Министерство сельского хозяйства Российской Федерации

Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Саратовский государственный аграрный университет
имени Н. И. Вавилова»

ЭКОНОМЕТРИКА

Краткий курс лекций

для обучающихся 3 курса

Направление подготовки
38.03.02 Менеджмент

Направленность (профиль)
«Производственный менеджмент
агропромышленного комплекса»

Саратов 2017

УДК 330.45:519.862
ББК 65.05
Ш55

Рецензенты:

Эконометрика: краткий курс лекций для обучающихся направления подготовки «Менеджмент» / Сост.: В. А. Шибайкин // ФГБОУ ВО
Ш55 Саратовский ГАУ. – Саратов, 2017. – с.

Краткий курс лекций по дисциплине «Эконометрика» составлен в соответствие с рабочей программой дисциплины и предназначен для обучающихся направления подготовки «Менеджмент». Краткий курс лекций содержит теоретический материал по общим проблемам регрессионного анализа и вопросам применения математических методов при проведении эконометрического анализа. Рассмотрены примеры известных эконометрических моделей, направлен на формирование у обучающихся знаний об использовании основных эконометрических методов, на применение этих знаний для понимания процессов, происходящих в экономике, для применения эконометрических моделей в экономическом анализе.

УДК 330.45:519.862
ББК 65.05

Шибайкин В. А, 2017
© ФГБОУ ВО Саратовский ГАУ, 2017

Введение

В широком понимании эконометрика начинается с анализа связи между экономическими переменными (такими как количество и цена, доходы и расходы и т. д.) с помощью рассмотрения интересующего явления в абстрактном виде и изложения теории в математической форме. Цель эконометрики состоит в том, чтобы придать количественные меры взаимосвязям экономических отношений и процессов.

Эконометрика включает в себя элементы трёх наук – экономика, статистика, и математика. Используя все эти методы не трудно запутаться и запутать других. согласно Фришу: «сущность эконометрики составляет взаимное переплетение количественной экономической теории и статистических наблюдений» (Frisch, R. (1933) [ ].

Лекция 1

ВВЕДЕНИЕ В ЭКОНОМЕТРИКУ

1.1. Задачи изучения дисциплины

Эконометрика – это наука с помощью, которой количественно выражаются взаимосвязанные экономические явления и процессы. Цель эконометрики состоит в том, чтобы придать количественные меры взаимосвязям экономических отношений и процессов.

Эконометрика состоит из трёх элементов -- экономика, статистика, и математика:

- эконометрика определяет постановку задачи, а также интерпретирует полученный результат;
- статистика предоставляет необходимые данные для построения моделей;
- математика предоставляет методы необходимые для построения моделей.

Эконометрика имеет дело с конкретными экономическими данными и занимается количественным описанием конкретных взаимосвязей, то есть заменяет коэффициенты, представленные в общем виде в этих взаимосвязях, конкретными численными значениями.

С помощью эконометрики возможно построение экономических моделей, основываясь на экономической теории или на эмпирических данных, и определения возможность их использования для описания, анализа и прогнозирования реальных экономических процессов.

Эконометрика оформилась как наука в первой четверти прошлого века. В журнале «Эконометрика», основанном в 1933г. Р. Фришем, он дал следующее определение эконометрики: «Эконометрика - это не то же самое, что экономическая статистика. Она идентична и тому, что мы называем экономической теорией, хотя значительная часть этой теории носит количественный характер. Эконометрика не является синонимом приложений математики к экономике.

Свидетельством всемирного признания эконометрики является присуждение четырех нобелевских премий по экономике за разработки в этой области:

премия 1969г. была присуждена Р. Фришу и Я.Тинбергену за разработку математических методов анализа экономических процессов;

премия 1980г. - Л. Клейну за создание эконометрических моделей и их применение к анализу экономических колебаний и экономической политике;

премия 1989г. - Т. Хаавельмо за прояснение вероятностных основ эконометрики и анализ одновременных экономических структур;

премия 2000г. - Дж. Хекману за развитие теории и методов анализа селективных выборок и Д. Макфаддену за развитие теории и методов анализа моделей дискретного выбора.

Эконометрика имеет дело с конкретными экономическими данными и занимается количественным описанием конкретных взаимосвязей, то есть заменяет коэффициенты, представленные в общем виде в этих взаимосвязях, конкретными численными значениями.

С помощью эконометрики возможно построение экономических моделей, основываясь на экономической теории или на эмпирических данных, и определения возможность их использования для описания, анализа и прогнозирования реальных экономических процессов.

Например, микроэкономическая теория утверждает, что снижение цены товара приводит к увеличению спроса на данный товар (при условии неизменности всех остальных факторов), то есть устанавливается связь между спросом на товар и ценой на него. Однако микроэкономическая теория не дает количественных оценок данной связи, то есть не позволяет ответить на вопрос о том, насколько изменится спрос на данный товар в результате изменения его цены на определенную величину. Расчет количественных оценок и есть задача эконометрики.

В эконометрике основным методом исследования является корреляционно-регрессионный анализ. Поэтому в качестве этапов эконометрического исследования можно указать :

- постановку проблемы;
- получение данных, анализ их качества;
- выбор спецификации модели;
- отбор факторов;
- оценку параметров;
- проверка надёжности полученных параметров;
- интерпретацию результатов.

Этот список включает те стадии, которые проходит любое исследование, независимо от того, на использование каких данных оно ориентировано: пространственных или временных

1.2. Типы моделей в эконометрике

В эконометрике выделяют выделить три основных класса моделей, которые применяются для анализа и прогнозирования экономических систем:

- регрессионные модели с одним уравнением – используются изолированные уравнения линейного и нелинейного вида, с пространственными данными. В зависимости от количества включенных в модель факторов модели делятся на модель парной регрессии и модель множественной регрессии;

- модели временных рядов, использующие в качестве факторного признака моменты времени t и соответственно временные данные;

- системы эконометрических уравнений. Применяются если исследуемый процесс нельзя описать одним уравнением.

1.3. Используемые данные

Основной базой данных для эконометрических исследований служат данные официальной статистики либо данные бухгалтерского учета. Таким образом, проблемы экономического измерения – это проблемы статистики. Используя экономическую теорию, можно определить связь между признаками и показателями, а используя статистику – ответить на вопросы, связанные с конкретными значениями экономических показателей.

При моделировании экономических процессов используют два типа данных:

- пространственные данные;
- временные данные.

Пространственными данными является набор сведений по разным объектам, взятым за один и тот же период или момент времени. Например набор сведений по разным фирмам (объем производства, численность работников, размер основных производственных фондов, доход за определенный период и т.д.). Примером таких данных может служить данные об объеме, ценах потребления некоторого товара по потребителям.

Временными данными является набор сведений, характеризующих один и тот же объект, но за разные периоды или моменты времени. Примером временных данных могут служить ежемесячные или ежеквартальные данные о средней заработной плате, индексе потребительских цен, объеме выпуска или, например, ежедневный курс доллара или евро на бирже. Отличительной особенностью временных данных является то, что они естественным образом упорядочены по времени, кроме того, наблюдения в близкие моменты времени могут быть зависимы. Набор сведений представляет собой множество признаков, характеризующих объект исследования. Признаки являются взаимосвязанными, причем в этой взаимосвязи они могут выступать в одной из двух ролей:

- в роли результативного признака (аналог зависимой переменной y в математике);
- в роли факторного признака, значения которого определяют значения признака-результата (аналог независимой переменной x в математике).

В эконометрической модели результативный признак называют объясняемой переменной, а факторный признак называют объясняющей переменной.

1.3. Простые количественные взаимосвязи между перменными

Как уже упоминалось ранее, главным инструментом эконометрических исследований является эконометрическая модель. Математические модели широко применяются в бизнесе, в экономике, общественных науках и даже в политике. Их удобно использовать для более полного понимания сущности происходящих процессов, анализа данных, прогнозирования и т.д. Можно выделить три класса моделей, используемых в эконометрике:

1. Модель временных рядов. Модель, в которой результативный признак является функцией переменной времени или переменных, относящихся к другим моментам времени.

К моделям, представляющим собой зависимость результативного признака от времени, относятся следующие модели:

• тренда (тренд представляет собой устойчивое изменение уровня показателя результативного признака в течение длительного времени):

$y(t) = T(t) + \varepsilon_t$,где T(t)– временной тренд заданного параметрического вида (например, линейный),

•сезонности (ее характеризуют устойчивые внутригодовые колебания уровня показателя результата):

$y(t) = S(t) + \varepsilon_t$, где S_t – периодическая (сезонная) компонента,

• тренда и сезонности:

$y(t) = T(t) + S(t) + \varepsilon_t$ – аддитивная

$y(t) = T(t) \cdot S(t) + \varepsilon_t$ – мультипликативная,

где – временной тренд заданного параметрического вида, – периодическая (сезонная) компонента, – случайная (стохастическая) компонента.

К моделям временных данных, представляющих собой зависимость результативного признака от переменных, датированных другими моментами времени, относятся модели:

• объясняющие поведение результативного признака в зависимости от предыдущих значений факторных признаков (модели с распределенным лагом);

• объясняющие поведение результативного признака в зависимости от предыдущих значений результативных признаков (модели авторегрессии);

• объясняющие поведение результативного признака в зависимости от будущих значений факторных или результативных признаков (модели ожидания);

Общей чертой всех моделей этого класса является то, что они объясняют поведение временного ряда исходя только из его предыдущих и будущих значений.

2. Регрессионная модель с одним уравнением. В таких моделях результативный признак (зависимая или объясняемая переменная) представляется в виде функции факторных признаков (независимых или объясняющих переменных):

$f(x,\beta) = f(x_1,...,x_k,\beta_1,...,\beta_p)$, где $x_1,\ldots,x_k$ – факторные признаки, $\beta_1,\ldots\beta_p$– параметры при этих факторах.

В зависимости от количества факторных признаков различают парную и множественную регрессию, кроме того, в зависимости от вида функции объясняющей связь переменных модели подразделяются на линейные и нелинейные. Например, можно исследовать спрос на мороженое как функцию от времени, температуры воздуха, среднего дохода или зависимость зарплаты от возраста, квалификации, образования, пола и т.д.

Область применения моделей этого вида, даже самых простых, линейных, намного шире, чем моделей временных рядов. Именно проблемам таких моделей посвящены несколько глав этой работы. Также эти модели более широко освещаются и в другой литературе.

Ниже приведено несколько примеров регрессионных моделей с одним уравнением.

• функция цены: P=f(Q,Pk), где цена определенного товара P зависит от объёма его поставки Q и от цен конкурирующих товаров Pk;

• функция спроса: Qd=f(P,Pk,I), где величина спроса на определенный товар Qd зависит от цены данного товара P, от цен конкурирующих товаров Pk и от реальных доходов потребителей I;

• производственная функция: Q=f(L,K), представляет собой зависимость объёма производства определенного товара от производственных факторов – затрат капитала K и затрат труда L;

3. Системы взаимосвязанных уравнений.. Системы могут состоять из тождеств и регрессионных уравнений, каждое из которых может включать в себя не только объясняющие переменные, но и объясняемые переменные из других уравнений системы. Таким образом, мы имеем набор объясняемых переменных, связанных через уравнения системы. Уравнения системы могут быть либо тождествами, либо поведенческими уравнениями. Для тождеств характерно, что их вид и значения параметров известны. В поведенческих уравнениях значения параметров требуется оценить.

Примером системы одновременных уравнений является модель спроса и предложения, включающая 3 уравнения: где – предложение товара в момент времени t;

Q_t^d – спрос на товар в момент времени t;

P_t – цена товара в момент времени t;

P_{t-1} – цена товара в предыдущий момент момент времени t-1;

I_t – доход потребителей в момент времени t.

Данная модель «объясняет» две результативные переменные:

1) Q_t – объем спроса, равный объему предложения в момент времени t;

2) P_t цену товара в момент времени t.

$$\begin{cases} Q_t^s = a_0 + a_1 \cdot P_t + a_2 \cdot P_{t-1} & \text{уравнение предложения,} \\ Q_t^d = b_0 + b_1 \cdot P_t + b_2 \cdot I_t & \text{уравнение спроса,} \\ Q_t^s = Q_t^d & \text{тождество равновесия,} \end{cases}$$

Приведённые коэффициенты a и b называются структурными коэффициентами модели.

Вопросы для самоконтроля

1. Эконометрика. Понятие, цель и задачи эконометрики. Типы данных. Примеры типов данных.
2. Модели в эконометрике. Классы моделей.
3. Какие виды моделей используются в эконометрике? От чего зависит выбор модели?
4. Какие типы данных используются в эконометрике?
5. От чего зависит тип используемых данных?
6. Определите источники информации используются при сборе данных и основные проблемы возникают при сборе данных?
7. Сформулируйте основные направления поиска измерителя исследуемого признака.
8. Какие виды моделей используются в эконометрике? От чего зависит выбор модели?

СПИСОК ЛИТЕРАТУРЫ

Основная

1. *Елисеева, И. И* Эконометрика : учебник / И. И. Елисеева, С.В. Курышева, Ю. В. Нерадовская и др.] ; под ред. И. И. Елисеевой. - М. :Проспект, 2010. - 288 с.- ISBN 978-5-392-00922-076 с.

Дополнительная

1. *Валентинов, В. А.* Эконометрика: практикум -2-е изд. – М. : Дашков и К, 2009. - 436 с. – ISBN 978-5-394-00428-5

Лекция 2

МОДЕЛЬ РЕГРЕССИИ С ДВУМЯ ПЕРЕМЕННЫМИ

2.1. Концепция популярной регрессионной функции

С помощью эконометрики возможно построение экономических моделей, основываясь на экономической теории или на эмпирических данных, и определения возможность их использования для описания, анализа и прогнозирования реальных экономических процессов.

