Министерство образования и науки Российской Федерации

Амурский государственный университет

Н.Н. Двоерядкина, Т.А. Юрьева, Т.Е. Гришкина

Методические указания к лабораторным занятиям по дисциплине **«Эконометрика»**

для направления подготовки 38.03.02 – «Менеджмент»

Благовещенск

Издательство АмГУ

2015

ББК
Ч

Рекомендовано
учебно-методическим советом университета

Рецензент:
Труфанова Т.В., канд. тех. наук, доц. кафедры математического анализа и моделирования АмГУ

Двоерядкина Н.Н.

Методические указания к лабораторным занятиям по дисциплине «Эконометрика» для направления подготовки 38.03.02 «Менеджмент». / Н.Н. Двоерядкина, Т.А. Юрьева, Т.Е. Гришкина. – Благовещенск: Изд-во АмГУ, 2015. – 37 с.

Методические указания предназначены для студентов очной формы обучения направления подготовки «38.03.02 «Менеджмент». В них приводятся материалы для проведения лабораторных занятий, а также контрольные вопросы для самопроверки знаний.

ВВЕДЕНИЕ

В настоящее время широкое распространение получило использование моделирования и количественного анализа, на базе которых сформировалось одно из направлений экономических исследований – эконометрика.

Эконометрика – наука, исследующая количественные закономерности и взаимозависимости в экономике при помощи методов математической статистики. Основа этих методов – корреляционно – регрессионный анализ.

Эконометрика занимается построением, статистической оценкой и анализом экономических зависимостей и моделей на основе изучения эмпирических данных. Одним из важных направлений эконометрики является прогнозирование различных экономических показателей.

Эконометрические методы и модели широко применяются во всех производственных и коммерческих фирмах для принятия практических решений в прогнозировании, банковском деле, бизнесе, а так же для разработки вариантов перспективного развития предприятия.

В пособии представлены задания для организации работы на занятии с краткими методическими указаниями по их выполнению в программе *Statistica* и задания для домашней самостоятельной работы.

Методические указания к лабораторным занятиям по дисциплине «Эконометрика» ориентированы на студентов направления подготовки «Менеджмент». Также он будет полезен преподавателям и всем интересующимся статистическими методами анализа социально - экономических процессов.

Предполагается, что студенты, изучающие эконометрику, уже прослушали базовые курсы по высшей математике, теории вероятностей и математической статистики, микро- и макро- экономике.

1. Линейная модель множественной регрессии

Цель: сформировать умения использовать метод наименьших квадратов для отыскания параметров линейных моделей парной и множественной регрессии.

Основные вопросы:

1. Линейная модель парной и множественной регрессий
2. Метод наименьших квадратов.
3. Описательные статистики.
4. Корреляционная матрица.

Указания для выполнения лабораторной работы:

Программа *Statistica* состоит из отдельных модулей, каждый из которых располагается в отдельном окне. Первым модулем является *Basic Statistics* – основные статистики. Окно *Basic Statistics* – представлено в виде таблицы с данными и пунктов меню (рис.1).

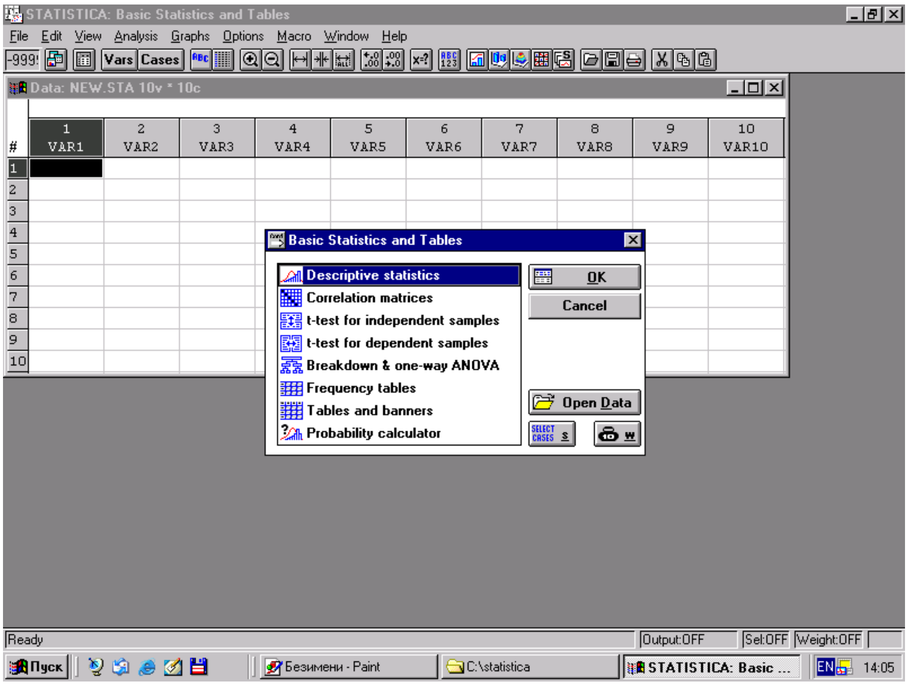


Рис.1

Таблица данных состоит из строк и столбцов. Столбцы используют для задания имен переменных (Variables), строки - для заполнения наблюдений (Cases). Строки и столбцы можно редактировать, выполнив двойной щелчок мыши на названии строки, столбца.

Диалоговое окно задания переменной позволяет:

а) name - задать имя переменной,

б) column width - ширину столбца в символах,

в) decimals - количество знаков после запятой,

г) category - тип данных (например, number - числовой).

Строки, столбцы в таблице можно добавлять(Add), удалять(Delete), перемещать(Move) и др. Данные действия выполняются при помощи команд Edit/Variables (работа со столбцами) (рис.2), Edit/Cases (работа со строками) (рис.3).

Команда Правка/Удалить переменные удаляет столбцы.

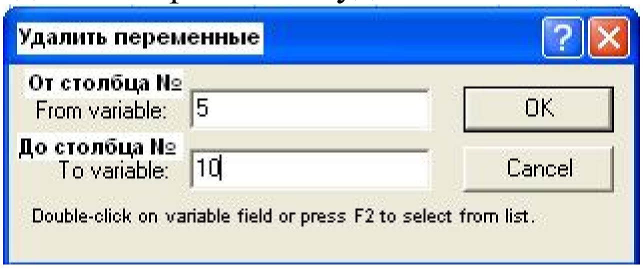


Рис. 2

Команда Вставка/Добавить регистры добавляет строки.

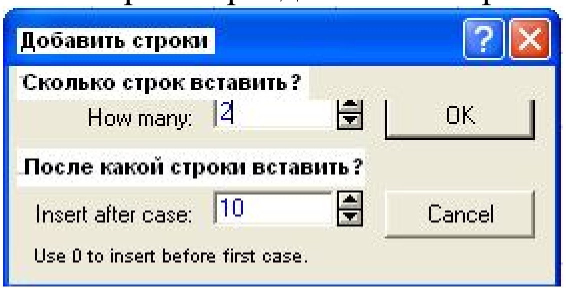


Рис.3

Например, нужно добавить один столбец после столбца с переменной So. Окно диалога будет выглядеть, как на рис. 4.

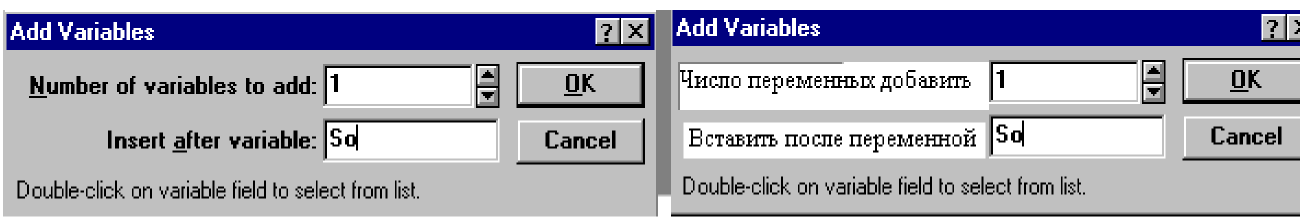


Рис. 4

Для нахождения значений основных статистик, в том числе и среднего значения, в программе *Statistica* необходимо выбрать пункт *Analysis/Descriptive Statistiscs*. В открывшемся окне диалога выбрать переменные, которые будут анализироваться (кнопка «*Variables*»). Клавиша «*Detailed Descriptive Statistisc*» вычисляет выборочное среднее значение (*Mean*), минимальное значение (*min*), максимальное значение *(max)* и др.

