,31	2,11	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,45	2,28	2,08	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,42	2,25	2,05	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,40	2,23	2,03	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,38	2,20	2,00	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,36	2,18	1,98	1,73
25	4,24	3,38	2,99	2,76	2,60	2,49	2,34	2,16	1,96	1,71
26	4,22	3,37	2,98	2,74	2,59	2,47	2,32	2,15	1,95	1,69

k_2	k_1									
	1	2	3	4	5	6	8	12	24	
27	4,21	3,35	2,96	2,73	2,57	2,46	2,30	2,13	1,93	1,67
28	4,20	3,34	2,95	2,71	2,56	2,44	2,29	2,12	1,91	1,65
29	4,18	3,33	2,93	2,70	2,54	2,43	2,28	2,10	1,90	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,27	2,09	1,89	1,62
35	4,12	3,26	2,87	2,64	2,48	2,37	2,22	2,04	1,83	1,57
40	4,08	3,23	2,84	2,61	2,45	2,34	2,18	2,00	1,79	1,51
45	4,06	3,21	2,81	2,58	2,42	2,31	2,15	1,97	1,76	1,48
50	4,03	3,18	2,79	2,56	2,40	2,29	2,13	1,95	1,74	1,44
60	4,00	3,15	2,76	2,52	2,37	2,25	2,10	1,92	1,70	1,39
70	3,98	3,13	2,74	2,50	2,35	2,23	2,07	1,89	1,67	1,35
80	3,96	3,11	2,72	2,49	2,33	2,21	2,06	1,88	1,65	1,31
90	3,95	3,10	2,71	2,47	2,32	2,20	2,04	1,86	1,64	1,28
100	3,94	3,09	2,70	2,46	2,30	2,19	2,03	1,85	1,63	1,26
125	3,92	3,07	2,68	2,44	2,29	2,17	2,01	1,83	1,60	1,21
150	3,90	3,06	2,66	2,43	2,27	2,16	2,00	1,82	1,59	1,18
200	3,89	3,04	2,65	2,42	2,26	2,14	1,98	1,80	1,57	1,14
300	3,87	3,03	2,64	2,41	2,25	2,13	1,97	1,79	1,55	1,10
400	3,86	3,02	2,63	2,40	2,24	2,12	1,96	1,78	1,54	1,07
500	3,86	3,01	2,62	2,39	2,23	2,11	1,96	1,77	1,54	1,06
1000	3,85	3,00	2,61	2,38	2,22	2,10	1,95	1,76	1,53	1,03
∞	3,84	2,99	2,60	2,37	2,21	2,09	1,94	1,75	1,52	1

Приложение 2

Значения d_L d_U (уровень значимости 5 %)

n	*k*									
	1		2		3		4		5	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
6	0,61	1,40								
7	0,70	1,36	0,47	1,90						
8	0,76	1,33	0,56	1,78	0,37	2,29				
9	0,82	1,32	0,63	1,70	0,46	2,13				
10	0,88	1,32	0,70	1,64	0,53	2,02				
11	0,93	1,32	0,66	1,60	0,60	1,93				
12	0,97	1,33	0,81	1,58	0,66	1,86				
13	1,01	1,34	0,86	1,56	0,72	1,82				
14	1,05	1,35	0,91	1,55	0,77	1,78				
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	0,85	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	1,99
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