Рассмотрим приемы некоторых эконометрических моделей.

Исследование динамики социальных и экономических процессов выявило довольно сильную распространенность эффекта насыщения: выхода на асимптоту при достижении определенных значений показателей. В силу этого в эконометрике большое распространение получили так называемые кривые с насыщением. этому типу кривых относится кривая Гомперца — s-образная кривая, предложенная Б. Гомперцем (1799-1865), которая имеет вид:

$$y = K \cdot a^{bt}$$

где K, a, b – параметры; t – время (1,2,...).

Кривая Гомперца используется для аналитического выражения тенденции развития показателя во времени, имеющего ограничения на рост (рис. 2.1 а,б).

Если log a < 0, то верхний предел для показателя у равен параметру K, а нижний — 0. Если log a > 0, то кривая имеет лишь нижний предел, равный величине параметра K. (рис. 1. в, г).

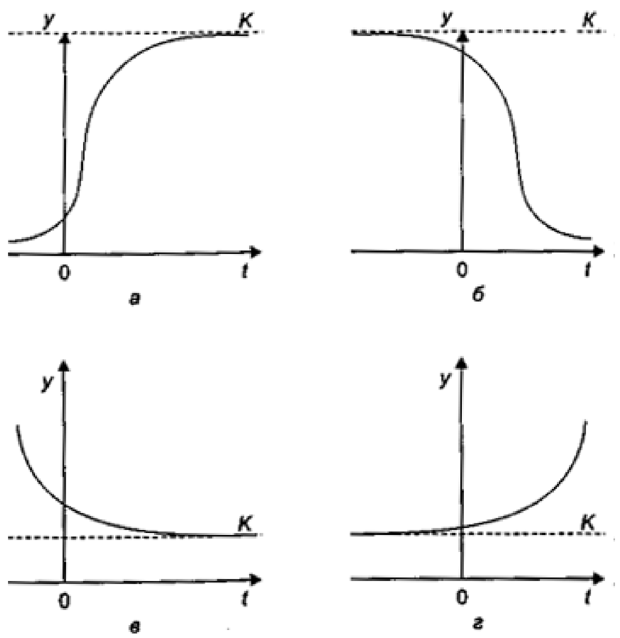

Рис. 2.1 Кривая Гомперца

Для определения параметров тренда a и b может использоваться метод наименьших квадратов, только если задан параметр К. В противном случае возможно лишь приближенное оценивание параметров. Кривая Гомперца применяется в демографических расчетах и страховом деле [ ].

Как уже отмечалось ранее, в эконометрике широко применяются методы статистики и математики в отношении экономических процессов. Для начала всегда строится математическая модель. Одна из наиболее простых моделей – нормальная простая (парная) регрессия. Хотя в чистом виде она встречается достаточно редко, её использование помогает понять суть процессов и исследовать их. Вообще изучение парной регрессии является, пожалуй, базовым во всем курсе эконометрики. Обычно парная регрессия используется в том случае, когда из всего круга факторов, влияющих на результат, можно выделить один, оказывающий наиболее сильное влияние. Он и используется в качестве объясняющей переменной. Представим, что у нас есть два ряда данных:

$$X=x_1,x_2,\ldots,x_n;$$

$$Y=y_1,y_2,\ldots,y_n;$$ где n – число наблюдений.

Каждое из наблюдений характеризуется двумя переменными xi, yi. В парной линейной регрессии связь между переменными определяется следующим образом:

$$y=\hat{y}(x)+\varepsilon=a+b\cdot x+\varepsilon,$$

где y – зависимая (объясняемая) переменная, реальная или фактическая или, как её ещё называют, эмпирическая, то есть, наблюдавшаяся в действительности;

x – независимая (экзогенная) переменная;

$\hat{y}(x)$ – зависимая переменная (эндогенная), рассчитанная с помощью уравнения регрессии, ещё её называют теоретической (в данном случае она рассчитывается по линейному уравнению регрессии);

a, b – константы, параметры уравнения линейной регрессии;

ε – случайная компонента, или возмущение.

Уравнение простой регрессии характеризует связь между двумя переменными, которая проявляется как некоторая закономерность лишь в среднем в целом по совокупности наблюдаемых данных. Например, если зависимость потребления электроэнергии y от объема выпускаемой продукции x можно представить в следующем виде $y=1500+24{,}8\cdot x$, то это означает, что при увеличении объема выпуска на 1 единицу потребление электроэнергии в среднем увеличивается на 24,8 единиц. То есть, в уравнении регрессии связь между результатом и фактором представляется, как бы в качестве функциональной связи. Причем представление функции может быть как линейным, так и нелинейным.

2.2. Подбор кривой

Самым простым способом определения вида связи, является построение поля корреляции. Каждую пару наблюдений (x_i,yi) можно отобразить в качестве точки на плоскости XY (рис. 2.2.). В этом случае наилучшей функцией считается та, чей график проходит через наибольшее количество точек или как можно ближе к ним.

В каждом из наблюдений величину случайной компоненты можно определить как разность между фактическим значением результата и рассчитанным по уравнению регрессии: $\varepsilon_i = y_i - \hat{y}_i$

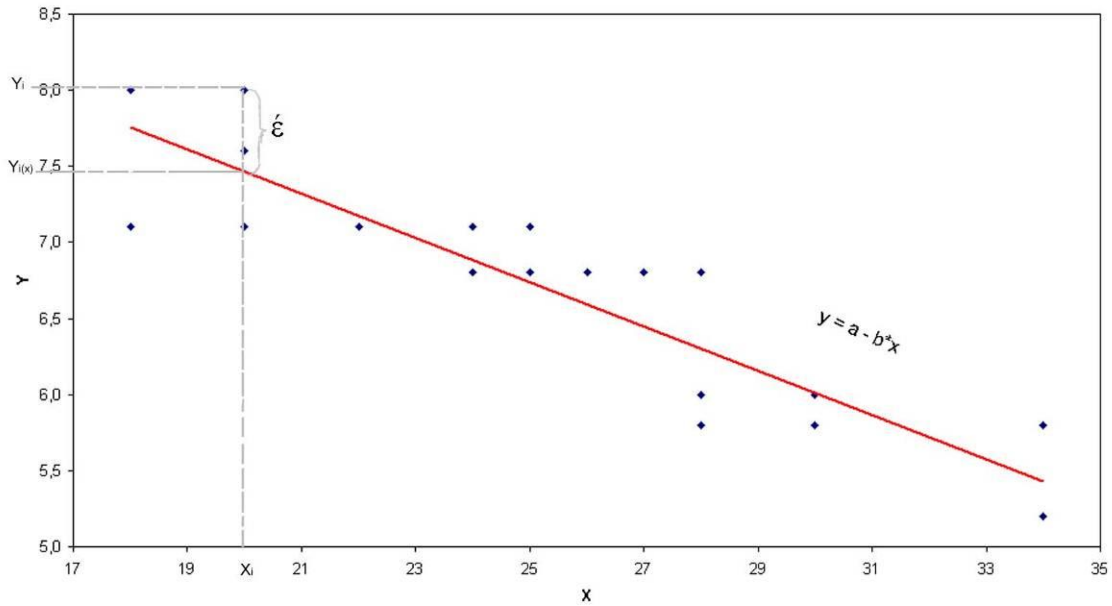

Рис. 2.2. Распределение значений X и Y на координатной плоскости

Если на графике все точки (x_i,y_i) совпадают с линией регрессии, тогда между результативным признаком Y и фактором X существует строгая функциональная связь и выполняется следующее равенство:

$\varepsilon_i = 0$ для каждого i=1,2,…,n.

В экономических процессах такое практически не встречается, экономические данные обычно не связаны строгой функциональной связью. Но во всех случаях, когда применение МНК оправдано, верно, что $\sum_i \varepsilon_i = 0$, поэтому в качестве меры отклонений используется сумма квадратов отклонений $\sum_i \varepsilon_i^2$.

Случайная компонента ε по своей сути есть случайная величина, также она называется возмущением. Она характеризует собой влияние не учтенных в модели факторов, каких-либо случайных влияний, неправильный выбор специфики модели и, в некоторых случаях, может быть связана с особенностями измерения. Как уже было сказано, данные, который характеризуют экономический процесс, не могут быть связаны строгой функциональной связью, кроме того, в эконометрическом исследовании используется случайная выборка данных, что и обуславливает постоянное наличие случайной компоненты. Размер остаточной дисперсии Dост также может зависеть от выбранного вида уравнения регрессии. Соответственно, чем она меньше, тем лучше подобрана функция.

При эконометрических изысканиях часто встречаются ошибки различного рода. Их принято делить на 3 вида:

1. Ошибки спецификации. От правильно выбранной спецификации модели зависит величина случайных ошибок: чем ближе рассчитанные с помощью уравнения регрессии значения результативного признака $\hat{y}_x$ подходят к фактическим данным y.

К ошибкам спецификации относятся не только неправильный выбор той или иной математической функции для расчета $\hat{y}_x$, но и отсутствие в построенном уравнении регрессии какого-либо существенного фактора, то есть, использование парной регрессии вместо множественной или двухфакторной вместо трехфакторной. Например, ценность работника зависит не только от его квалификации, здесь влияние могут оказывать и такие факторы как возраст, пол и т.д. Ошибки спецификации – это ошибки, которые могут быть сведены исследователем к минимуму с помощью изменения вида уравнения регрессии, заменой линейной связи на нелинейную, как лучше подходящую (рис.2.3).

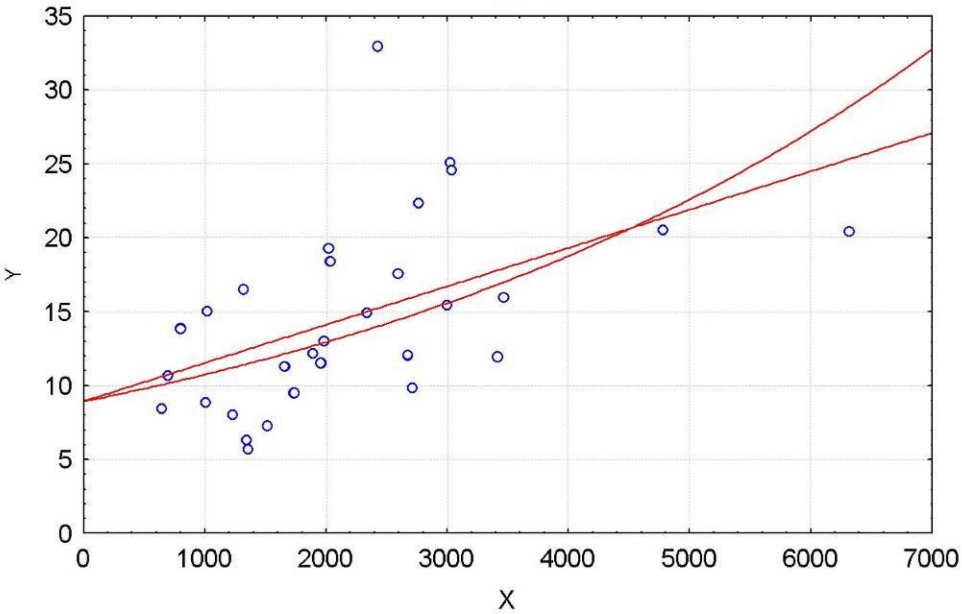

Рис. 2.3. Подбор функций графически линейная и нелинейная

В парной регрессии выбор вида математической функции $\hat{y}_x$=f(x) может быть осуществлен тремя методами:

•графическим;

•аналитическим, то есть исходя из теории изучаемой взаимосвязи;

•экспериментальным.

При изучении зависимости между двумя признаками графический метод подбора вида уравнения регрессии достаточно нагляден. Он основан на поле корреляции. Просто подбирается функция, график которой проходит через наибольшее количество точек или как можно ближе к ним. Существует набор функций, используемых наиболее часто

Аналитический метод подразумевает хорошей знание экономики и основан на изучении материальной природы связи исследуемых признаков, теоретических соображениях и опыте подобных предыдущих исследований.

Экспериментальный метод основан на переборе нескольких функций., то есть для каждой строится уравнение регрессии, для каждой оценивается качество регрессии и по этим показателям выбирается лучшая. Наиболее часто этот метод используется при обработке данных на компьютере. Большинство статистических пактом включает в себе автоматический перебор наиболее распространенных функций, а также пользователю предоставляется возможность добавлять дополнительные интересующие его функции. Основным показателем при переборе функций является остаточная дисперсия $D_{ост}$. Необходимо заметить, что нельзя сравнивать $D_{ост}$ по разным данным, сравнению поддаются только остаточные дисперсии по одним и тем же данным, но рассчитанные по разным уравнениям регрессии. Если для нескольких функций $D_{ост}$ одинаковы, то предпочтение отдается наиболее простому по виду уравнению регрессии, так как они более удобны для понимания, для них требуется меньшее количество наблюдений и расчеты, связанные с оцениванием качества регрессии, более просты.

Результаты многих исследований подтверждают, что число наблюдений должно в 6-7 раз превышать число рассчитываемых параметров при переменной x. Это означает, что искать линейную регрессию, имея менее 7 наблюдений, вообще не имеет смысла. Если вид функции усложняется, то требуется увеличение объема наблюдений, ибо каждый параметр при x должен рассчитываться хотя бы по 7 наблюдениям. Значит, если мы выбираем параболу второй степени $\hat{y}_x = a + b \cdot x + c \cdot x^2$, то требуется объем информации уже не менее 14 наблюдений. Учитывая, что эконометрические модели часто строятся по данным рядов динамики, ограниченным по протяженности (10, 20, 30 лет), при выборе спецификации модели предпочтительна модель с меньшим числом параметров при x [ ].