Можно указать два варианта рассмотрения взаимосвязей между двумя переменными. В первом случае обе переменные считаются равноценными, они не подразделяются на зависимую и независимую. Основным в этом случае является вопрос о наличии и силе взаимосвязи между переменными. При исследовании силы линейной связи обращаются к корреляционному анализу, основной мерой в котором является коэффициент корреляции.

Для нахождения коэффициента корреляции используется пункт *Analysis/Correlation matrices,* который позволяет просмотреть корреляционную матрицу, т.е. матрицу элементами которой являются коэффициенты корреляции, вычисленные для выбранных переменных по формуле:

$r_{xy} = \dfrac{\overline{xy} - \overline{x} \cdot \overline{y}}{\sqrt{\overline{x^2} - \overline{x}^2} \cdot \sqrt{\overline{y^2} - \overline{y}^2}}$, где $\overline{x}$, $\overline{y}$ - средние значения переменных.

Парный коэффициент находится на пересечение соответствующей строки и столбца.

Клавиша «*Two lists*» используется для выбора переменных, размещаемых в корреляционной матрице по строкам (first var) и столбцам (second var).

Коэффициент корреляции Пирсона показывает линейную меру зависимости между переменными.

Коэффициент корреляции изменяются в пределах от -1.00 до +1.00.

Отрицательное значение корреляции означает наличие убывающей зависимости. Положительное значение означает возрастающую зависимость.

Таблица 1

r<0	r>0	Мера зависимости
[-1 . .-0,9]	[0,9..1]	Отличная (очень значительная)
(-0,9 .. -0,7]	[0,7 .. 0,9)	Хорошая (значительная)
(-0,7 .. -0,5]	[0,5 .. 0,7)	Средняя (заметная)
(-0,5 .. -0,3]	[0,3..0,5)	Низкая (умеренная)
(-0,3 .. -0,05]	[0,05..0,3)	Очень низкая (слабая)
(-0,05 ..0]	[0 ..0,05)	Отсутствует

Другой вариант рассмотрения взаимосвязи между переменными выделяет одну из величин как независимую *x*, а другую как зависимую *y.* И изучает влияние этих переменных друг на друга.

Зависимость $y=f(x)$ называется функцией регрессии *y* на *x*.

Если рассматривается зависимость двух величин, то регрессия называется парной.

Для определения вида парной регрессии в декартовой системе координат строят точки наблюдений и соединяют их отрезками. Полученную линию называют эмпирической линией регрессии.

По внешнему виду эмпирической линии регрессии определяют плавную кривую, около которой группируются все точки наблюдений. Эту кривую называют теоретической линией регрессии или регрессией.

Самой простой парной регрессией является линейная регрессия с уравнением: $y = \alpha + \beta \cdot x + \varepsilon$, где α и β – коэффициенты регрессии, которые находятся с помощью метода наименьших квадратов, ε – случайный член.

β – показывает на сколько изменяется *y* при изменении переменной *x* на 1 единицу; α –это первоначальное значение *y* при $x=0$. Коэффициент α не имеет четкого экономического объяснения. Если $\alpha <0$, то результат изменяется быстрее фактора. Если $\alpha >0$, то результат изменяется медленнее фактора.

В программе *Statistica* пункт *Graphs/Stat 2D Graphs/ Scatterplot* позволяет отображать в декартовой системе координат диаграмму рассеяния точек наблюдения и применяет метод наименьших квадратов для построения модели.

После выбора пункта меню открывается диалоговое окно, в котором задаются:

переменные по осям OX (независимая переменная) и OY (зависимая переменная) (рис.5);

выбирается модель: *Regular* – парная, *Multiple* – множественная и др.;

задается вид регрессионной модели: *Linear* – линейная, *Logarifmic* – логарифмическая, *Exponention* - экспоненциальная и др.

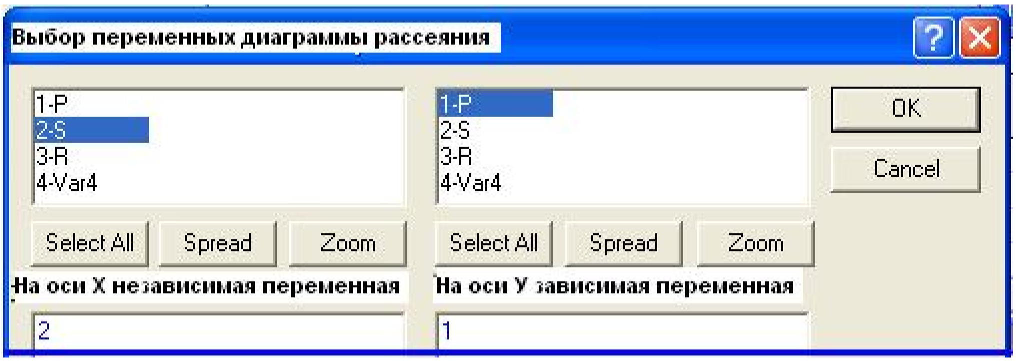


Рис.5

<u>Задания</u>

1. Выберите пункты меню Пуск/Программы/ Statistica/ Basic Statistics. Откройте новый файл, выбрав пункт File/New Data. Сохраните файл с англоязычным именем.

2. Задайте переменные:

Price: ширина столбца - 6, количество знаков после запятой - 0;

So: ширина столбца - 5, количество знаков после запятой - 1;

R: ширина столбца - 4, количество знаков после запятой - 1.

3. В вашей таблице 10 строк. Добавьте две строки, чтобы у Вас получилось 12 наблюдений. Оставьте в таблице три столбца, остальные удалите. В итоге у Вас получится таблица размером 3x12.

4. Внесите в таблицу 1 данные об условной стоимости квартир в г. N Price - цены, тыс. руб, So - общая площадь, м^2, R - расстояние до центра, км.

Таблица 2

	PRICE	SO	R
1	460	46	7
2	350	44	3,5
3	490	57,6	5
4	470	53,1	1,2
5	350	50	5
6	450	52,1	1,5
7	300	48	9
8	370	53	6,5
9	380	49	5
10	430	42	1,2
11	400	44	2
12	420	41	0

5. Найдите выборочное среднее значение переменных.

6. Дайте определение коэффициента корреляции. Постройте корреляционную матрицу, в которой в строке находится переменная Price, в столбцах переменные So и R. Найдите коэффициенты корреляции r(Price, So) и r(Price,R), объясните их значения.

7. Отобразить на графиках точки наблюдений Price(So), Price(R) зарисуйте схематично графики в тетрадь;

8. Оцените регрессии: Price = $\alpha + \beta$ * So и Price = $\alpha + \beta$ * R. Дайте экономическое объяснение полученным регрессиям.

2. Показатели качества регрессии

Цель: научиться оценивать качество уравнения регрессии в целом и параметров уравнения регрессии

Основные вопросы:

1. Стандартизированные и нестандартизированные коэффициенты регрессии.

2. Коэффициент детерминации.

3. F-статистика для оценки надежности уравнения.

4. t-статистика для оценки статистической значимости коэффициентов регрессии.

Указания по выполнению лабораторной работы:

Рассмотрим модель множественной линейной регрессии вида:

$y = \alpha + \beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \ldots + \beta_n \cdot x_n + \varepsilon$,

где y – зависимая переменная (*depended var*),

$x_1, x_2, \ldots, x_n$ – независимые переменные (*undepended var),*

ε –случайная величина.

После того как построено уравнение регрессии, необходимо оценить значимость как уравнения в целом, так и отдельных его параметров.

Для оценки коэффициентов множественной регрессии в программе *Statistica* используется модуль *Multiple Regression*, который открывается пунктами *Analysis /Other Statistics / Multiple Regression..* Первоначально в окне множественной регрессии появляется окно диалога, в котором кнопка *«Variables»* позволяет указать зависимые и независимые переменные (рис.6).