2.3. Метод наименьших квадратов

После того как выбрана модель, определен её вид, следующим шагом является оценивание параметров модели. Для линейной парной регрессии вида $y = \hat{y}(x) + \varepsilon = a + b \cdot x + \varepsilon$ необходимо оценить (найти фактическое значение) свободный член уравнения регрессии (константу) a и коэффициент регрессии b.

Для определения параметров модели можно использовать следующие критерии:

1) Сумму квадратов отклонений фактических значений результата y от рассчитанных с помощью уравнения регрессии $\hat{y}_x$:

$$S = \sum (y - \hat{y}_x)^2$$

Метод, использующий эту сумму, называется методом наименьших квадратов и является одним из основных методов эконометрики. В дальнейшем мы будем кратко его называть МНК.

2) Сумму модулей отклонений наблюдаемых значений зависимой переменной от её расчетных величин:

$$S = \sum |y - \hat{y}_x|$$

3) Отклонения, включаемые в сумму с определенной мерой:

$S = \sum g \cdot (y - \hat{y}_x)$, где g – мера, с которой отклонение для i-ого наблюдения входит в функционал.

При использовании МНК оптимальными будут значения параметров регрессии, минимизирующие функционал S.

Для оценки параметров модели линейной регрессии наиболее часто используется МНК, согласно которому в качестве оценок параметров принимают величины a и b, минимизирующие сумму квадратов отклонений наблюдаемых значений результативного признака y от расчетных (теоретических).

$$S=\sum(y-\hat{y}_x)^2=\sum(y-(a+b\cdot x))^2\rightarrow\min$$

Значения рядов наблюдений x и y нам известны. В уравнении необходимо найти два параметра a и b. Поэтому записывают частные производные данной функции по параметрам a и b и приравнивают их к нулю.

$$\frac{\partial S}{\partial a}=-2\cdot\sum_{i=1}^{n}(y_i-a-b\cdot x_i)=0,$$

$$\frac{\partial S}{\partial b}=-2\cdot\sum_{i=1}^{n}x_i\cdot(y_i-a-b\cdot x_i)=0,$$

Из этих равенств мы получим систему нормальных уравнений для оценки параметров a и b:

$$\begin{cases}\sum y=a\cdot n+b\cdot\sum x\\ \sum y\cdot x=a\cdot\sum x+b\cdot\sum x^2\end{cases}$$

Решая данную систему , найдем оценки параметров регрессии:

$$b=\frac{n\cdot\sum x\cdot y-\sum x\cdot\sum y}{n\cdot\sum x^2-(\sum x)^2}=\frac{\sum x\cdot y-n\cdot\bar{x}\cdot\bar{y}}{\sum x^2-n\cdot(\bar{x})^2}=\frac{\overline{x\cdot y}-\bar{x}\cdot\bar{y}}{\sigma_x^2}=\frac{cov(x,y)}{\sigma_x^2}.$$

Так как $cov(x,y)=\overline{x\cdot y}-\bar{x}\cdot\bar{y}$ – ковариация признаков;

$\sigma_x^2=\overline{x^2}-\bar{x}^2$ – дисперсия признака x, то b можно представить в несколько ином виде: $b=\dfrac{\overline{x\cdot y}-\bar{x}\cdot\bar{y}}{\overline{x^2}-\bar{x}^2}$.

Параметр b называют коэффициентом регрессии. Его величина показывает, насколько в среднем изменяется значение результативного признака при изменении факторного на одну единицу. Значения параметра b не имеют ограничений. Если коэффициент регрессии больше 0, то при увеличении фактора результат увеличивается и линия регрессии . Если же коэффициент регрессии меньше нуля, то при возрастании фактора результат уменьшается, в этом случае линия регрессии имеет отрицательный наклон.

Параметр a оценивается по следующей формуле:

$$a=\bar{y}-b\cdot\bar{x}$$

Правильность расчета параметров уравнения регрессии может быть проверена с помощью сравнения сумм $\sum y=\sum\hat{y}_x$. На практике из-за округления при расчетах возможно некоторое расхождение.

2.4. Коэффициент аппроксимации

Практически всегда фактические значения результативного признака отличаются от теоретических, рассчитанных по уравнению регрессии. Чем меньше это отличие, тем ближе теоретические значения подходят к эмпирическим, тем лучше подобрано уравнение регрессии. Величина отклонений фактических и расчетных значений результативного признака $(y_i - \hat{y}_x)$ по каждому наблюдению представляет собой ошибку аппроксимации. Число ошибок соответствует размеру совокупности. В отдельных случаях ошибка аппроксимации может оказаться равной нулю (когда в одном наблюдении фактическое и теоретическое значения результата совпадают). Отклонения $(y_i - \hat{y}_x)$ несравнимы между собой, исключая величину, равную нулю. Так, если для совокупности данных (или для одних и тех же наблюдений, но для разных моделей) $(y_i - \hat{y}_x)=7$, а для другой она равна 14, то это не означает, что во втором случае модель дает вдвое худший результат. Отклонения можно рассматривать как абсолютную ошибку аппроксимации.

Для сравнения используются величины отклонений, выраженные в процентах к фактическим значениям. Так, если для первого наблюдения y = 20, а для второго y = 50, ошибка аппроксимации составит 35 % для первого наблюдения и 28 % – для второго.

Поскольку $(y_i - \hat{y}_x)$ может быть как величиной положительной, так и отрицательной, то ошибки аппроксимации для каждого наблюдения принято определять в процентах по модулю:

$$A_i = \frac{1}{n} * \sum \left| \frac{y_i - \hat{y}_x}{y} \right| \cdot 100\% \cdot$$

Эти ошибки уже поддаются сравнению, но они оценивают каждое наблюдение в отдельности. Такую ошибку принято называть относительной ошибкой аппроксимации.

Модель считается подобранной достаточно хорошо, если средняя ошибка аппроксимации не превышает 8-10%. Предпочтение отдается модели с наименьшей ошибкой [ ].

Вопросы для самоконтроля

1. Сформулируйте основные направления поиска измерителя исследуемого признака.
2. Сформулируйте задачи корреляционного анализа в эконометрике.
3. Основные виды уравнений используемых в эконометрике.
4. Как осуществляется выбор вида уравнения регрессии. В чем состоят ошибки спецификации?
5. Способы выбора уравнения регрессии. Ошибки, встречающиеся при построении модели и оказывающие влияние на её качество.
6. Сформулируйте основной принцип метода наименьших квадратов.
7. Опишите основные этапы оценивания с помощью метода инструментальных переменных.
8. Алгоритм метода наименьших квадратов
9. Методика расчёта и применения коэффициента аппроксимации.

СПИСОК ЛИТЕРАТУРЫ

Основная

1. *Валентинов, В. А*. Эконометрика: практикум -2-е изд. – М. : Дашков и К, 2009. - 436 с. – ISBN 978-5-394-00428-5

Дополнительная

1. *Кремер, Н.Ш., Путко Б.*А. Эконометрика : учебник / Н.Ш. Кремер., Б.А. Путко - 2-е изд., стереотип - М. : ЮНИТИ- Дана, 2008. - 311 с. ISBN 978-5-238-01286-5
2. *Афанасьев, Владимир Николаевич*. Эконометрика : учебник / В. Н. Афанасьев, М. М. Юзбашев, Т. И. Гуляева ; ред. : В. Н. Афанасьев. - М. : Финансы и статистика, 2005. - 255 с. ISBN: 5-279-02738-3

Лекция № 3

ЛИНЕЙНАЯ МОДЕЛЬ МНОЖЕСТВЕННОЙ РЕГРЕССИИ

3.1. Спецификация модели

Выбор вида регрессионного уравнения. Выбор формы модели регрессии с одним уравнением и модели временных рядов. Форма может быть линейно и нелинейной. К линейным относятся математическая функция уравнения прямой. Различают два класса нелинейных регрессий: регрессии, нелинейные относительно включенных в анализ объясняющих переменных, но линейные по оцениваемым параметрам. К ним, например, относятся полиномы различных степеней и равносторонняя гипербола. Линейные по переменным и нелинейными по параметрам. Этот класс регрессий подразделяется на две группы: нелинейные модели внутренне линейные и нелинейные модели внутренне нелинейные. В парной регрессии выбор формы математической функции $\hat{y}_x$=f(x) может быть осуществлен тремя методами:

•графическим;

•аналитическим, то есть исходя из теории изучаемой взаимосвязи;

•экспериментальным.

При изучении зависимости между двумя признаками графический метод подбора вида уравнения регрессии достаточно нагляден. Он основан на поле корреляции (слайд 3). Просто подбирается функция, график которой проходит через наибольшее количество точек или как можно ближе к ним. Существует набор функций, используемых наиболее часто

Аналитический метод подразумевает хорошей знание экономики и основан на изучении материальной природы связи исследуемых признаков, теоретических соображениях и опыте подобных предыдущих исследований.

Экспериментальный метод основан на переборе нескольких функций, то есть для каждой функции строится уравнение регрессии, для каждой оценивается качество регрессии и по этим показателям выбирается лучшая. Наиболее часто этот метод используется при обработке данных на компьютере. Большинство статистических пакетов включает в себе автоматический перебор наиболее распространенных функций, а также пользователю предоставляется возможность добавлять дополнительные интересующие его функции. Основным показателем при переборе функций является остаточная дисперсия $D_{ост}$. Необходимо заметить, что нельзя сравнивать $D_{ост}$ по разным данным, сравнению поддаются только остаточные дисперсии по одним и тем же данным, но рассчитанные по разным уравнениям регрессии. Если для нескольких функций $D_{ост}$ одинаковы, то предпочтение отдается наиболее простому по виду уравнению регрессии, так как они более удобны для понимания, для них требуется меньшее количество наблюдений и расчеты, связанные с оцениванием качества регрессии, более просты.

3.2. Обор факторов

Естественным продолжением парной линейной регрессии является множественная линейная регрессионная модель с p переменными.

$$y = a + b_1 \cdot x_1 + b_2 \cdot x_2 + K + b_p \cdot x_p + \varepsilon.$$

Такого рода уравнение может использоваться при изучении роста товарооборота (необходимо заметить, что каждый фактор x_i представляет собой набор из n наблюдений по одному и тому же признаку). Тогда коэффициенты b_i – частные производные товарооборота y по соответствующим факторам x_i:

$$b_1 = \frac{dy}{dx_1},\ b_2 = \frac{dy}{dx_2},\ \ldots,\ b_p = \frac{dy}{dx_p},$$

при условии, что все остальные факторы неизменны.

Множественная регрессия широко используется в решении проблем спроса и потребления, доходности акций, при изучении функции издержек производства, в макроэкономических расчетах и целого ряда других вопросов эконометрики. В настоящее время множественная регрессия — один из наиболее распространенных методов в эконометрике. Основная цель множественной регрессии – построить модель с большим числом факторов, определив при этом влияние каждого из них в отдельности, а также совокупное их воздействие на моделируемый показатель.

Модель линейной множественной регрессии, для которой выполняются условия Гаусса-Маркова, называется нормальной линейной множественной регрессией. Самыми первыми проблемами, с которыми сталкивается исследователь при построении множественной регрессии, являются отбор факторов, которые будут учитываться в регрессионном уравнении, и выбор вида уравнения регрессии. Про выбор уравнения регрессии мы подробнее расскажем в главе 4, где будут рассматриваться нелинейные модели, как парные, так и множественные. А при отборе факторов существуют определенные правила, которые должны выполняться, иначе оценки параметров уравнения, да и само регрессионное уравнение будут недостоверными и не будут отражать связь результативного признака с факторными.

Включение в уравнение множественной регрессии тех или иных факторов связано прежде всего с представлением исследователя о природе взаимосвязи моделируемого показателя с другими экономическими явлениями, то есть исследователь должен иметь какие-либо экономические знания, позволяющие сделать вывод о наличии достаточно сильной связи между рассматриваемыми факторами и моделируемым показателем и включить эти факторы в модель. Факторы, включаемые во множественную регрессию, должны отвечать следующим требованиям:

Они должны быть количественно измеримы. Если необходимо включить в модель качественный фактор, не имеющий количественного измерения, то ему нужно придать количественную определенность (например, в модели урожайности качество почвы задается в виде баллов; в модели стоимости объектов недвижимости учитывается место нахождения недвижимости: районы могут быть проранжированы), наличию или отсутствию какого-либо признака также должно придаваться числовой значение (например, мужчина – 0, женщина – 1).

Каждый фактор должен быть достаточно тесно связан с результатом (то есть коэффициент парной линейной корреляции между каждым включаемым в модель фактором и результатом должен отличаться от нуля, причем достаточно сильно, чтобы подтвердить наличие связи).

Факторы не должны быть тесно связаны между собой и тем более находиться в строгой функциональной связи (они не должны коррелировать друг с другом, или, можно сказать по другому, они не должны быть интеркоррелированы). Включение в модель факторов с высокой интеркорреляцией, когда коэффициент корреляции между двумя факторами больше, чем коэффициент корреляции для любого из этих факторов с результатом (что обозначает более тесную связь между факторами, чем между фактором и результатом $R_{x_i y} < R_{x_i x_j}$) для зависимости $y = a + b_1 \cdot x_1 + b_2 \cdot x_2 + K + b_p \cdot x_p + \varepsilon$ может привести к нежелательным последствиям – система нормальных уравнений может оказаться плохо обусловленной и повлечь за собой неустойчивость и ненадежность оценок коэффициентов регрессии.

Причем чем больше факторов, включаемых в модель, интеркоррелированы друг с другом, тем ненадежнее уравнение регрессии. Если между факторами существует высокая корреляция, то нельзя определить их изолированное влияние на результативный показатель и параметры уравнения регрессии оказываются смещенными. Так, в уравнении $y = a + b_1 \cdot x_1 + b_2 \cdot x_2 + \varepsilon$ предполагается, что факторы x_1и x_2 независимы друг от друга, то есть $r_{x_1 x_2} = 0$. Тогда можно говорить, что параметр b_1 измеряет силу влияния фактора x_1 на результат y при неизменном значении фактора x_2. Если же $r_{x_1 x_2} = 1$, то с изменением фактора x_1 фактор x_2 не может оставаться неизменным. Отсюда b_1 и b_2 нельзя интерпретировать как показатели раздельного влияния x_1 и x_2 на y.

Включаемые во множественную регрессию факторы должны объяснить вариацию независимой переменной. Если строится модель с набором p факторов, то для нее рассчитывается показатель детерминации R^2, который фиксирует долю объясненной вариации результативного признака за счет рассматриваемых в регрессии p факторов. Влияние других не учтенных в модели факторов оценивается как $1 - R^2$ с соответствующей остаточной дисперсией S^2 [ ].

Таким образом, хотя теоретически регрессионная множественная модель позволяет учесть любое число факторов, на практике в этом нет необходимости. Отбор факторов для включения в модель обычно осуществляется в два шага: на первом подбираются факторы исходя из экономической сущности проблемы (то есть набор факторов определяется непосредственно самим исследователем); на втором на основе матрицы показателей парной корреляции определяют тесноту связи для параметров регрессии.

3.2. Частные уравнения регрессии

Уравнение множественной линейной регрессии характеризует весь исследуемый процесс в целом. На основе этого уравнения могут быть построены частные уравнения регрессии. Это уравнения, которые связывают результативный признак с соответствующими факторами x при закреплении других, учитываемых в уравнении множественной регрессии, факторов на среднем уровне:

Если несколько конкретизировать, то частные уравнения регрессии будут иметь следующий вид:

$$\begin{cases} y_{x_1 \cdot x_2, x_3, \ldots, x_p} = a + b_1 \cdot x_1 + b_2 \cdot \bar{x}_2 + b_3 \cdot \bar{x}_3 + \ldots + b_p \cdot \bar{x}_p + \varepsilon, \\ y_{x_2 \cdot x_1, x_3, \ldots, x_p} = a + b_1 \cdot \bar{x}_1 + b_2 \cdot x_2 + b_3 \cdot \bar{x}_3 + \ldots + b_p \cdot \bar{x}_p + \varepsilon, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ y_{x_p \cdot x_1, x_2, \ldots, x_{p-1}} = a + b_1 \cdot \bar{x}_1 + b_2 \cdot \bar{x}_2 + b_3 \cdot \bar{x}_3 + \ldots + b_{p-1} \cdot \bar{x}_{p-1} + b_p \cdot x_p + \varepsilon. \end{cases}$$

Так как средние значения факторов неизменны, то при их подстановке эти уравнения примут вид линейных парных регрессионных уравнений:

$$\begin{cases} \hat{y}_{x_1 \cdot x_2, x_3, \ldots x_p} = A_1 + b_1 \cdot x_1, \\ \hat{y}_{x_2 \cdot x_1, x_3, \ldots x_p} = A_2 + b_2 \cdot x_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ \hat{y}_{x_p \cdot x_1, x_2, \ldots x_{p-1}} = A_p + b_p \cdot x_p, \end{cases}$$

Частные уравнения регрессии характеризуют изолированное влияние определенного фактора на результат, так как другие факторы закреплены на неизменном среднем уровне. Эффект влияния этих факторов присоединены в уравнениях частной регрессии к свободному члену уравнения.

Поэтому во множественном регрессионном анализе возникает проблема определения тесноты связи между двумя признаками в чистом виде, то есть при устранении воздействия других факторов. Мы можем устранить влияние только учтенных в модели факторов.

Показателем чистого влияния фактора на результат при устранении влияния прочих факторов, включенных в модель множественной линейной регрессии, является частный коэффициент корреляции.

Кроме того, коэффициенты частной корреляции используются не только для определения влияния факторов на результат, их можно использовать при отборе факторов для модели множественной регрессии: целесообразность включения в модель того или иного фактора доказывается величиной коэффициента корреляции.

Частные коэффициенты корреляции характеризуют тесноту связи между соответствующим фактором и результатом при устранении влияния других факторов, включенных в модель множественной регрессии.

Коэффициенты частной корреляции представляют собой отношение сокращения остаточной дисперсии за счет включения в модель дополнительного фактора к остаточной дисперсии, которая имела места до его введения в модель.

Рассмотрим это на примере двухфакторной модели $\hat{y} = a + b_1 \cdot x_1 + b_2 \cdot x_2$:

Сокращение остаточной дисперсии в этом случае за счет включения дополнительного фактора x_2 составит:

$$S^2_{yx_1} - S^2_{yx_1x_2}.$$

Чем больше доля этой величины в остаточной дисперсии до введения дополнительного фактора, тем теснее связь между y и x_2 при неизменном воздействии фактора x_1. Коэффициент частной корреляции, показывающий в чистом виде тесноту связи фактора и результата, для двухфакторной модели рассчитывается по следующим формулам:

$$r_{yx_2/x_1} = \sqrt{\frac{S^2_{yx_1} - S^2_{yx_1x_2}}{S^2_{yx_1}}},$$

$$r_{yx_1/x_2} = \sqrt{\frac{S^2_{yx_2} - S^2_{yx_1x_2}}{S^2_{yx_2}}}.$$

Если выразить остаточную дисперсию через коэффициент детерминации $S^2 = \sigma^2_y \cdot (1 - R^2)$, то получим следующие формулы:

$$r_{yx_2/x_1} = \sqrt{1 - \frac{1 - R^2_{yx_1x_2}}{1 - R^2_{yx_1}}},$$

$$r_{yx_1/x_2} = \sqrt{1 - \frac{1 - R^2_{yx_1x_2}}{1 - R^2_{yx_2}}}.$$

Это частные коэффициенты корреляции первого порядка, так как они фиксируют тесноту связи фактора и результата при постоянном воздействии одного фактора [ ].

3.3. Графическое представление модели множественной регрессии

Начнём с рассмотрения примера с факторами, определяющими величину заработка. Расширим первоначальную модель, включив учёт влияния способности к познанию наряду с образованием и допустим что зависимость можно выразить следующим образом:

$$EARNINGS = a + b_1 S + b_2 ASVABC + \varepsilon$$

где EARNINGS –часовой заработок;

S- продолжительность обучения, ASVABC - общий результат выполнения тестов на о познавательные способности, ε – случайный член.

Такая модель, разумеется, является значительным упрощением как с точки зрения состава независимых переменных, включенных в зависимость, так и с точки зрения математической формулы связи [ ].

Для геометрической иллюстрации этой зависимости необходима трехмерная диаграмма с отдельными осями для EARNINGS, S и ASVABC (рис. 3.1)

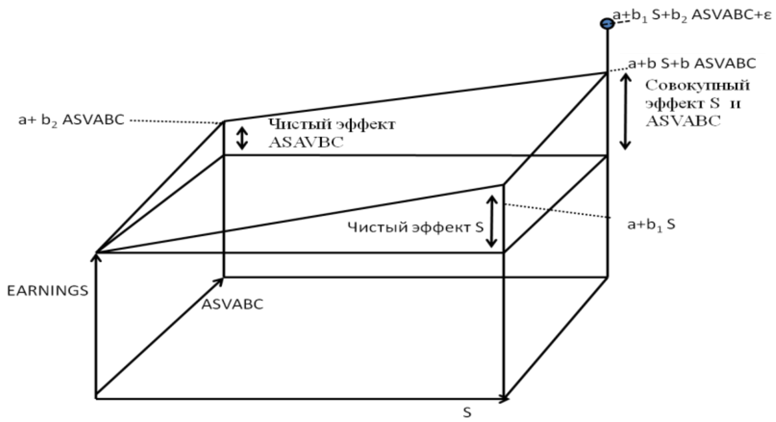

Рис. 3.1. Модель с двумя объясняющими переменными

3.4. Метод наименьших квадратов

Традиционный метод наименьших квадратов для множественной регрессии. При его применении должна минимизироваться остаточная сумма квадратов отклонений фактических величин от теоретических. Для уравнения множественной регрессии $y=a+b_1\cdot x_1+b_2\cdot x_2+\ldots+b_p\cdot x_p+\varepsilon$

Это выглядит следующим образом:

$$S=\sum(y-\hat{y})^2=\sum(y-(a+b_1\cdot x_1+b_2\cdot x_2+\ldots+b_p\cdot x_p))^2\to\min\cdot$$

В данном случае неизвестными являются параметры регрессии a, b_1, b_2,…,b_p. Чтобы найти их, мы дифференцируем остаточную сумму квадратов отклонений по этим переменным и приравниваем их к нулю.

В итоге строится система нормальных уравнений, решение которой и позволяет получить оценки параметров регрессии:

$$\begin{cases}\sum y=n\cdot a+b_1\cdot\sum x_1+b_2\cdot\sum x_2+\ldots+b_p\cdot\sum x_p,\\ \sum y\cdot x_1=a\cdot\sum x_1+b_1\cdot\sum x_1^2+b_2\cdot\sum x_2\cdot x_1+\ldots+b_p\cdot\sum x_p\cdot x_1,\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ \sum y\cdot x_p=a\cdot\sum x_p+b_1\cdot\sum x_1\cdot x_p+b_2\cdot\sum x_2\cdot x_p+\ldots+b_p\cdot\sum x_p^2.\end{cases}$$

Эта система может быть решена методом определителей:

$$a=\frac{\Delta a}{\Delta},\qquad b_1=\frac{\Delta b_1}{\Delta},\qquad b_2=\frac{\Delta b_2}{\Delta},\qquad\ldots\qquad,\qquad b_p=\frac{\Delta b_p}{\Delta},$$

где Δ – определитель системы,

Δa, Δb_1, Δb_2, …, Δb_p – частные определители.

$$\Delta = \begin{vmatrix} n & \sum x_1 & \sum x_2 & \dots & \sum x_p \\ \sum x_1 & \sum x_1^2 & \sum x_2 \cdot x_1 & \dots & \sum x_p \cdot x_1 \\ \sum x_2 & \sum x_1 \cdot x_2 & \sum x_2^2 & \dots & \sum x_p \cdot x_2 \\ \dots & \dots & \dots & \dots & \dots \\ \sum x_p & \sum x_1 \cdot x_p & \sum x_2 \cdot x_p & \dots & \sum x_p^2 \end{vmatrix},$$

а Δa, Δb_1, Δb_2, …, Δb_p получаются заменой соответствующего столбца матрицы определителя системы данными левой части системы.

Кроме того, для линейной множественной регрессии существует другой способ оценки параметров – через β-коэффициенты [ ].

Вопросы для самоконтроля

1. Какие виды моделей используются в эконометрике? От чего зависит выбор модели?
2. Сформулируйте задачи корреляционного анализа в эконометрике.
3. Как осуществляется выбор вида уравнения регрессии. В чем состоят ошибки спецификации?
4. Способы выбора уравнения регрессии. Ошибки, встречающиеся при построении модели и оказывающие влияние на её качество.
5. Определите показатель детерминации и скорректированный коэффициент детерминации
6. Определите частные коэффициенты корреляции
7. Методика оценки параметров множественной линейной регрессии

СПИСОК ЛИТЕРАТУРЫ

Основная

1. . *Елисеева, И. И* Эконометрика : учебник / И. И. Елисеева, С.В. Курышева, Ю. В. Нерадовская и др.] ; под ред. И. И. Елисеевой. - М. :Проспект, 2010. - 288 с.- ISBN 978-5-392-00922-076 с.

Дополнительная

1. *Доугерти, Кристофер*. Введение в эконометрику: учебник: пер. с англ : учебник / Кристофер Доугерти. - М. : Инфра-М, 2009. - 465 с. - ISBN: 978-5-16-003640-3

Лекция 4

СВОЙСТВА ОЦЕНОК МНК

4.1. Мультиколлинеарность

Так как одним из условий построения уравнения множественной регрессии является независимость действия факторов, то коллинеарность факторов нарушает это условие. Если факторы явно коллинеарны, то они дублируют друг друга и один из них рекомендуется исключить из регрессии.

Намного сложнее выявить так называемую мультиколлинеарность факторов, когда более чем два фактора связаны между собой нестрогой линейной зависимостью, то есть наибольшие трудности мы встречаем, когда необходимо выявить совокупное воздействие нескольких факторов друг на друга. Если при исследовании модели мы сталкиваемся в мультиколлинеарностью, это обозначает, что некоторые из включаемых в модель факторов всегда будут действовать вместе. И опять же возникает проблемы при оценке воздействия каждого фактора x_i на результат y в отдельности при неизменном значении прочих факторов, так как вариации мультиколлинеарных факторов перестают быть независимыми друг от друга. Чем сильнее мультиколлинеарность факторов, тем менее надежна оценка распределения суммы объясненной вариации по отдельным факторам с помощью метода наименьших квадратов.

На практике о наличии мультиколлинеарности судят по матрице коэффициентов парной корреляции между факторами (почти такую мы уже рассматривали ранее, в нее не включается только корреляция с результативным признаком), точнее, по её определителю [ ].

4.2. Показатели качества регрессии

Практическая значимость уравнения множественной регрессии оценивается с помощью показателя множественной корреляции и его квадрата – коэффициента детерминации.