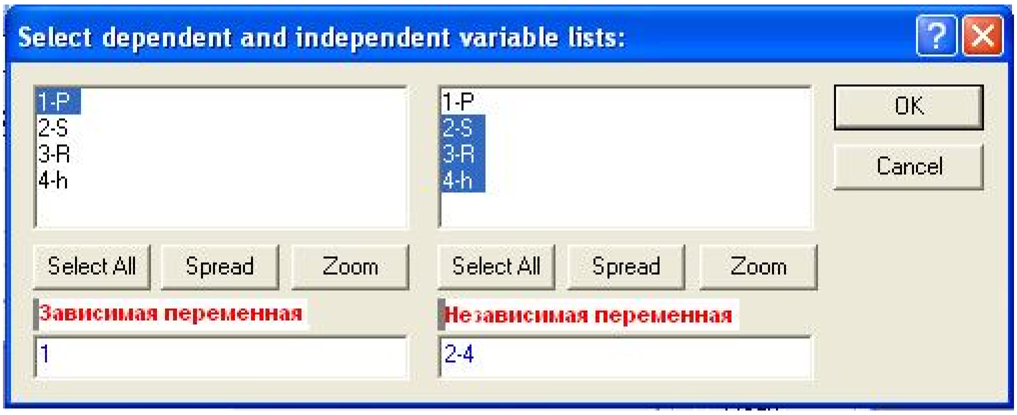


Рис.6

После нажатия на клавишу ОК появляется окно результата множественной регрессии (рис.7).

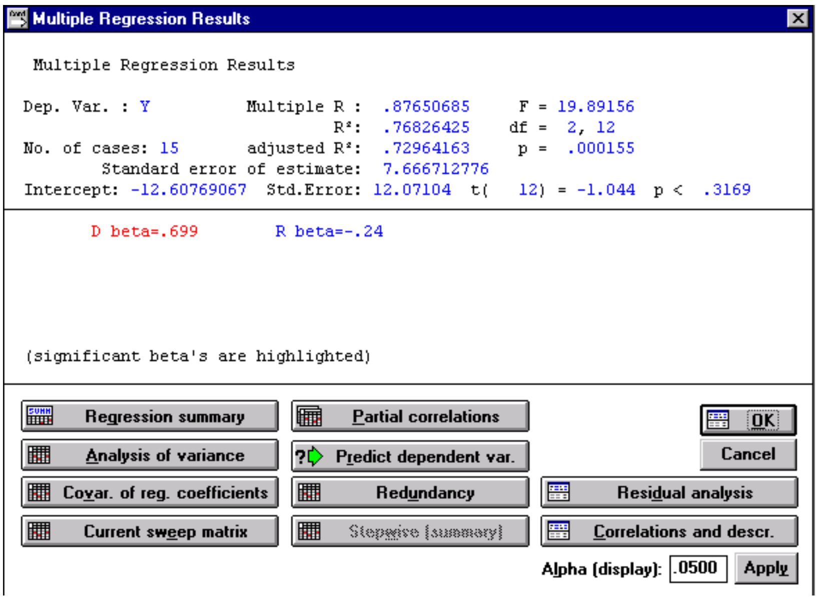


Рис.7

Dep var- зависимая переменная;

No of cases – количество наблюдений (объём выборки);

R^2 - коэффициент детерминации;

R^2_{adj} - скорректированный коэффициент детерминации;

F – статистика;

dF(k, n-k-1) – число степеней свободы критерия Фишера;

Standart Error of estimate – стандартная ошибка регрессии;

p – уровень значимости (вероятность ошибки прогноза).

Оценки коэффициентов модели появляются при нажатии клавиши *«Regression Summary»* (рис.8).

Regression Summary for Dependent Variable: Y (ahha.sta)

Continue...

R= .87650685 RI= .76826425 Adjusted RI= .72964163
F(2,12)=19.892 p<.00015 Std.Error of estimate: 7.6667

N=15	BETA	St. Err. of BETA	B	St. Err. of B	t(12)	p-level
Intercpt			-12.6077	12.07104	-1.04446	.316858
D	.698990	.187161	5.4995	1.47255	3.73471	.002849
R	-.238113	.187161	-.0219	.01722	-1.27224	.227389

Рис.8

Третий столбец *«B»* показывает оценку коэффициентов, столбец *«St.Err.of B»* указывает ошибку измерения коэффициентов, а пятый столбец – вычисленную *t* – статистику.

Согласно приведенным в таблице данным модель имеет вид:

$$Y = -12.6077 + 5.4995 D - 0.0219 R$$

Статистическая значимость коэффициентов множественной линейной регрессии проверяется на основе *t-* статистики по основной схеме проверки статистических гипотез.

Формулируются нулевая и альтернативная гипотезы:

$H_0 : \beta_i = 0$ - коэффициент β_i статистически незначим, т.е. переменная x_i линейно не связана с зависимой переменной *y*. Её наличие среди объясняющих переменных не оправдано со статистической точки зрения. Не оказывая серьезного влияния на зависимую переменную она лишь искажает реальную картину взаимосвязи.

$H_1 : \beta_i \neq 0$ коэффициент β_i статистически значим.

С помощью компьютерного пакета определяются $t_{набл}$ для каждой переменной, включенной в модель, а по таблицам распределения Стьюдента при заданном уровне значимости α - $t_{крит}$ имеющее *n-k-1* степеней свободы, где *n* – объём выборки, *k* – количество независимых переменных.

Строятся доверительные интервалы, и выделяется область принятия нулевой гипотезы:

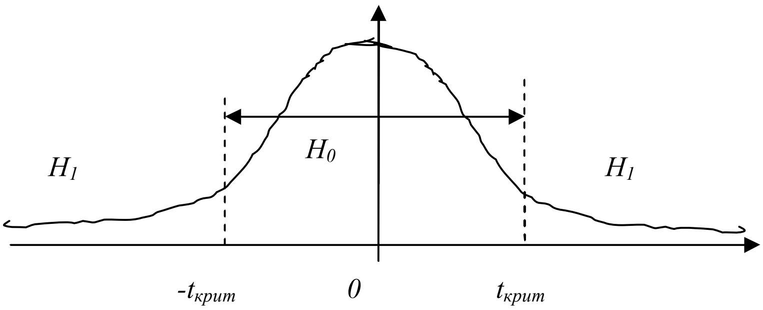


Рис.9

Если $t_{набл}$ попадает в область принятия нулевой гипотезы, то соответствующий параметр считаем статистически незначимым.

После проверки значимости каждого коэффициента регрессии проверяется общее качество уравнения. Для этой цели используется коэффициент детерминации R^2, который показывает долю дисперсии объясненной регрессией. Чем больше дисперсии объясняется регрессией, тем значимее уравнение регрессии, тем ближе значение R^2 к 1.

Коэффициент детерминации, показывает качество модели.

Таблица 3

r>0	Мера зависимости
[0,9..1]	Отличная (очень значительная)
[0,7 .. 0,9)	Хорошая (значительная)
[0,5 .. 0,7)	Средняя (заметная)
[0,3..0,5)	Низкая (умеренная)
[0,3..0,05)	Очень низкая (слабая)
[0 ..0,05)	Отсутствует

Оценив индивидуально каждый коэффициент необходимо проанализировать их совокупную значимость.

$$H_0 : \beta_1 = \beta_2 = \beta_3 = 0$$

$$H_1 : \beta_1 \neq 0, \beta_2 \neq 0, \beta_3 \neq 0$$

Для проверки данных гипотез используется следующая F – статистика:

$$F_{набл} = \frac{R^2}{1-R^2} \cdot \frac{n-k-1}{k},$$

где R^2 - коэффициент детерминации,

n – объём выборки,

k – количество независимых переменных.

$F_{крит}$ рассчитывается по таблице распределения Фишера при заданном уровне значимости α и числе степеней свободы $v_1=n-k-1;\ v_2=k$.

Если нулевая гипотеза не отклоняется , то делается вывод о том, что совокупное влияние всех объясняющих переменных модели на зависимую переменную *y* можно считать несущественным, а качество модели невысоким.

Качественные модели используют для прогнозирования. Клавиша *«Predict depended var»* осуществляет прогнозирование. Результат появляется в диалоговом окне представленном в табл.4.

Таблица 4

Predictd	*17,90579*	*Прогноз*
-95,0%CL	*5,40461*	*Доверительные интервалы прогнозирования*
+95,0%CL	*30,40697*	

P=17,9 – точечный прогноз

5,4<17,9<30,4 – интервалы прогнозирования

Задания:

1. Имеется файл с данными о стоимости квартир в г. N (табл. 2). Постройте модель зависимости цены от всех переменных. Запишите полученную модель в тетрадь. Объясните значение R^2. Проведите F – статистику, укажите проверяемые гипотезы.