Показатель множественной корреляции характеризует тесноту связи рассматриваемого набора факторов с исследуемым признаком, или, иначе, оценивает тесноту совместного влияния факторов на результат.

По аналогии с парной регрессией коэффициент множественной корреляции можно определить как долю дисперсии результата, объясненной вариацией включенных в модель факторов, в его общей дисперсии. Точнее, эту величину характеризует множественный коэффициент детерминации:

$$R^2_{yx_1x_2...x_p} = 1 - \frac{\sigma^2_{ост}}{\sigma^2_y} = 1 - \frac{\sum(y - \hat{y}_{x_1x_2...x_p})^2}{\sum(y - \bar{y})^2},$$

Границы изменения множественного коэффициента корреляции от 0 до 1. Чем ближе его значение к 1, тем теснее связь между результатом и всеми факторами в совокупности и уравнение регрессии лучше описывает фактические данные.

Если отбор факторов, включаемых в регрессию был проведен правильно, то множественный коэффициент корреляции значительно отличается от частных коэффициентов корреляции, так как характеризует тесноту связи между результатом и всеми факторами в совокупности, и каждое включение в уравнение регрессии нового фактора (если, конечно, дополнительный фактор оказывает значимое влияние на результат), будет значимо увеличивать величину множественного коэффициента корреляции. Однако, если дополнительные факторы не оказывают существенного влияния на результат, величина множественного коэффициента корреляции изменится несущественно. Если же в уравнении регрессии учитывается какой-либо фактор, оказывающий наиболее сильное влияние на результативный признак, то частный коэффициент корреляции будет достаточно близок к множественному коэффициенту корреляции, но ни в коем случае не больше него.

Множественный коэффициент корреляции1 рассчитывается для линейной регрессии, поэтому его часто называют линейным коэффициентом множественной корреляции, а так иногда его упоминают как совокупный коэффициент корреляции [ ].

4.3. Условия применения моделей множественной регрессии.

Свойства коэффициентов регрессии существенным образом зависят от свойств случайной компоненты. В самом деле, для того, чтобы регрессионный анализ, основанный на обычном методе наименьших квадратов, давал наилучшие из всех возможных результаты, случайный член должен удовлетворять четырем условиям, известным как условия Гаусса—Маркова. Если эти условия не выполнены, исследователь должен это сознавать. Если корректирующие действия возможны, то аналитик должен быть в состоянии их выполнить. Если ситуацию исправить невозможно, исследователь должен быть способен судить, насколько серьезно это может повлиять на результаты. Более подробно последние три условия изучите самостоятельно.

<u>1-е условие Гаусса-Маркова: $E(\varepsilon i) = 0$ для всех наблюдений</u>

Первое условие состоит в том, что математическое ожидание случайного члена в любом наблюдении должно быть равно нулю. Иногда случайный член будет положительным, иногда отрицательным, но он не должен иметь систематического смещения, ни в каком из двух возможных направлений. Фактически, если уравнение регрессии включает постоянный член, то обычно бывает разумно предположить, что это условие выполняется автоматически, так как роль постоянного члена состоит в отражении любой систематической, но постоянной составляющей в Y, которую не учитывают объясняющие переменные, включенные в уравнение регрессии.

[1] Дело в том, что по ранее указанным формулам можно рассчитывать показатель и для нелинейной множественной регрессии, в этом случае он будет называться индексом множественной корреляции.

2-е условие Гаусса-Маркова: теоретическая дисперсия ε_i постоянна для всех наблюдений

Второе условие состоит в том, что дисперсия случайного члена должна быть постоянна для всех наблюдений. Иногда случайный член будет больше, иногда меньше, однако не должно быть априорной причины для того, чтобы он порождал большую ошибку в одних наблюдениях, чем в других. Эта постоянная дисперсия обычно обозначается σ2ε, или часто в более краткой форме σ_2, а условие записывается следующим образом:

$\sigma^2_{\varepsilon_i} = \sigma^2_{\varepsilon}$ для всех i.

Так как $E(\varepsilon_i)$ равно нулю, и теоретическая дисперсия ε_i, равна $E(\varepsilon_{i2})$ условие можно переписать в виде

$E(\varepsilon_{i2}) = \sigma_{\varepsilon 2}$ для всех i.

Величина σε, конечно, неизвестна. Одна из задач регрессионного анализа состоит в оценке стандартного отклонения случайного члена. Если рассматриваемое условие не выполняется, то коэффициенты регрессии, найденные по обычному методу наименьших квадратов, будут неэффективны, и можно получить более надежные результаты путем применения модифицированного метода регрессии. Это будет рассмотрено в главе 8.

3-е условие Гаусса-Маркова: εi, распределено независимо от ε_j ($i \neq j$)

Это условие предполагает отсутствие систематической связи между значениями случайного члена в любых двух наблюдениях. Например, если случайный член велик и положителен в одном наблюдении, это не должно обусловливать систематическую тенденцию к тому, что он будет большим и положительным в следующем наблюдении (или большим и отрицательным, или малым и положительным, или малым и отрицательным). Значения случайного члена должны быть абсолютно независимы друг от друга. Из данного условия следует, что $Cov(\varepsilon_i,\varepsilon_j)$ — ковариация между ε_i и ε_j равна нулю, поскольку

Если это условие не выполняется, то регрессия, оцененная по обычному методу наименьших квадратов, вновь дает неэффективные оценки.

4-е условие Гаусса-Маркова: случайный член должен быть распределен независимо от объясняющих переменных

Заключительное условие Гаусса—Маркова может быть рассмотрено в двух формах — слабой и сильной. Сильная форма заключается в том, что объясняющая переменная должна быть нестохастической, т. е. не содержать случайных составляющих. Это в действительности весьма нереалистично для экономических переменных, и мы в конечном счете перейдем к слабой форме этого условия, согласно которому объясняющие переменные могут содержать случайные компоненты, но они должны быть распределены независимо от случайного члена. Тем не менее, некоторое время мы будем использовать сильную форму данного условия, поскольку она упрощает анализ свойств оценок.

Трудно представить действительно нестохастические переменные, отличные от переменной времени. Поэтому приводимый ниже пример кажется несколько искусственным. Предположим, что мы связываем величину заработка с продолжительностью обучения S (числом завершенных лет учебы). Допустим, что согласно данным переписи населения 1% респондентов имеет S= 8; 3% имеют S= 9; 5% имеют S= 10; 7% имеют S= 11; 43% имеют S= 12 (выпускники средней школы) и т. д. Предположим, мы решили сделать обзор для выборки размера 1000, и мы хотим, чтобы выборка в максимальной степени соответствовала генеральной совокупности. В этом случае мы можем взять так называемую расслоенную (стратифицированную) случайную выборку, построенную так, что она включает 10 индивидов с S= 8, 30 индивидов с S= 9 и т. д. Величины S в выборке были бы в этом случае предопределенными и, следовательно, нестохастическими. Уровень образования и другие демографические переменные в больших выборках, построенных таким образом, чтобы быть представительными для всего населения, как, например, Всеамериканский опрос молодежи (NLSY) в США, вероятно, вполне удовлетворяют этому условию.

Если данное условие выполнено, то Cov (Xi,ε) - теоретическая ковариация между объясняющей переменной и случайным членом равна нулю. Поскольку E(εi) равно нулю и слагаемое, включающее X, является нестохастическим,

В главах 9 и 10 обсуждаются два важных случая, в которых это условие не выполнено, и последствия этого.

Предположение о нормальности

Наряду с условиями Гаусса-Маркова обычно также предполагает нормальность распределения случайного члена. Читатели должны знать о нормальном распределении из вводного курса статистики. Дело в том, что случайный член ε нормально распределен, то так же будут распределены и коэффициенты регрессии.

Предположение о нормальности основывается на центральной предельной теореме. В сущности, теорема утверждает, что если случайная величина является общим результатом взаимодействия большого числа других случае величин, ни одна из которых не является доминирующей, то она будет иметь приблизительно нормальное распределение, даже если отдельные составляющие не имеют нормального распределения. Случайный член ε определяется несколькими факторами, которые не входят в явной форме в уравнение регрессии. Поэтому, даже если мы ничего не знаем о распределении этих факторов (или даже об их сущности), мы имеем право предположить, что они нормально распределены [ ].

4.4. Тест Голдфелда – Квандта

Совершенно необходимым для получения по МНК состоятельных оценок параметров регрессии является соблюдение 2 го условия Гаусса-Маркова (или 3-й предпосылки МНК) . В соответствии со вторым условием дисперсия остатков должна быть постоянной т.е. гомоскедастичной. Если это условие применения МНК не выполняется, то имеет место гетероскедастичность. Гомоскедастичность остатков означает, что дисперсия остатков εi одинакова для каждого наблюдения x. Предполагается монотонная связь дисперсии остатков с одним из регрессоров

При малом объёме выборки, что наиболее характерно для эконометрических исследований, для оценки гетероскедастичности -Кет использоваться метод Гольдфельда - Квандта, разработанный в 1965 г. Для того чтобы оценить нарушение гомоскедастичности, они предложили параметрический тест, который включает в себя следующие шаги.

Шаг 1. Упорядочение и наблюдений по мере возрастания переменной x. Если используется множественная регрессия то квлючаются все факторы, аранжирование производится по факторуимеющему наиболее тесную связь с результатом. При необходимости тест проводится по каждому фактору с исходных данных. При этом протестированные факторы остаются в модели.акторы остаются в модели

Шаг 2. Исключение из рассмотрения C центральных наблюдений; Центральнее наблюдения должны составлять не более одной четверти всей выборки. Центральными моментами называются наблюдения равноудалённые от медианы. Отнимаем от количества всех наблюдений количество центральных моментов и делим на 2 группы. При этом (n - C): 2 > k , где k -- число оцениваемых параметров

Шаг 3. Получаем две группы из (n - C):2. первая группа S1 с малыми и Вторая группа S2 с большими значениями факторах X. Определяем по каждой из групп уравнений регрессии.

Шаг 4. Определение остаточной суммы квадратов (SSостаточ) ш первой (S1) и второй (S2) групп и нахождение их отношениея: F фак = S2: S1, где S2 >S1.

При выполнении нулевой гипотезы о гомоскедастичности отношение F фак будет удовлетворять F-критерию с (n - C – 2k)степенями свободы для каждой остаточной суммы квадратов. Чем больше величина F фак превышает табличное значение F-критерия тем более нарушена предпосылка о равенстве дисперсий остаточных величин. Принимается решение о существенности различий дисперсий остатков при наибольших и наименьших значениях регрессора, а следовательно о гетероскедастичности остатков [ ].

Вопросы для самоконтроля

1. Дайте определение мультиколлинеарности и её влияния на результаты оценки.
2. Сформулируйте условия Гаусса-Маркова
3. Сформулируйте необходимость нормального распределения случайной компоненты
4. Сформулируйте задачи решаемые с помощью теста Чоу
5. Сформулируйте задачи решаемые с помощью теста Голдфелда-Квандта

СПИСОК ЛИТЕРАТУРЫ

Основная

1. . *Доугерти, Кристофер.* Введение в эконометрику: учебник: пер. с англ : учебник / Кристофер Доугерти. - М. : Инфра-М, 2009. - 465 с. - ISBN: 978-5-16-003640-

Дополнительная

1. *Уткин, В.Б.* Эконометрика : Учебник / под ред. проф. В.Б. Уткина – М. : Издательско-торговая корпорация «Дашков и К°», 2007. – 564 с. –ISBN 978-5-91131-346-3

Лекция 5

РЕГРЕССИОННЫЕ МОДЕЛИ С ПЕРМЕНННОЙ СТРУКТУРОЙ

5.1. Причины изменчивости структуры модели

Как правило, независимые переменные в регрессионных моделях имеют «непрерывные» области изменения (национальный доход, уровень безработицы, размер зарплаты и т.п.). Однако теория не накладывает никаких ограничений на характер регрессоров, в частности, некоторые переменные могут принимать всего два значения или, в более общей ситуации, дискретное множество значений. Необходимость рассматривать такие переменные возникает довольно часто в тех случаях, когда требуется принимать во внимание какой-либо качественный признак. Например, при исследовании зависимости зарплаты от различных факторов может возникнуть вопрос, влияет ли на ее размер, и если да, то в какой степени, наличие у работника высшего образования. Также можно задать вопрос, существует ли дискриминация в оплате труда между мужчинами и женщинами. В принципе можно оценивать соответствующие уравнения внутри каждой категории, а затем изучать различия между ними, но введение дискретных переменных позволяем оценивать одно уравнение сразу по всем категориям.

Использование моделей с переменной структурой позволяет решить проблемы с не включением в модель существенных факторов и неверной функциональной формой модели.

В англоязычной литературе по эконометрике переменные указанного выше типа называются dummy variables, что на русский язык часто переводится как «фиктивные переменные» (см., например, Джонстон, 1980). Следует, однако, ясно понимать, что d такая же «равноправная» переменная, как и любой из регрессоров $x_{1,2,..j.}$ Ее «фиктивность» состоит только в том, что она количественным образом описывает качественный признак.

Покажем, как это можно сделать в примере с зарплатой [ ].

5.2. Фиктивные переменные

Ранее рассматривались регрессионные уравнения с постоянными значениями коэффициентов на множестве исходных данных. Их также называют эконометрическими моделями с постоянной структурой. При построении таких моделей предполагается, что взаимосвязи между зависимой и независимой переменной постоянны и не подвержены изменениям во времени и пространстве.

При анализе реальных экономических данных возможно воздействие качественных факторов, меняющих взаимосвязи между зависимыми и независимыми переменными. В таких случаях модели с постоянной структурой являются недостаточно точными, и для их анализа прибегают к построению моделей с переменной структурой.