2. Просмотрите результаты оценки коэффициентов. Выпишите значения t – статистик, укажите проверяемые гипотезы, проверьте их для каждого коэффициента.

3. Спрогнозируйте цену квартиры, значения переменных задайте самостоятельно. Для прогнозирования используйте клавишу *«Predict depended var».*

3. Мультиколлинеарность

Цель: выяснить суть мультиколлинеарности. Научиться обнаруживать и при необходимости устранять последствия мультиколлинеарности в модели множественной регрессии

Основные вопросы:

1. Выявление мультиколлинеарности.

2. Обнаружение последствий мультиколлинеарности.

3. Вычисление частных коэффициентов корреляции их связь с парными коэффициентами корреляции.

4. Устранение коллинеарности в модели. Спецификация модели по переменным-факторам.

Указания к лабораторной работе

Мультиколлинеарность - это наличие линейной связи между двумя или несколькими объясняющими переменными. Мультиколлинеарность бывает полная и частная. Полная мультиколлинеарность наблюдается в случае, когда объясняющие переменные связаны строгой функциональной зависимостью ($r_{xy}=1$). Частичная мультиколлинеарность наблюдается в случае, когда объясняющие переменные связаны сильной корреляционной зависимостью, но эта зависимость не является функциональной ($0,5<r_{xy}<1$).

Признаки мультиколлинеарности:

1) Коэффициент детерминации достаточно высок, но некоторые из коэффициентов регрессии статистически незначимы, т.е. они имеют низкие t – статистики.

2) Небольшое изменение исходных данных приводит к существенному изменению оценок коэффициентов регрессии.

3) Оценки коэффициентов имеют неправильные с точки зрения теории знаки или неоправданно большие значения.

4) Частные коэффициенты корреляции имеют большие значения.

5) Сильная вспомогательная регрессия.

Методы устранения мультиколлинеарности.

1) Исключение переменных из моделей, однако в этой ситуации возможны ошибки спецификации.

2) Попробовать добавить новую переменную, которая возможно была упущена, однако, если она тоже будет иметь сильную зависимость от других

переменных, то её введение может ещё больше усугубить проблему мультиколлинеарности.

3) Увеличить число наблюдений, однако получение новой выборки не всегда возможно или связано с серьезными издержками.

Задания

1. Найти частные коэффициенты корреляции переменной PRICE (табл. 1) со всеми независимыми переменными, сравнить их с обычными коэффициентами корреляции;

2. Выбрать переменные, которые сильнее всего влияют на PRICE, и построить уравнение регрессии, включив в него эти переменные;

3. Найти частные коэффициенты корреляции независимых переменных, включенных в модель. Сделать предположение о наличии мультиколлинеарности в модели;

4. Проверить предположение о наличии мультиколлинеарности с помощью вспомогательной регрессии;

5. Избавиться от мультиколлинеарности путем удаления коррелированных переменных из модели. Полученное уравнение регрессии записать в тетрадь и дать экономическую интерпретацию коэффициентам регрессии.

4. Линейные регрессионные модели с гетероскедастичными и автокоррелированными остатками

Цель: сформировать умения определять и при необходимости устранять гетероскедастичность остатков в моделях множественной регрессии.

Основные вопросы:

1. Построение линейной регрессионной модели с гетероскедастичными остатками.

2. Расчет остатков регрессионной модели. Построение графика распределения остатков.

3. Зависимость остатков и независимых переменных. Устранение гетероскедастичности с помощью обобщенного метода наименьших квадратов.

Указания по выполнению лабораторной работы

Рассмотрим модель множественной регрессии: $y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n + \varepsilon$. Переменная y зависит не только от объясняющих переменных x_i , коэффициентов регрессии α, β_i, но и от случайной величины ε. Для того, чтобы регрессия имела хорошее качество необходимо, чтобы разброс значений случайного члена был стабильным, т.е. дисперсия ошибки была независимой от переменных модели ($\sigma^2 = const$).

Условие независимости дисперсии ошибки называется гомоскедастичностью. Если разброс остатков подчиняется некоторому закону, то наблюдается гетероскедастичность в модели ($\sigma^2 \neq const$).

Для обнаружения гетероскедастичности существует несколько тестов и критериев, но все они проверяют справедливость гипотез:

$H_0 : \sigma^2 = const$ - модель гомоскедастична;

$H_1 : \sigma^2 \neq const$ - модель гетероскедастична.

Графический анализ остатков. По оси абсцисс откладывают значения объясняющей переменной, а по оси ординат либо отклонения ε_i, либо их квадраты ε_i^2. Если все отклонения находятся внутри полосы постоянной ширины, параллельной оси абсцисс, то модель гомоскедастична.

Для анализа остатков в окне *Multiple Regression* нажмите клавишу *Results Residual analysis*, а затем в открывшемся окне *Display residuals $ pred.* Третий столбец таблицы *Residual* показывает остатки ε_i.

Predicted & Residual Values (ahha.sta)

Continue... Dependent variable: Y

Case No.	Observed Value	Predictd Value	Residual	Standard Pred. v.	Standard Residual	Std.Err. Pred.Val
1	10.00000	15.99879	-5.99879	-.06715	-.78245	3.286880
2	50.00000	29.74325	20.25675	.99634	2.64217	2.936944
3	0.00000	6.65839	-6.65839	-.78987	-.86848	3.559447
4	40.00000	39.74978	.25022	1.77060	.03264	4.205288
5	5.00000	13.82545	-8.82545	-.23532	-1.15114	3.954526
6	20.00000	20.96604	-.96604	.31719	-.12600	3.466517
7	3.00000	-1.57328	4.57328	-1.42681	.59651	3.531309
8	2.00000	-2.12765	4.12765	-1.46970	.53839	3.777624
9	8.00000	7.10541	.89459	-.75529	.11668	2.595940
10	12.00000	8.88469	3.11531	-.61761	.40634	5.061101
11	20.00000	22.05271	-2.05271	.40127	-.26774	2.156269
12	6.00000	5.01735	.98265	-.91685	.12817	2.810307
13	30.00000	27.76253	2.23747	.84308	.29184	2.966625
14	20.00000	24.24811	-4.24811	.57115	-.55410	2.448325
15	27.00000	34.68843	-7.68843	1.37897	-1.00283	3.529928
Minimum	0.00000	-2.12765	-8.82545	-1.46970	-1.15114	2.156269
Maximum	50.00000	39.74978	20.25675	1.77060	2.64217	5.061101
Mean	16.86667	16.86667	.00000	.00000	.00000	3.353460

Рис.10

Тест Уайта (White, 1980г.).

К обычной модели применяют обычный метод наименьших квадратов и находят остатки ε_i. Осуществляется регрессия ε_i^2 на все независимые переменные x_i. Если построенная регрессия имеет хорошее качество (коэффициент детерминации близок к 1, F – статистика имеет большое значение), то нулевая гипотеза не принимается и говорят о наличии гетероскедастичности в модели.

Тест Голдфелда – Квандта (Goldfeld - Quandt).

Суть этого теста состоит в следующем:

1) Все n наблюдений упорядочиваются по независимой переменной x_i.

2) Упорядоченная выборка визуально разбивается на три части размерностей m, n-$2m$, m соответственно.

3) Оцениваются отдельно регрессии для первой подвыборки (m первых наблюдений) и для третьей (m последних наблюдений). Если предположение о наличии гетероскедастичности верно, то дисперсия регрессии по первой подвыборке ESS_1 будет значительно меньше дисперсии регрессии по третьей подвыборке ESS_3 .

4) Для сравнения соответствующих дисперсий строится следующая F – статистика: $F_{набл} = \frac{ESS_3}{ESS_1}$, по таблице распределения Фишера при заданном уровне значимости α находится $F_{крит}$ с числами степеней свободы $v_1=v_2=m-k-1$, где m – объем подвыборки, k – количество независимых переменных в уравнении регрессии.

5) Если $-F_{крит} < F_{набл} < F_{крит}$, то принимаем нулевую гипотезу, т.е. модель гомоскедастична, в противном случае говорим о наличии гетероскедастичности.

Для упорядочения наблюдений по переменной в *Basic Statistics* выберите *Analysis / Other Statistics/ Data Management/ Sort Cases*. Укажите переменную, по которой хотите упорядочить наблюдения.

Для устранения гетероскедастичности воспользуемся взвешенным методом наименьших квадратов (ВМНК). Он применяется к преобразованным данным – взвешенным по гетероскедастичному регрессору x_1.

Разделим обе части равенства $y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n + \varepsilon$ на переменную x_1 получим:

$$\frac{y}{x_1} = \frac{\alpha}{x_1} + \beta_1 + \beta_2 \frac{x_2}{x_1} + \ldots + \beta_n \frac{x_n}{x_1} + \frac{\varepsilon}{x_1}$$

Переобозначим $\frac{y}{x_1} = y^*$, $\frac{1}{x_1} = x_1^*$, $\frac{x_2}{x_1} = x_2^*$, …, $\frac{x_n}{x_1} = x_n^*$, $\frac{\varepsilon}{x_1} = \varepsilon^*$, применим к преобразованным данным обычный метод наименьших квадратов и получим модель вида: $y^* = \beta_1 + \alpha \cdot x_1^* + \beta_2 \cdot x_2^* + \ldots + \beta_n \cdot x_n^* + \varepsilon^*$.

В ней остатки не зависят от x_1, а значит имеют постоянную дисперсию, т.е. модель гомоскедастична. В этой модели β_1 играет роль постоянного числа. Оценки коэффициентов $\alpha, \beta_1, \beta_2, \ldots \beta_n$ становятся точнее и, как правило, ниже. Для того чтобы использовать эту модель для дальнейшего анализа необходимо вернуться к прежним переменным.

Если в уравнении регрессии присутствуют несколько гетероскедастичных переменных, то в качестве «весов» можно использовать либо их линейную комбинацию, либо наиболее подходящую, исходя из графического представления.

Для применения взвешенного метода наименьших квадратов в программе *Statistica* достаточно на стартовой панели множественной линейной регрессии нажать на клавишу *Select cases* и в открывшемся окне задать переменную, по которой необходимо взвесить наблюдения.

Задания

1. Создайте файл GDP. sta:

EE- государственные расходы на образование;

GDP – валовой внутренний продукт;

P – численность населения.

Таблица 5

EE	GDP	P
0,34	5,67	0,36
0,22	10,13	2,9
0,32	11,34	2,39
1,23	18,88	3,44
1,81	20,94	3,87
1,02	22,16	10,71
1,27	23,83	3,1
1,07	24,67	9,93
0,67	27,56	5,07
1,25	27,57	11,1
0,75	40,15	9,6
2,8	51,62	4,78
4,9	57,71	4,09
3,5	63,03	22,34
4,45	66,32	5,12
1,6	66,97	44,92
4,26	76,88	7,51
5,31	101,65	6,37
6,4	115,97	8,37
7,15	119,49	9,86
11,22	124,15	8,31
8,66	140,98	14,62
5,56	153,85	27,06
13,41	169,38	14,14
5,46	186,33	67,4
4,79	211,78	37,43
8,92	249,72	123,03
18,9	261,41	23,94
15,95	395,52	57,04
29,9	534,97	55,95
33,59	655,29	53,71
38,62	815	61,56
61,61	1040,45	116,78
181,3	2586,4	227,64

2. Постройте модель линейной множественной регрессии EE(GDP, P). Запишите полученную модель в тетрадь. Просмотрите остатки ε_i в модели. В рабочем документе (файл GDP) добавьте переменную RES, которая рассчитывает остатки как разность между реальным значением EE и ее прогнозным значением, полученным в модели.

Сравните ε_i и RES.

3. Постройте графики зависимостей $(RES)^2$ от GDP и $(RES)^2$ от P. Сделайте предположение о наличии гетероскедастичности в модели для каждой переменной.

4. Проверьте наличие гетероскедастичности в модели по каждой переменной используя тесты Уайта и Голдфелда – Квандта.

5. Устраните гетероскедастичность в модели используя взвешенный метод наименьших квадратов.

5. Регрессионные модели с переменной структурой

Цель: сформировать знания о фиктивных переменных и особенностях их использования в регрессионном анализе.

Основные вопросы:

1. Задание фиктивных переменных: текстовые надписи и числовые значения.

2. Выявление структурного сдвига модели, тест Чоу.

3. Регрессионные модели с фиктивными переменными.

Указания по выполнению лабораторной работы:

При более детальном изучении модели на этапе спецификации требуется определить, совпадают ли уравнения регрессии для отдельных групп наблюдений. Распространенным тестом для поверки данной гипотезы является тест Чоу, суть которого состоит в следующем:

Пусть имеются две выборки объёма n_1 и n_2. Для каждой из этих выборок оценено уравнение вида:

$y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n$ - для n_1

$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n$ - для n_2

Проверяется нулевая гипотеза о равенстве друг другу соответствующих коэффициентов регрессий, т.е. другими словами будет ли уравнение регрессии одним и тем же для обеих выборок.

$H_0 : \alpha_0 = \beta_0;\ \alpha_1 = \beta_1;\ \alpha_2 = \beta_2;\ \ldots;\ \alpha_n = \beta_n.$

$H_1 : \alpha_0 \neq \beta_0;\ \alpha_1 \neq \beta_1;\ \alpha_2 \neq \beta_2;\ \ldots;\ \alpha_n \neq \beta_n.$

Для проверки нулевой гипотезы используется F- статистика:

$$F_{набл.} = \frac{ESS_0 - ESS_1 - ESS_2}{ESS_1 + ESS_2} \cdot \frac{n_1 + n_2 - 2k - 2}{k + 1}$$, где

ESS_1, ESS_2 - доли дисперсии необъясненные регрессиями , построенными для выборок объёмов n_1 и n_2 соответственно;

k – количество независимых переменных, входящих в уравнение.

ESS_0 - доля дисперсии необъясненная регрессией, построенной для объединенной выборки объёма ($n_1 + n_2$).

По таблице Фишера определяется $F_{крит.}$, которое имеет следующие степени свободы $v_1=k+1$ $v_2=n_1+n_2 - 2k - 2$.

Строим доверительные интервалы и выделяем область принятия нулевой гипотезы при заданном уровне значимости α.

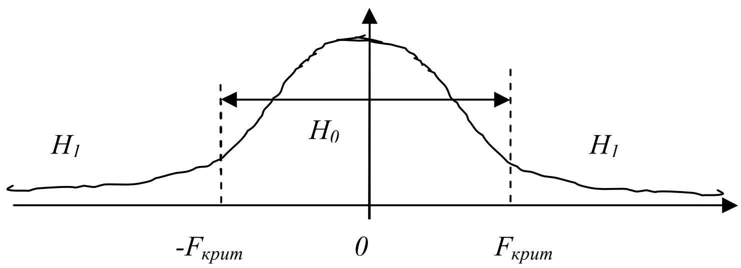
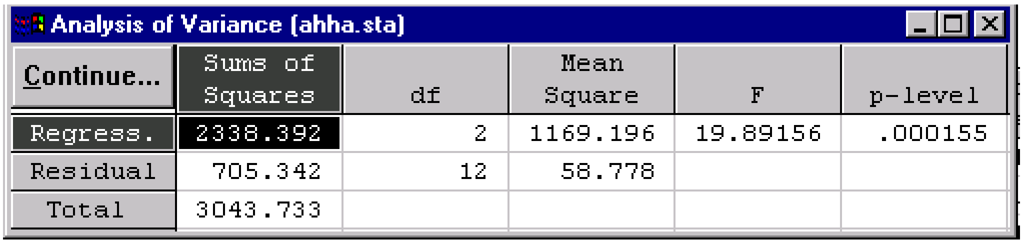


Рис.11

Если $F_{набл.}$ близка к нулю, то это означает, что коэффициенты регрессии совпадают и нецелесообразно рассматривать две различные регрессии для двух групп наблюдений. В противном случае, говорят о наличии структурного сдвига в модели и используют для прогнозирования модели для подвыборок.

Для отбора наблюдений в программе *Statistica* используется «*Select Cases*», которая появляется при открытии стартовой панели для построения множественной регрессии. Например, нужно построить регрессию для 1 – комнатных квартир в этом случае отбор переменных будет происходить при условии *Kol=1*.