Для оценки параметров регрессии, на которые оказывают влияние сопутствующие переменные, привлекают введение так называемых фиктивных переменных.

Под фиктивной переменной понимают переменную, которая равна 1 для конкретной части выборочной совокупности и 0- для оставшейся части.

Рассмотрим зависимость заработной платы работника и уровня образования.

EARNINGS = a+b1S +z1+ε

Определим фиктивную переменную d значения которой равны 1, если респондент мужчина, и 0- женщина.

$$d_i = \begin{cases} 0, & \text{если } i \text{ – индивид женщина;} \\ 1, & \text{если } i \text{ – индивид мужчина;} \end{cases}$$

В уравнении регрессии

$$EARNINGS = a+b_1S +b_2d_i + \varepsilon$$

Коэффициент b_2 отражает различие заработной платы мужчин и женщин. Положительный знак коэффициента означает, что выше средний заработок мужчин, отрицательный – выше средний заработок женщин [ ].

Для того чтобы выяснить, имеются ли различия в заработной плате у мужчин и женщин, необходимо с помощью t- критерия проверить гипотезу $H_0 : b_2 =0$.

Замечание.

Качественное различие можно формализовать с помощью любой переменной, принимающей два значения, а не обязательно значения 0 или 1. Однако в эконометрической практике почти всегда используют лишь фиктивные переменные типа «0-1», поскольку в этом случае интерпретация выглядит наиболее просто. Если бы в рассмотренном выше примере переменная d принимала значение, скажем, 5 для индивидуума мужчина и 2 для индивидуума женщина образования, то коэффициент при этом регрессоре равнялся бы трети среднего изменения заработной платы, если индивид мужчина [ ].

Вопросы для самоконтроля

1. Дайте определение фиктивным переменным?
2. В каких случаях строится уравнение с фиктивными переменными

СПИСОК ЛИТЕРАТУРЫ

Основная

1. *Магнус, Я.Р.* Эконометрика. Начальный курс: учеб. пособие. / Я.Р Магнус, П.К.Катышев, А.А.Пересецкий. - М. : Дело, 2005. - 575 с.- ISBN: 5-7749-0055-X
2. *Елисеева И.И.* Эконометрика : учебник: [для вузов по специальности 080601"Статистика" и другим междисциплинар. специальностям / И. И. Елисеева, С.В. Курышева, Ю. В. Нерадовская и др.] ; под ред. И. И. Елисеевой. - М. :Проспект, 2010. - 288 с.- ISBN 978-5-392-00922-0

Дополнительная

1. *Доугерти, Кристофер.* Введение в эконометрику: учебник: пер. с англ : учебник / Кристофер Доугерти. - М. : Инфра-М, 2009. - 465 с. - ISBN: 978-5-16-003640-
2. *Уткин, В.Б.* Эконометрика : Учебник / под ред. проф. В.Б. Уткина – М. : Издательско-торговая корпорация «Дашков и К°», 2007. – 564 с. –ISBN 978-5-91131-346-3

Лекция 6

НЕЛИНЕЙНЫЕ МОДЕЛИ РЕГРЕССИИ И ИХ ЛИНЕАРИЗАЦИЯ

6.1. Классификация моделей. Подбор линеаризующего преобразования

При наличии прямолинейной зависимости между признаками устанавливается постоянное соотношение. Увеличение факторного признака x на единицу в среднем вызывает увеличение (или уменьшение) y на постоянную величину b. Линейные связи являются основными, параметры уравнения регрессии имеют экономическую интерпретацию. Однако при изучении социально-экономических явлений исследователь часто сталкивается с криволинейными зависимостями. Например, с увеличением осадков урожай (при прочих равных условиях) увеличивается до определенного предела, если же осадки являются излишними, то урожай уменьшается. С увеличением урожайности сельскохозяйственных культур себестоимость 1ц. продукции снижается, однако этот показатель никогда не будет равен нулю.

Если между экономическими явлениями существуют нелинейные соотношения, то они выражаются с помощью соответствующих нелинейных функций (например, равносторонней гиперболы $y=a+\frac{b}{x}+\varepsilon$, степенной функции $y=a\cdot x^b\cdot\varepsilon$ или параболы второй степени $y=a+b\cdot x+c\cdot x^2+\varepsilon$).

Различают два класса нелинейных регрессий:

- регрессии, нелинейные относительно включенных в анализ объясняющих переменных, но линейные по оцениваемым параметрам. К ним, например, относятся полиномы различных степеней и равносторонняя гипербола.

Оценка параметров регрессий нелинейных по объясняющим переменным. Суть оценки состоит в замене «нелинейных» объясняющих переменных новыми «линейными» переменными и сведение нелинейной регрессии к линейной. Рассмотрим применение данного подхода к параболе второй степени: $y=a+b\cdot x+c\cdot x^2+\varepsilon$ $x_1=x,\ x_2=x^2,$ и получаем множественную линейную регрессию с двумя факторами $y=a+b\cdot x_1+c\cdot x_2+\varepsilon$,причем b> 0, c<0, c<b (в этом случае оценка параметров происходит с помощью МНК для множественной регрессии). Примером может служить зависимость душевого дохода от возраста физического лица. Начиная примерно с 14-15 лет доход на 1 лицо постепенно повышается по мере получения образования, квалификации и продвижения по службе. Однако после 45-50 лет душевой доход уже не возрастает (если взять среднюю величину по всему населению) и начинает снижаться по мере перехода все большей части лиц на пенсию, или на более легкую, но ниже оплачиваемую работу.

Среди класса нелинейных, по оцениваемым параметрам, следует назвать равностороннюю гиперболу: $y=a+\frac{b}{x}+\varepsilon$, заменим 1/x $z=\frac{1}{x},$ и получаем простую линейную регрессию $y=a+b\cdot z+\varepsilon$. Может быть использована, например, для характеристики связи удельных расходов сырья, материалов и топлива с объемом выпускаемой продукции. При b>0 имеем обратную зависимость, которая при $x\to\infty$ характеризуется нижней асимптотой, минимальным предельным значением y, оценкой которого служит параметр a. При b<0 имеем медленно повышающуюся функцию с верхней асимптотой при $x \to \infty$, т.е. максимальным предельным уровнем y, оценкой которого служит параметр a.

Для оценки параметров этих регрессий уже возможно использование МНК.

• регрессии, нелинейные по оцениваемым параметрам. К данному классу регрессий относятся уравнения, в которых y нелинейно связан с параметрами. Примером таких нелинейных регрессий являются функции:

степенная

показательная

экспоненциальная

Иначе обстоит дело с регрессиями, нелинейными по параметрам. Этот класс регрессий подразделяется на две группы: нелинейные модели внутренне линейные и нелинейные модели внутренне нелинейные.

Если нелинейная модель внутренне линейна, то с помощью соответствующих преобразований её можно свести к линейному виду, и для полученной регрессии использовать метод наименьших квадратов.

Для степенной функции $y = a \cdot x^b \cdot \varepsilon$ соответствующим преобразованием будет логарифмирование $\ln y = \ln a + b \cdot \ln x + \ln \varepsilon$. Соответственно оценка параметров этой функции производиться не по изначальным рядам наблюдений X и Y, а по их логарифмам. Однако, чтобы рассчитать теоретические значения результативного признака, используется первоначальный вид функции. Также оценка качества уравнения регрессии оценивается по первоначальным данным, а не по преобразованным.

Если же нелинейная регрессия внутренне нелинейна, привести к линейному виду её невозможно, и для неё нельзя использовать МНК. В этом случае используются специальные итеративные процедуры. Есть специальный раздел математики, изучающий различные итеративные методы. Однако модели такого вида достаточно редко применяются в эконометрике, предпочтение отдается регрессиям, приводимых к линейному виду [ ].

6.2. Корреляционное отношение

Для нелинейных регрессий, к которым можно применить метод наименьших квадратов, используются те же формулы, что и к линейным регрессиям. Исключение составляет коэффициент корреляции. Для нелинейных регрессий показателем тесноты связи выступает индекс корреляции p_{xy} (корреляционное отношение), который рассчитывается по следующей формуле:

$$p_{xy} = \sqrt{1 - \frac{\sum (y - \hat{y}_x)^2}{\sum (y - \bar{y})^2}}.$$

Вместо коэффициента детерминации используется термин индекс детерминации, причем между индексом корреляции и индексом детерминации сохраняется то же соотношение, что между коэффициентами корреляции и детерминации для линейной регрессии.

Вопросы для самоконтроля

1. Классы нелинейных регрессий.
2. Методы оценки нелинейных регрессий.
3. По каким признакам классифицируются методы оценки параметров нелинейных регрессий?

СПИСОК ЛИТЕРАТУРЫ

Основная

1.*Балдин, К. В.* Эконометрика: учебник / К. В Балдин., В. Н Башлыков., Н. А. Брызгалов; ред. В. Б. Уткин. - М. : Дашков и К, 2009. - 564 с. - ISBN 978-5-394-00431-5.

2.*Мхитарян, В.С.* Эконометрика: учебник / В.С. Мхитарян, М.Ю Архипова, В.А. Балаш; под ред. д-ра. экон.наук. проф. В. С. Мхитаряна. – М. :Проспект, 2009. – 384 с. ISBN 978-5-392-00188-0

Дополнительная

3. *Валентинов, Вячеслав Аркадьевич*. Эконометрика: практикум -2-е изд. – М. : Дашков и К, 2009. - 436 с. – ISBN 978-5-394-00428-5

Лекция 7

ХАРАКТЕРИСТИКА ВРЕМЕННЫХ РЯДОВ

7.1. Элементы ряда

Временной ряд — это совокупность значений какого-либо показателя за несколько последовательных моментов (периодов) времени. В общем виде модель временного ряда представлена следующим образом

$$y_t = T + S + C + \varepsilon$$

Каждый уровень временного ряда формируется под воздействием следующих факторов:

где Т-временной тренд;

S- сезонная компонента

С – циклическая компонента

ε –случайная компонента

При различных сочетаниях, этих факторов зависимость уровней ряда от времени может принимать разные формы. Если из временного ряда удалить тренд и периодические составляющие, то останется нерегулярная компонента.

Если временной ряд представляется в виде суммы соответствующих компонент см. слайд то модель носит название аддитивной. Представление Временного ряда ввиде произведения компонент соответствует мультипликативной модели.

Решение любой задачи по анализу и прогнозированию временных рядов начинается с построения графика исследуемого показателя, так как на этом этапе можно исследовать компонентный состав временных рядов, а также сделать первые шаги к выбору модели для описания их $y_t = T \cdot S \cdot C \cdot \varepsilon$ динамики.

По графику исследуемого показателя можно определить графического анализа можно определить характер сезонных колебаний. Для аддитивной модели амплитуда сезонных колебаний остаётся примерно постоянной и неизменной во времени.

Моделируют обычно стационарный ряд. Рассмотрим формальное определение стационарности. Так как нас интересует не всё распределение, а средние значения и ковариации, используется понятие слабой стационарности.

В этом случае стационарность временного ряда связывается требованием того, что бы он имел среднее, дисперсию и ковариацию, не зависимо от момента времени t [3].

7.2. Выделение во временном ряду циклической и трендовой компоненты

Временной ряд — это совокупность значений какого-либо показателя за несколько последовательных моментов (периодов) времени. Каждый уровень временного ряда формируется под воздействием следующих факторов; (тренд); циклическую компоненту; случайные факторы.

При наличии тенденции и циклических колебаний значения каждого последующего уровня ряда зависят от предыдущих значений. Корреляционную зависимость между последовательными уровнями временного ряда называют автокорреляцией уровней ряда.

Количественно ее можно измерить с помощью линейного коэффициента корреляции между уровнями исходного временного и уровнями этого ряда, сдвинутыми на несколько шагов во времени.

Определим коэффициент корреляции между рядами yt, yt-1.

Одна из рабочих формул для расчета коэффициента корреляции имеет вид:

$$r_{xy} = \frac{\sum (x_j - \bar{x})(y_j - \bar{y})}{\sqrt{\sum (x_j - \bar{x})^2 \cdot \sum (y_j - \bar{y})^2}}$$

В качестве переменной мы рассмотрим ряд, y2,y3, ..., y8; в качестве переменной y – ряд y1 y2,.... y7. Тогда приведенная выше формула примет вид:

$$r_1 = \frac{\sum_{t=2} (y_t - \bar{y}_1)(y_{t-1} - \bar{y}_2)}{\sqrt{\sum (y_t - \bar{y}_1)^2 \cdot \sum (y_{t-1} - \bar{y}_2)^2}}$$

где

$$\bar{y}_1 = \frac{\sum_{t=2}^{n} y_t}{n-1}; \qquad \bar{y}_2 = \frac{\sum_{t=2}^{n} y_{t-1}}{n-1};$$

Эту величину называют коэффициентом автокорреляции уровней ряда первого порядка, так как он измеряет зависимость между соседними уровнями ряда t и t - 1, т. е. при лаге 1.

При различных сочетаниях, этих факторов зависимость уровней ряда от времени может принимать разные формы [ ].

При анализе многих экономических показателей (особенно в макроэкономике) зачастую используют ежегодные, ежеквартальные, ежемесячные, ежедневные данные. Например, это могут быть годовые данные по ВНП, ВВП, объему чистого экспорта, инфляции и т. д., месячные данные по объему продажи продукции, ежедневные объемы выпуска какой-либо фирмы. Для рационального анализа необходимо систематизировать моменты получения соответствующих статистических данных.