Для определения значений ESS_i в диалоговом окне *Multuple Regression Result* нажмите клавишу *Analisis of Variance* , появится таблица первый столбец которой состоит из сумм квадратов отклонений:


Analysis of Variance (ahha.sta)

Continue...	Sums of Squares	df	Mean Square	F	p-level
Regress.	2338.392	2	1169.196	19.89156	.000155
Residual	705.342	12	58.778		
Total	3043.733				

Рис.12.

Regression Sums of Squares (RSS) – объясненная регрессией часть дисперсии;
Residual Sums of Squares (ESS) – необъясненная регрессией часть дисперсии;
Total Sums of Squares (TSS) – общая дисперсия.

Задания:

1. Имеется предположение о наличии структурного сдвига ценообразования для 1-комнатных, 2-комнатных, 3-комнатных и 4- комнатных квартир. Запишите в тетради сущность теста Чоу для проверки этого предположения.

2. Создайте файл home.sta со следующим содержанием:

PRICE – цена квартиры,

SO – общая площадь,

R – расстояние до центра,

Y – возраст дома,

KOL – количество комнат

Таблица 6

PRICE	SO	R	Y	KOL
450	35	3,5	17	1
450	29	0,5	36	1
360	31	1	30	1
510	41	1	7	1
500	32	1,5	9	1
370	28	1,7	40	1
470	36	3,2	15	1
320	32	6	33	1
539	34	1,5	6	1
360	27	0	32	1
700	50	3	22	2
570	45	0,8	30	2

530	54	0	27	2
650	54	1,2	9	2
550	53	3	15	2
700	48	0,6	25	2
650	55	0,9	25	2
530	42	0	30	2
570	46	0,6	30	2
550	42	0	30	2
550	45	0	25	2
550	52	3	25	2
999	74	3	0	2
600	62	8	2	2
600	47	3	9	2
650	53	1,3	25	2
200	41	16	25	2
750	54	0	14	2
550	47	0,6	25	2
1380	102	0	9	2
620	60	3	40	3
930	65	0	12	3
960	71	0,7	14	3
1560	130	0	0	3
650	60	3,5	9	3
650	50	3	30	3
780	64	0	14	3
600	62	0,8	40	3
700	58	0	16	3
750	71	0,6	25	3
820	86	3	23	4
700	72	3,2	18	4
620	86	8	8	4
950	87	0,4	19	4
910	80	0	22	4
800	71	3	16	4
670	61	0,8	28	4
750	85	2,8	17	4
775	70	2,3	20	4

3. Постройте отдельно регрессии зависимости цены (PRICE) от переменных SO, R, Y для 1-комнатных. 2 – комнатных, 3,4 – комнатных квартир. Осуществите тест Чоу.

6. Нелинейные модели регрессии и их линеаризация

Цель: сформировать умения строить нелинейные регрессионные модели и производить линеаризацию моделей.

Основные вопросы:

1. Регрессии нелинейные по переменным.

2. Стандартные уравнения подгона: линейная, линейно-логарифмическая, лог-линейная, двойная логарифмическая модели регрессии.

3. Построение регрессий, нелинейных по параметрам, их линеаризация.

4. Степенная и экспоненциальная модели регрессии.

5. Спецификация модели по виду уравнения.

Указания по выполнению лабораторной работы

Для случая парной регрессии подбор модели обычно осуществляется по виду расположения наблюдаемых точек на корреляционном поле, с учетом фактов известных из экономической теории. В случае если зависимость может быть описана несколькими функциями, необходимо выбрать ту из них, которая обладает наилучшим качеством. Но следует помнить, что чем сложнее модель, тем менее интерпретируемы ее параметры.

Рассмотрим наиболее часто встречающиеся модели. Для простоты изложения и возможности графической интерпретации ограничимся моделями парной нелинейной регрессии.

При визуализации данных на корреляционном поле возможны следующие результаты (рис.13).

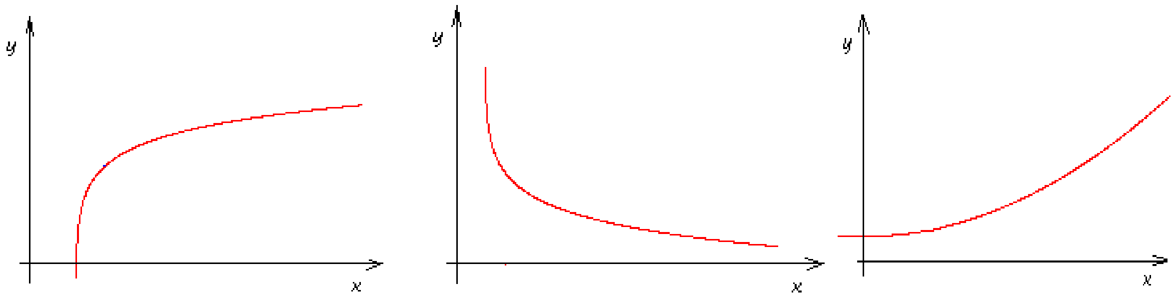


Рис. 13

1. Логарифмические модели.

– $\ln Y = \beta_0 + \beta \cdot \ln X + \varepsilon$ - двойная логарифмическая модель

коэффициент β в данной модели определяет эластичность переменной Y по переменной X, т.е. процентное изменение Y для данного процентного изменения X.

– $\ln y = a + b \cdot x + \varepsilon$ - лог-линейная модель, используется, например, при исследовании зависимости прироста объема выпуска от относительного увеличения затрат ресурса.

Коэффициент b в данной модели имеет смысл темпа прироста переменной у по переменной х , т.е. характеризует отношение относительного изменения у к абсолютному изменению х. Умножив b на 100%, получим процентный темп прироста переменной у.

– $y = a + b \cdot \ln x + \varepsilon$ - линейно-логарифмическая модель, используется, например, когда необходимо исследовать влияние процентного изменения независимой переменной на абсолютное изменение зависимой переменной.

Коэффициент b определяет изменение переменной у вследствии единичного относительного прироста х, например, если предположить, что у – валовой национальный продукт, а х – денежная масса, то b показывает, что с увеличением предложения денег на 1 % ВНП в среднем вырастет на b единиц.

2. Гиперболическая модель

$y = a + b \cdot \frac{1}{x} + \varepsilon$ - применяется в тех случаях, когда неограниченное увеличение значений объясняющей переменной х асимптотически приближает зависимую переменную у к некоторому пределу a. Подобная регрессия может отражать зависимости между объемом выпуска (х) и средними фиксированными издержками (у), между доходом (х) и спросом (у) на товары первой необходимости или предметы относительной роскоши (функция Торнквиста), между уровнем безработицы (х) и изменением заработной платы (у) и др.

3. Полиномиальная модель.

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \ldots + \beta_k X^k$$

При k = 2 квадратичная функция может отражать зависимость между расходами на рекламу (х) и прибылью (у), но большое применение имеет при анализе временных рядов; при k = 3 кубическая функция моделирует зависимость общих издержек (у) от объема выпуска (х).

4. Степенная модель.

$y = ax^b$ - отражает, например, зависимость спроса y на благо от его цены или от дохода x. Данная регрессия, путем математических преобразований сводится к двойной логарифмической модели. Коэффициент b является коэффициентом эластичности переменной y по переменной x.

5. Показательная модель.

$y = a \cdot e^{b \cdot x}$ - используется чаще всего в той ситуации, когда анализируется изменение переменной y с постоянным темпом прироста во времени. Переменная x заменяется на t, а модель используется при анализе временных рядов.

Коэффициент b показывает постоянный темп прироста переменной y во времени.

Показательная модель, путем логарифмирования сводится к лог-линейной модели.

Линеаризируем функцию путем математических преобразований (логарифмирования), например для модели $y = ax^b$ линеаризация выглядит следующим образом: $\ln y = \ln ax^b$, $\ln y = \ln a + \ln x^b$, $\ln y = \ln a + b \ln x$.

Обозначим $Y = \ln y$, $X = \ln x$, $A = \ln a$, получим $Y = A + bX$.

Задания

1. Создать файл с данными (табл.7).

Таблица 7

y	x
5	0,8
7	1
13	1,8
15	2,5
20	4
25	5,7
22	7,5
20	8,3
27	9,8
26	7

2. Изобразить корреляционное поле данных.

3. Оценить параметры нелинейных парных регрессий вида: $y = a \cdot e^{b \cdot x}$, $y = ax^b$, $y = a + b \cdot \frac{1}{x}$, $y = a + b \cdot \ln x$, $\ln y = a + b \cdot x$. Проинтерпретировать коэффициенты регрессии (если возможно).

4.Оценить качество модели. Изобразить график функции регрессии на корреляционном поле.

7. Системы одновременных регрессионных уравнений

Цель: сформировать знания о системах эконометрических уравнениях и особенностях оценивания параметров структурной формы модели с помощью косвенного, двухшагового и трехшагового методов наименьших квадратов.

Основные вопросы:

1. Системы независимых уравнений.

2. Структурная и приведенная формы эконометрической модели.

3. Проблема идентификации при переходе приведенной формы к структурной форме модели.

4. Оценивание параметров структурной модели. Косвенный, двухшаговый и трехшаговый методы наименьших квадратов.

Указания к выполнению лабораторной работы

Некоторые экономические процессы моделируются не одним, а несколькими уравнениями, содержащими как повторяющиеся, так и собственные переменные.

В силу этого возникает необходимость использования систем уравнений. При рассмотрении систем уравнений переменные делятся на два больших класса. Эндогенные и экзогенные переменные

Эндогенные переменные – это переменные, значения которых определяются внутри модели.

Экзогенные переменные – это внешние по отношению к модели переменные. Их значения определяются вне модели и поэтому они считаются фиксированными.

Уравнения, составляющие исходную модель называют структурными уравнениями модели.

Уравнения, в которых эндогенные переменные выражены только через экзогенные называют приведенными уравнениями.

При решении систем уравнений могут возникнуть проблемы идентификации. Под проблемой идентификации понимается возможность численной оценки параметров структурных уравнений по оценкам коэффициентов приведенных уравнений.

Исходную систему уравнений называют идентифицируемой, если по коэффициентам приведенных уравнений можно однозначно определить значения коэффициентов структурных уравнений.

Исходную систему уравнений называют сверхидентифицируемой, если по коэффициентам приведенных уравнений можно получить несколько вариантов значений коэффициентов структурных уравнений.

Исходную систему уравнений называют неидентифицируемой, если по коэффициентам приведенных уравнений невозможно определить значения коэффициентов структурных уравнений.

Необходимые условия идентифицируемости.

Пусть система уравнений включает в себя N - эндогенных и M – экзогенных переменных. Пусть количество эндогенных и экзогенных переменных в проверяемом уравнении равно n и m соответственно.

Переменные не входящие в данное уравнение, но входящие в другие уравнения системы, называют исключенными переменными. Их количество равно (N-n) - для эндогенных и (M-m) – для экзогенных переменных.

Первое необходимое условие идентифицируемости: уравнение идентифицируемо, если $(N-n)+(M-m) \geq N-1$;

Второе необходимое условие идентифицируемости: уравнение идентифицируемо, если $M-m \geq n-1$.

Знаки равенства соответствует точной идентификации уравнений, неравенство – сверхидентификации уравнения.

Необходимое и достаточное условие идентифицируемости.

В модели, содержащий N уравнений относительно N эндогенных переменных, условие идентифицируемости выполняется тогда и только тогда, когда ранг матрицы, составленной из исключенных из данных уравнений переменных, но входящих в другие уравнения системы, равен N-1.

Для оценки параметров точно идентифицируемых уравнений используют косвенный метод наименьших квадратов (КМНК), а для сверхидентифицируемых – двухшаговый метод наименьших квадратов (ДМНК).

Задания

1. Модель «спрос-предложение» имеет вид:

$$\begin{cases} q = \beta_0 + \beta_1 \cdot p + \varepsilon_1 \quad - \quad \textit{предложени е} \\ q = \alpha_0 + \alpha_1 \cdot p + \alpha_2 \cdot y + \varepsilon_2 \quad - \quad \textit{спрос} \end{cases},$$

где p – цена, q – количество товара, y – доход потребителя.

2. Определить оценки параметров идентифицируемых уравнений КМНК по следующим статистическим данным:

Таблица 8

p	n	2n	3n	4n	5n
q	8n	10n	7n	5n	n
y	2k	4k	3k	5k	2k

где n , к – параметр задаваемый преподавателем.

8. Временные ряды

Цель: Сформировать умения определять поведение временного ряда, выделять трендовую и сезонную компоненты ряда, строить прогнозы с помощью временного ряда при наличии аддитивной или мультипликативной сезонности.

Основные вопросы:

1. Автокорреляция уровней временного ряда
2. Аналитическое выравнивание временного ряда
3. График поведения временного ряда, уравнения тренда
4. Сезонные колебания временного ряда

Указания к выполнению лабораторной работы

Изучение динамики того или иного объекта, явления начинается с построения ряда динамики, или временного ряда. *Динамический ряд* – это таблица, в которой представлены значения показателя за последовательные периоды или на моменты времени. Каждое значение показателя *называется уровнем ряда*.

Динамический ряд является *интервальным*, если каждый уровень представляет собой итог процесса за некоторый интервал времени. Динамический ряд является *моментным*, если уровни отражают состояние объекта в отдельные моменты времени.

Ряд динамики состоит из двух строк или столбцов: промежутков или моментов времени, к которым относятся уровни, и самих уровней признака (показателя). Ряд, в котором время задано в виде промежутков – лет, месяцев, суток, называется *интервальным динамическим рядом.* Ряд, в котором время задано в виде конкретных дат (моментов времени), называется *моментным динамическим рядом.*

При статистическом изучении динамики необходимо четко разделить два её основных элемента – тенденцию и колеблемость, чтобы дать каждому из них количественную характеристику с помощью специальных показателей. Понятие об уравнении тенденции динамики было введено в статистику английским ученым Гукером в 1902 г. Он предложил называть такое уравнение трендом (анг. the trend – направление, тенденция). Для того чтобы нагляднее представить показатели, характеризующие тенденцию, следует абстрагироваться от колеблемости и выявить динамический ряд в форме «чистого» тренда при отсутствии колебаний.

Абсолютный прирост – это разность между сравниваемым уровнем и уровнем более раннего периода, принятым за базу сравнения. Если эта база – непосредственно предыдущий уровень, показатель называют *цепным*, если за базу взят, например, начальный уровень, показатель называют базисным. Формулы абсолютного прироста: цепное: $\Delta_i = y_i - y_{i-1}$; базисное: $\Delta_i = y_i - y_0$.

Абсолютный прирост не является константой тенденции. Оно со временем возрастает, т.е. уровни ряда изменяются с ускорением. *Ускорение* – это разность между абсолютным приростом за данный период и абсолютным приростом за предыдущий период одинаковой длительности:

$$\Delta_i^{/} = \Delta_i - \Delta_{i-1}$$

Показатель ускорения применяется только в цепном варианте, но не в базисном. Отрицательная величина ускорения говорит о замедлении роста или об ускорении снижения уровней ряда.

Ускорение является константой тенденции ряда, что свидетельствует о параболической форме этой тенденции. Её уравнение имеет вид:

$$y_i = a + b \cdot t_i + c \cdot t_i^2,$$

где a – уровень ряда в начальный период при t=0;

b – средний по ряду абсолютный прирост;

t_i – номер периода;

c – половина ускорения.

Система показателей должна содержать не только абсолютные, но и относительные статистические показатели.

Относительные показатели динамики необходимы для сравнения разных объектов, особенно если их абсолютные характеристики различны.

Темп роста – это отношение сравниваемого уровня (более позднего) к уровню принятому за базу сравнения (более раннему). Темп роста исчисляется в цепном варианте к предыдущему уровню или в базисном варианте – к одному и тому же, обычно начальному уровню. Он говорит о том, сколько процентов составляет сравниваемый уровень по отношению к уровню, принятому за базу, или во сколько раз сравниваемый уровень больше уровня, принятого за базу.

Рассмотрим связь абсолютных и относительных показателей динамики. Обозначим темп изменения через k. Тогда имеем: цепной темп роста: $k_{\frac{i}{i-1}} = \frac{y_i}{y_{i-1}}$; базисный темп роста: $k_{\frac{i}{0}} = \frac{y_i}{y_0}$.