В этом случае следует упорядочить данные по времени их получения и построить так называемые временные ряды. Пусть исследуется показатель Y. Его значение в текущий момент (период) времени t обозначают yt; значения Y в последующие моменты обозначаются y_t+1, y_t+2, ... , yt+k, ...; значения Y в предыдущие моменты обозначаются y_t−1, yt−2, ..., y_t−k, ... Нетрудно понять, что при изучении зависимостей между такими показателями либо при анализе их развития во времени в качестве объясняющих переменных используются не только текущие значения переменных, но и некоторые предыдущие по времени значения, а также само время T. Модели данного типа называют динамическими. В свою очередь переменные, влияние которых характеризуется определенным запаздыванием, называются лаговыми переменными. Обычно динамические модели подразделяют на два класса [3].

Одним из наиболее распространенных способов моделирования тенденции временного ряда является построение аналитической функции, характеризующей зависимость уровней ряда от времени, или тренда. Этот способ называют аналитическим выравниванием временного ряда. Поскольку зависимость от времени может принимать разные формы, для ее формализации можно использовать различные виды функций. Для построения трендов чаще всего применяются следующие функции:

• линейный тренд $\hat{y}_t = a + b \cdot t$;

• гипербола $\hat{y}_t = a + b/t$;

• экспоненциальный тренд $\hat{y}_t = a \cdot b^t$;

• тренд в форме степенной функции $\hat{y}_t = a \cdot t^b$;

парабола второго и более высоких порядков

$$\hat{y}_t = a + b_1 \cdot t + b_2 \cdot t^2 + \ldots + b_k \cdot t^k.$$

Параметры каждого из перечисленных выше трендов можно обычным МНК, используя в качестве независимой переменной время t=1,2,..., n, а в качестве зависимой переменой — фактические уровни временного ряда yt. Для нелинейных трендов предварительно проводят стандартную процедуру их линеаризации.

Известно несколько способов определения типа тенденции; к наиболее распространенным относятся качественный анализ изучаемого процесса, построение и визуальный анализ (графика) зависимости уровней ряда от времени, расчет некоторых основных показателей динамики. В этих же целях можно использовать и коэффициенты автокорреляции уровней ряда. Тип тенденции можно определить путем сравнения коэффициентов автокорреляции первого порядка, рассчитанных по исходным и преобразованным уровням ряда. Если временной ряд имеет линейную тенденцию, то его соседние уровни yt, и yt-1 тесно коррелируют. В этом случае коэффициент автокорреляции первого порядка уровней исходного ряда должен быть высоким. Если временной ряд содержит нелинейную тенденцию, например, в форме экспоненты, то коэффициент автокорреляции первого порядка по логарифмам уровней исходного ряда будет выше, чем соответствующий коэффициент, рассчитанный по уровням ряда. Чем сильнее выражена нелинейная тенденция в изучаемом временном ряде, тем в большей степени будут различаться значения указанных коэффициентов.

Выбор наилучшего уравнения в случае, если ряд содержит нелинейную тенденцию, можно осуществить путем перебора основных форм тренда, расчета по каждому уравнению скорректированного коэффициента детерминации R2 и выбора уравнения тренда с максимальным значением скорректированного коэффициента детерминации. Реализация этого метода относительно проста при компьютерной обработке данных [3].

7.3. Авторегрессионное преобразование

Авторегрессионное преобразование (AR)

Пусть Y − исследуемая величина, и ее изменение можно описать с помощью модели

$$(y_t - m) = \alpha\ 1(y_t{-}1 - m) + u_t$$

где m – среднее значение Y, u t – некоррелированные случайные отклонения с нулевым математическим ожиданием и постоянной дисперсией σ^2 (такие отклонения при рассмотрении временных рядов иногда называют белым шумом).

Преобразование (9.13) в этом случае называют авторегрессионным преобразованием первого порядка AR(1). При этом значение y_t переменной Y в момент времени t пропорционально ее же значению $y_t{-}1$ в момент времени (t – 1) плюс некоторое случайное отклонение. По аналогии

$$(y_t - m) = \alpha\ 1(y_t{-}1 - m) + \alpha\ 2(y_t{-}2 - m) + u_t$$

называется авторегрессионным преобразованием второго порядка AR.

$$(y_t - m) = \alpha\ 1\ (y_t{-}1 - m) + \alpha\ 2\ (y_t{-}2 - m) + \dots + \alpha\ p\ (y_t{-}p - m) + u_t$$

называется авторегрессионным преобразованием порядка P AR.

Во всех этих преобразованиях текущее значение y t переменной Y выражается только через ее предыдущие значения и случайную составляющую (белый шум) u t. *Преобразование методом скользящих средних (MA)*

Пусть поведение моделируется формулой:

$$y_t = \gamma + \beta_0 u_t + \beta_1 u_t{-}1,$$

где γ = const, u_t и $u_t{-}1$ – белый шум в текущий и предыдущий моменты времени.

В этом случае значение переменной Y в момент времени t равно сумме константы и скользящей средней между текущим и предыдущим значениями случайного отклонения (белого шума). Соотношение (9.16) называют преобразованием методом скользящих средних первого порядка MA(1). Соотношение

$$y_t = \gamma + \beta_0\ u_t + \beta_1\ u_t{-}1 + \beta_2\ u_t{-}2 + \dots + \beta_q\ u_t$$

называют преобразованием методом скользящих средних порядка q MA(q).

Преобразование ARMA. Сочетание преобразований AR и MA называется авторегрессионным преобразованием со скользящей средней ARMA. Например, для переменной Y преобразование ARMA(1,1) будет иметь вид:

$$y_t = \gamma + \alpha\ 1 \cdot y_t{-1} + \beta_0 \cdot u_t + \beta_1 \cdot u_t{-1}$$

В общем случае преобразование ARMA(p,q) включает в себя p авторегрессионных членов и q скользящих средних.

Преобразование ARIMA Преобразование ARMA в сочетании с переходом от объемных величин к приростным называется преобразованием ARIMA. В некоторых случаях такой переход позволяет получить более точную и явную модель зависимости. Здесь приращением (конечной разностью) первого порядка переменной Y называется разность $y_t - y_t{-1}$. Приращением порядка d переменной Y называют разность
$y_t - y_t{-1} - y_t{-2} - ... - y_t{-d}$.

В общем виде преобразование ARIMA(p,d,q) выражается формулой:

$$y_t * = \alpha_1 * 1_t\, y - + ... + \alpha_p * p_t\, y - + \beta_0 \cdot u_t + \beta_1 \cdot u_t{-1} + ... + \beta_q \cdot u_t{-q}\ ,$$

где α_i , i = 1, 2, ..., p; βi , i = 0, 1, 2, ..., q – неизвестные параметры.

Величины y_{t-i} , i = 0, 1, ... , p представляют собой конечные разности порядка d переменной Y.u_{t-i}, i = 0, 1, ... , q – независимые друг от друга нормально распределенные случайные величины с нулевым математическим ожиданием и постоянной дисперсией.

Отметим, что преобразования AR, MA и ARIMA целесообразно использовать тогда, когда достаточно ясен набор объясняющих переменных и общий вид уравнения регрессии, но в то же время сохраняется автокорреляция остатков [ ].

Вопросы для самоконтроля

1. Сформулируйте основные составляющие временного ряда
2. В чем состоит различие между моделями с распределёнными лагами и авторегрессионными моделями?
3. Сформулируйте этапы для преобразования Койка?
4. Какой процесс может быть назван процессом скользящего среднего первого порядка?
5. Опишите модель ARMA.

СПИСОК ЛИТЕРАТУРЫ

Основная

1. *Буравлев, Александр Иванович*. Эконометрика: учебное пособие для обучающихся высших учебных заведений, обучающихся по специальности "Статистика" и другим экономическим специальностям / А. Буравлев. - Москва: БИНОМ Лаборатория знаний, 2012. – 163 с. - ISBN: 978-5-9963-0741-8

Дополнительная

1. *Мхитарян, В.С.* Эконометрика: учебник / В.С. Мхитарян, М.Ю Архипова, В.А. Балаш; под ред. д-ра. экон.наук. проф. В. С. Мхитаряна. – М. : Проспект, 2009. – 384 с. ISBN 978-5-392-00188-0.
2. *Елисеева, И. И* Эконометрика : учебник / И. И. Елисеева, С.В. Курышева, Ю. В. Нерадовская и др.] ; под ред. И. И. Елисеевой. - М. :Проспект, 2010. - 288 с.- ISBN 978-5-392-00922-076 с.

Лекция 8

ПРОНОЗИРОВАНИЕ ВО ВРЕМЕННЫХ РЯДАХ

Одна из важнейших целей моделирования заключается в прогнозировании поведения исследуемого объекта. Обычно термин «прогнозирование» используется в тех ситуациях, когда требуется предсказать состояние системы в будущем. Для регрессионных моделей он имеет, однако, более широкое значение. Как уже отмечалось, данные могут не иметь временной структуры, но и в этих случаях вполне может возникнуть задача оценить значение зависимой переменной для некоторого набора независимых, объясняющих переменных, которых нет в исходных наблюдениях. Именно в этом смысле — как построение оценки зависимой переменной — и следует понимать прогнозирование в эконометрике.

Проблема прогнозирования имеет много различных аспектов. Можно различать точечное и интервальное прогнозирование. В первом случае оценка — это конкретное число, во втором — интервал, в котором истинное значение переменной находится с заданным уровнем доверия. Выделяют также безусловное и условное прогнозирование в зависимости от того, известны ли интересующие нас объясняющие переменные точно или приближенно. Кроме того, для временных рядов при нахождении прогноза существенно наличие или отсутствие корреляции по

8.1. Точеное прогнозирование

Одним из наиболее распространенных способов получения мно гофакторных рогнозов является упоминавшийся ранее классиче- ский метод наименьших квадратов и построение на его основе мо- дели множественной регрессии [16]. Неизвестные коэффициенты модели находятся из условия ми- нимума функционала рассогласований, который представляет собой сумму квадратов рассогласований реальных значений зависимой переменной и значений. В векторном виде функционал рассогласо. Условия минимума Ф реализуются при равенстве нулю первых производных функционала по неизвестным коэффициентам, (1.41) Данное условие эквивалентно выполнению векторного соотношения. Надежность получаемой с помощью оценок α модели определя- ется с помощью величины остаточной дисперсии, которая вычисля- ется по формуле коэффициента множественной корреляции, Величина R2 – мно- жественный коэффициент детерминации; она показывает, какая до- ля дисперсии функции объясняется изменениями входящих в урав- нение регрессии независимых переменных при полученных значе- ниях коэффициентов модели. Надежность коэффициента множест- венной корреляции определяется по критерию Фишера при заданном уровне надежности и степени свободы ν 1 = n, ν1 = N – n. Наличие связи между зависимой переменной yi и независимыми переменными xij определяется с помощью коэффициентов корреляции коэффициенты ковариации, определяемые по формуле где i k x , x - средние значения независимых переменных i x и k x . Доверительные интервалы Z1, Z2 для коэффициента корреляции опре- деляются как процентиль нормального распределения N (0, 1) с нулевым средним и единичной дисперсией, z - преобразование Фи- шера, определяемое по формуле. Истинное значение коэффициента корреляции заключено в пре- делах. Для определения надежности оценок строится доверительный интервал для полученных оценок α коэффициентов модели значения критерия Стъюдента с уровнем надежности p и степенями свободы. Поэтому истинное значение коэффициента a, модели будет лежать в интервале Использование вычислительной процедуры по методу наимень- ших квадратов с целью получения оценок коэффициентов модели, которые удовлетворяли бы условиям несмещенности, состоятельности, эффективности, предполагает выполнение ряда условий. Рассмотрим эти условия: – независимые переменные представляют собой неслучайный набор чисел, их средние значения и дисперсии конечны; – случайные ошибки j ε - имеют нулевую среднюю и конечную дисперсию – между независимыми переменными отсутствует корреляция и автокорреляция; – случайная ошибка не коррелированна с независимыми переменными; – случайная ошибка подчинена нормальному закону распреде- ления. Кроме того, можно выделить условие отсутствия мультиколли неарности, когда несколько независимых переменных связаны между собой линейной зависимостью, и условие гомоскедастичности, т. е. одинаковой дисперсии для всех случайных ошибок. Важным является условие линейной формы связи между зависимой и независимой переменными. Зависимость должна быть именно линейной или сводимой к линейной с помощью некоторых преобразований. Но иногда исследуемый процесс не может быть сведен к линейной зависимости никакими преобразованиями, как, например, в случае логистической зависимости. Тогда используется ряд методов, например, метод симплексов. Данный метод отличается сравнительной простотой, легкой реализуемостью на ЭВМ, эффективностью при определении оценок коэффициентов модели.