Если сравниваемый уровень y выразить через уровень предыдущего года плюс прирост или через уровень базисного года плюс базисный абсолютный прирост, получим:

$$k_{\frac{i}{i-1}} = \frac{y_{i-1} + \Delta_i}{y_{i-1}} = 1 + \frac{\Delta_i}{y_{i-1}} \quad \textit{или} \quad 100\% + \frac{\Delta_i}{y_{i-1}} \cdot 100;$$

$$k_{\frac{i}{0}} = \frac{y_0 + \Delta_{0i}}{y_0} = 1 + \frac{\Delta_{0i}}{y_0} \quad \textit{или} \quad 100\% + \frac{\Delta_{0i}}{y_0} \cdot 100.$$

Величину $\frac{\Delta_i}{y_{i-1}}$ *или* $\frac{\Delta_{0i}}{y_0}$, т.е. отношение абсолютного изменения к предыдущему или базисному уровню, часто называют *относительным приростом* (относительным изменением) или же *темпом прироста*. Он равен k-1 или k-100%. Темп прироста может иметь как положительные, так и отрицательные значения. Например, финансовый результат от реализации продукции предприятием может быть прибылью (+), а может быть убытком (-), тогда темп изменения и темп прироста применять нельзя.

В этом случае такие показатели теряют смысл и не имеют экономической интерпретации. Сохраняют смысл только абсолютные показатели динамики.

Для временных рядов гетероскедастичность проявляется в виде автокорреляции остатков. Автокорреляция может быть положительной и отрицательной. Чаще всего положительная автокорреляция вызывается направленным воздействием некоторых неучтенных в модели факторов.

Например, пусть исследуется спрос y на прохладительные напитки в зависимости от дохода x по ежемесячным данным. Зависимость, отражающая увеличение спроса с ростом дохода, может быть представлена линейной функцией. $y = \alpha + \beta \cdot x$

Однако фактические точки наблюдений обычно будут превышать линию графика в летние периоды и будут ниже ее в зимние.

Отрицательная автокорреляция означает, что за положительным отклонением следует отрицательное и наоборот. На практике отрицательная автокорреляция встречается редко.

Наиболее известным критерием обнаружения автокорреляции является критерий Дарбина – Уотсона, общая схема применения которого следующая:

Для построенного уравнения регрессии определяют значение статистики Дарбина – Уотсона DW. По таблице критических точек распределения для заданного уровня значимости α, числа наблюдений n и количества независимых переменных k определяются два значения d_l - нижняя граница, d_u – верхняя граница.

Далее осуществляем выводы по правилу:

$0<DW<d_l$ – существует положительная автокорреляция;

$d_l<DW<d_u$ – зона неопределенности;

$d_u<DW<4-d_u$ – автокорреляция отсутствует;

$4-d_u<DW<4-d_l$ – зона неопределенности;

$4-d_l<DW<4$ – существует отрицательная автокорреляция.

Графически выводы можно представить следующим образом:

Рис. 14

Для устранения автокорреляции необходимо, прежде всего, скорректировать саму модель. Возможно, автокорреляция вызвана отсутствием в модели некоторой важной объясняющей переменной. И добавление этой переменной поможет устранить автокорреляцию. Чаще всего приходится добавлять фиктивные переменные, отвечающие за сезонность.

Например, при изучении спроса на прохладительные напитки добавим переменную $s = \begin{cases} 0, & \textit{если холодное время года}, \\ 1, & \textit{если теплое время года}. \end{cases}$

Тогда y может быть представлено в виде: $y = \alpha + \beta_1 \cdot x + \beta_2 \cdot s$.

При s=0 y покажет объем продаж прохладительных напитков в холодное время года, а при s=1 – в теплое. Коэффициент β_2 показывает на сколько изменяется объем продаж в теплое время года по сравнению с холодным.

Иногда приходится добавлять несколько фиктивных переменных.

Для определения поведения временного ряда строят корреляционное поле переменной у в зависимости от времени.

а)

б)

Рис.15

Если амплитуда изменения переменной у с течением времени остается постоянной (рис. 15а), можно предположить наличие аддитивной сезонности. В этом случае значение у в каждый момент времени найдется из условия: y=T+S, где Т – тренд, S – сезонная компонента.

Если амплитуда изменения переменной у с течением времени увеличивается (рис. 15 б), то можно предположить наличие мультипликативной сезонности. В этом случае значение объема экспорта в каждый момент времени найдется из условия: y=TS, где Т – тренд, S – сезонная компонента.

Для оценки сезонной компоненты и определения уравнения тренда проводят центрирование данных, методом скользящей средней. Скользящая средняя определяется как среднее значение между двумя соседними значениями $y_{ср.}$

$$y_{ск\,ср} = \frac{(y_{ср})_{i-1} + (y_{ср})_i}{2}$$

Разница между соответствующими значениями у и значениями скользящей средней определяет оценку сезонной компоненты.

Для определения значений сезонной компоненты в каждом квартале необходимо среднюю оценку сезонной компоненты за год вычесть из соответствующих средних за каждый квартал.

Уравнение линейного тренда временного ряда записывается в виде:

$y_t = \Delta y(абс) + y_{t-1}$ или $y_t = y_0 + \Delta y(абс) \cdot t$, где $\Delta y(абс) = y_i - y_{i-1}$ - абсолютный прирост переменной y.

Задания

1. Создайте файл Avto.sta;

2. Постройте модель зависимости

3. Проверьте наличие автокорреляции в модели;

4. Добавьте 4 фиктивных переменных, отвечающих за времена года (зима, весна, лето, осень), постройте новую модель, включив в нее фиктивные переменные, проверьте наличие автокорреляции в этой модели.

5. Дайте экономическую интерпретацию всем коэффициентам модели.

6. Поквартальные данные о прибыли некоторого предприятия в течение 6 лет представлены в таблице. Создать соответствующий файл (табл.9) и спрогнозировать прибыль предприятия на 7 год поквартально.

Таблица 9

Год	1				2				3				4				5				6			
Квартал	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
прибыль	11,86	14,34	15,22	14,53	10,84	13,78	16,41	15,68	12,19	15,65	17,69	16,37	13,21	16,4	17,89	16,85	14,08	16,3	18,31	17,7	14,73	17,74	19,44	18,23

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Артамонов Н.В. Введение в эконометрику: курс лекций / Н.В. Артамонов. – М.: МЦНМО, 2011. – 204 с.

2. Балдин, К. В. Математические методы и модели в экономике: учеб. / К. В. Балдин, В. Н. Башлыков, А. В. Рокосуев ; ред. К. В. Балдин. – М. : Флинта, 2012. – 328 с.

3. Валентинов В.А. Эконометрика: учеб. / В.А.Валентинов. – 2-е изд.– М: Дашков и К°, 2012. – 448 с.

4. Ермолаев М. Б. Эконометрика: учеб. пособие: рек. УМО / М.Б. Ермолаев, Г. Г. Кадамцева, С. Б. Лапшинов. – Иваново : Институт бизнеса, информационных технологий и финансов, 2011. – 111 с.

5. Кремер Н.Ш. Эконометрика: учеб. /Н. Ш. Кремер, Б. А. Путко ; под ред. Н. Ш. Кремера. – 3-е изд., перераб. и доп. – М. : ЮНИТИ-ДАНА, 2012. – 328 с.

6. Мхитарян В. С.Эконометрика: учеб.-практ. пособие / В. С. Мхитарян, М. Ю. Архипова, В. П. Сиротин. - М.: Евразийский открытый ин-т, 2012. – 221 с.

7. Новиков А.И. Эконометрика: учеб. пособие /А.И. Новиков. – М.: Дашков и К°, 2013. – 224 с.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ 3

1. Линейная модель множественной регрессии 4

2. Показатели качества регрессии 9

3. Мультиколлинеарность 13

4. Линейные регрессионные модели с гетероскедастичными и автокоррелированными остатками 15

5. Регрессионные модели с переменной структурой 20

6. Нелинейные модели регрессии и их линеаризация 23

7. Системы одновременных регрессионных уравнений 27

8. Временные ряды 29

БИБЛИОГРАФИЧЕСКИЙ СПИСОК 36

Наталья Николаевна Двоерядкина,
канд. пед. наук, доц. кафедры общей математики и информатики АмГУ;

Татьяна Александровна Юрьева,
канд. пед. наук, доц. кафедры общей математики и информатики АмГУ;

Татьяна Евгеньевна Гришкина,
старший преподаватель кафедры общей математики и информатики АмГУ

*Методические указания к лабораторным занятиям по дисциплине «**Эконометрика**»*

……………………………………………………………………………………

Изд-во АмГУ. Подписано к печати Формат Усл. печ. л. Тираж. Заказ. Отпечатано в типографии АмГУ