8.2. Интервальное прогнозирование

Важной характеристикой реализованной модели является оценка ошибки прогноза. Так, предлагается следующая оценка дис персии прогноза. Важным условием получения надежных оценок для модели по методу наименьших квадратов является отсутствие автокорреляции. Оценка автокорреляции для полученной по МНК модели осуществ- ляется по критерию Дарбина – Уотсона где T - длина временного ряда. Полученное расчетное значение d сравнивается с нижней и верхней границей d1 и d2, критерия. Если d < d1, то гипотеза отсутст- вия автокорреляции отвергается; если d > d2, то гипотеза отсутствия автокорреляции принимается; если d1≤ d ≤ d2, то необходимо даль- нейшее исследование. Одним из известных способов уменьшения автокорреляции является авторегрессионное преобразование для исходной информации или переход к разностям, т. е. ΔYt=Yt+1-Yt; ΔXt=Xt+1-Xt. Если же автокорреляцию устранить не удается, то по- лученные оценки считаются состоятельными, и среднеквадратиче- ское отклонение корректируется на величину Δj для j-го коэффициента, где r1 - коэффициент автокорреляции случайных слагаемых первого порядка; R1j - коэффициент автокорреляции для j-й независимой пе- ременной первого порядка. Другим условием, необходимым для получения состоятельных оценок, является отсутствие мультиколлинеарности. Действительно, при наличии мультиколлинеарности определитель квадратной матрицы равен или близок нулю, следовательно, матрица вы- рождена, и поэтому решения системы нормальных уравнений не существует. Эффективный подход к определению мультиколлинеарности предполагает следующую последовательность расчетов. Пусть рас- сматривается уравнение регрессии y=f(x1, …, xn). Тогда для выявле- ния существования мультиколлинеарности предлагается критерий T - определитель матрицы, имеющий асимптоти- ческое распределение Пирсона с n − степенями свободы. В формуле N – число наблюдений по каждому переменному, n – число независимых переменных, матрица X включает значения пе- ременных, преобразованных по i S , x – соответственно оценки среднеквадратического откло- нения и среднее значение для i-й независимой переменной. Далее вычисляются величины которые при неколлинеарности переменных близки единице, а при наличии мультиколлинеарности близки к бесконечности, что дает основание оставить или отбросить показатель xi, что определяется 39 статистикой, имеет распределение Фишера с v = N − n 1 и 1 v2 = n − степенями свободы. Существует еще ряд способов определения мультиколлинеарности, В целях устранения или уменьшения ее можно переходить к разностям для исходной информации, использовать метод факторного анализа или метод главных компонент. Получение прогнозов с помощью многофакторных регрессион- ных моделей предполагает неизменность значений коэффициентов этих моделей во времени. Тем не менее, в процессе исследования объекта возможно появление новой информации, что позволяет с помощью рекуррентного оценивания корректировать значения оце- нок коэффициентов моделей. В то же время исходная информация может содержать в себе различные динамики изменения независи- мых переменных, которые возникают в результате различных «ре- жимов» функционирования исследуемого объекта. В этом случае важным является, как сам факт установления различия динамик про- цессов на разных временных интервалах, так и выбор такого интер- вала для построения на нем модели прогнозирования, который был бы наиболее адекватным будущему поведению объекта. Если ока- зывается, что для одного интервала времени построена многофак- торная модель, а для другого интервала то прогноз будет смещен, а, следователь- но, резко возрастает дисперсия прогноза. Построение адекватных регрессионных моделей для целей про- гнозирования с помощью метода наименьших квадратов предъявля- ет к исходной информации весьма жесткие требования. В ряде слу- чаев эти требования для реальных наблюдений оказываются невы- полненными, поэтому получаемые оценки оказываются неэффективными, а прогноз – недостоверным. Действительно, требование нормальности распределения ошибок, предъявляемое к исходной информации процедурой метода наименьших квадратов, в большом числе случаев оказывается невыполненным. Так, говорится: «Нормальность – это миф. В реальном мире никогда не было и никогда не будет нормального распределения». Поэтому в последнее время ин- тенсивно разрабатывается новое направление в статистике – так на- зываемая робастная статистика, задача которой в том и состоит, 40 чтобы получать эффективные оценки в случаях невыполнения неко- торых предпосылок, например, нормальности распределения, нали- чия аномальных наблюдений. Использование робастных методов получения статистических оценок для информации, содержащей аномальные «выбранные» наблюдения, позволяет значительно повысить надежность получаемых оценок по сравнению с обычным методом наименьших квадратов.

Вопросы для самоконтроля

1. В чем состоит суть корреляционного анализа?
2. Какую роль в корреляционном анализе играет оценка показателей F-статистики Фишера и t-статистики Стьюдента?
3. В чем состоит проблема мультиколлинеарности?
4. Чем затруднен процесс построения адекватных прогнозов на ос- нове регрессионных моделей?

СПИСОК ЛИТЕРАТУРЫ

Основная

1. *Балдин, К. В.* Эконометрика: учебник / К. В Балдин., В. Н Башлыков., Н. А. Брызгалов; ред. В. Б. Уткин. - М. : Дашков и К, 2009. - 564 с. - ISBN 978-5-394-00431-5.
2. *Буравлев, Александр Иванович.* Эконометрика: учебное пособие для обучающихся высших учебных заведений, обучающихся по специальности "Статистика" и другим экономическим специальностям / А. Буравлев. - Москва: БИНОМ Лаборатория знаний, 2012. – 163 с. - ISBN: 978-5-9963-0741-8

Дополнительная

1. *Елисеева, И. И* Эконометрика : учебник / И. И. Елисеева, С.В. Курышева, Ю. В. Нерадовская и др.] ; под ред. И. И. Елисеевой. - М. :Проспект, 2010. - 288 с.- ISBN 978-5-392-00922-076 с.

Лекция 9

СИСТЕМА ЛИНЕЙНЫХ УРАВНЕНИЙ

8.1. Виды систем уравнений

Объектом статистического изучения в социальных науках являются сложные системы, изменение тесноты связи между переменными, построение изолированных уравнений регрессии недостаточно для описания таких систем и объяснения механизма их функционирования. Не всегда получается описать сложное социально-экономическое явление с помощью только одного уравнения. Кроме того, некоторые переменные могут оказывать взаимные воздействия и трудно однозначно определить, какая из них является зависимой, а какая независимой переменной. Поэтому при построении эконометрической модели прибегают к системам уравнений.

Системы эконометрических уравнений включают множество эндогенных (зависимых) переменных и множество предопределенных переменных, к которым относятся лаговые и текущие экзогенные переменные, а также лаговые эндогенные переменные. Все эконометрические модели предназначены для объяснения текущих значений эндогенных переменных по значениям предопределенных переменных.

Система уравнений в эконометрических исследованиях может быть построена по-разному. Выделяют три вида эконометрических систем:

1) Система независимых уравнений, когда каждая зависимая переменная **У** рассматривается как функция только от предопределенных переменных **Х.** Каждое уравнение системы независимых уравнений может использоваться самостоятельно, для нахождения его параметров используют МНК. В принципе каждое уравнение системы является независимым уравнением регрессии, поскольку никогда нет уверенности, что факторы, включаемые в уравнение регрессии полностью объясняют результативный признак, то в каждом уравнении системы также присутствует случайная компонента. В итоге система независимых уравнений при 3-х зависимых переменных и 4-х факторах примет следующий вид:

$$\begin{cases} y_1=a_{01}+a_{11}x_1+a_{21}x_2+a_{31}x_3+a_{41}x_4+e_1 \\ y_2=a_{02}+a_{12}x_1+a_{22}x_2+a_{32}x_3+a_{42}x_4+e_2 \\ y_3=a_{03}+a_{13}x_1+a_{23}x_2+a_{33}x_3+a_{43}x_4+e_3 \end{cases}$$

Однако, если одна зависимая переменная У одного уравнения выступает в виде фактора Х в другом уравнении этой же системы, то можно построить модель, которая будет называться системой рекурсивных уравнений.

2) Система рекурсивных уравнений, когда в каждом последующем уравнении системы зависимая переменная представляет функцию от всех зависимых и предопределенных переменных предшествующих уравнений:

$$\begin{cases} y_1=a_{11}x_1+a_{12}x_2+\ldots+a_{1m}x_m+e_1 \\ y_2=b_{21}y_1+a_{21}x_1+a_{22}x_2+\ldots+a_{2m}x_m+e_2 \\ y_3=b_{31}y_1+b_{32}y_2+a_{31}x_1+a_{32}x_2+\ldots+a_{3m}x_m+e_3 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots. \\ Y_n=b_{n1}y_1+b_{n2}y_2+\ldots+b_{nn\text{-}1}y_{n\text{-}}+a_{n1}x_1+a_{n2}x_2+\ldots+a_{nm}x_m+e_n \end{cases}$$

В этой системе зависимая переменная У включает в уравнение в качестве факторов все зависимые переменные предшествующих уравнений. Наряду с набором собственных факторов Х. Примером такой системы может служить модель производительности труда и фондоотдачи вида:

$$\begin{cases} y_1=a_{11}x_1+a_{12}x_2+a_{13}x_3+e_1 \\ y_2=b_{21}y_1+a_{21}x_1+a_{22}x_2+a_{23}x_3+e_2 \end{cases}$$

где y_1 – производительность труда

y_2 – фондоотдача

x_1 – фондовооруженность труда

x_2 – энерговооруженность труда

x_3 – квалификация рабочих

Как и в предыдущей системе каждое уравнение может использоваться самостоятельно и его параметры можно определить с помощью МНК.

Наибольшее распространение в эконометрических исследованиях получила система взаимозависимых уравнений. Система взаимозависимых (совместных, одновременных) уравнений, когда зависимые переменные в одних уравнениях входят в левую часть (т.е. выступают в роли признаков-результатов), а в других уравнениях – в правую часть системы (т.е. выступают в роли признаков-факторов). В общем виде система таких уравнений выглядит следующим образом:

$$\begin{cases} y_1=b_{12}y_2+b_{13}y_3+\ldots+b_{1n}y_n+a_{11}x_1+a_{12}x2+\ldots+a_{1m}x_m+e_1 \\ y_2=b_{21}y_1+b_{23}y_3+\ldots+b_{2n}y_n+a_{21}x_1+a_{22}x_2+\ldots+a_{2m}x_m+e_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots. \\ y_n=b_{n1}y_1+b_{n2}y_2+\ldots+b_{nn\text{-}1}y_{n\text{-}1}+a_{n1}x_1+a_{n2}x_2+\ldots+a_{nm}x_m+e_n \end{cases}$$

Система таких уравнений также называется системой совместных одновременных уравнений. Тем самым подчеркивается, что в системе одни и те же переменные одновременно используется как зависимые в одних и независимые в др. уравнениях. В эконометрике такая система уравнений также называется структурной формой модели.

В отличии от всех систем, система одновременных уравнений не может использоваться как отдельный набор уравнений, каждое из которых является уравнением регрессии и параметры которого могут быть определены с помощью традиционного МНК. Для нахождения параметров такой системы используются специальные приемы оценивания. Примером системы одновременных уравнений может служить система, отображающая модель в динамике цены и з/пл.

$$\begin{cases} Y_1=b_{12}y_2+a_{11}x_1+e_1 \\ Y_2=b_{12}y_1+a_{22}x_2+a_{23}x_3+e_2 \end{cases}$$

где y_1 – темп изменения месячной з/пл.
y_2 – темп изменения цен
x_1 - % безработных
x_2 – темп изменения постоянного капитала
x_3 – темп изменения цен на импорт сырья

Система одновременных уравнений обычно содержит эндогенные и экзогенные переменные.

Эндогенные переменные – обозначают в приведенной ранее структурной форме модели как У – зависимые переменные, число таких переменных равно числу уравнений в системе.

Экзогенные переменные были обозначены как Х – это предопределенные переменные, влияющие на эндогенные переменные, но независящие от них. Простейшая структурная форма модели имеет следующий вид:

$$\begin{cases} y_1=b_{12}y_2+a_nx_1+e_1 \\ y_2=b_{21}y_1+a_{22}x_2+e_2 \end{cases}$$

где *y* – эндогенные переменные

x – экзогенные переменные.

Классификация переменных на эндогенные и экзогенные зависят от теоретической концепции, принятой исследователем модели. Экономические показатели могут выступать в одних моделях как эндогенные, а в др., как экзогенные переменные. Климатические условия в эконометрических моделях всегда входят в систему как экзогенные переменные. В качестве экзогенных переменных также могут использоваться эндогенные переменные за предшествующий период времени. В этом случае они называются лаговыми переменными. Структурная форма модели позволяет увидеть влияние любой экзогенной переменной на значение эндогенной переменной. Целесообразно в качестве экзогенных переменных выбирать такие показатели, которые могут быть, объектом регулирования. Меняя их и управляя ими можно заранее иметь наименьшие значения эндогенных переменных. Структурная форма модели в правой части содержит коэффициенты b_i и a_j, где b_i – коэффициенты при эндогенных переменных, a_j – коэффициенты при экзогенных переменных . Коэффициенты b_i и a_j называются структурными коэффициентами модели. Все переменные в модели выражены в отклонениях от среднего уровня, т.е. $x=x-x_{ср}$, $y=y-y_{ср}$, поэтому свободный член в каждом уравнении системы одновременных уравнений отсутствует.

Приведенными называют уравнения, полученные из структурных путем подстановки взамен эндогенной переменной в правую часть уравнения ее выражения из другого структурного уравнения, в котором эта эндогенная переменная находится в левой части. После такой подстановки производят преобразования, при которых члены уравнения, содержащие эндогенную переменную, переносят в левую часть, а в правой части остаются только экзогенные переменные. Приведенная форма модели выглядит следующим образом:

$$\begin{cases} y_1=\sigma_{11}x_1+\sigma_{12}x_2+\ldots+\sigma_{1m}x_m \\ y_2=\sigma_{21}x_1+\sigma_{22}x_2+\ldots+\sigma_{2m}x_m \\ y_n=\sigma_{n1}x_1+\sigma_{n2}x_2+\ldots+\sigma_{2m}x_m \end{cases}$$

где σ_{ij} – коэффициенты приведенной формы модели

По своему виду приведенная форма модели ничем не отличается от системы независимых уравнений, параметры которой оцениваются с помощью обычного МНК. Применяя МНК, можно оценить коэффициенты уравнения регрессии, а затем и значения эндогенных переменных всей системы с помощью экзогенных. Коэффициенты приведенной формы модели представляют собой нелинейные функции от коэффициентов структурной формы модели. Рассмотреть такое положение очень просто на примере структурной модели с 2-мя уравнениями и 2-мя экзогенными переменными следующего вида:

$$\begin{cases} y_1=b_{12}y_2+a_{11}x_1 \\ y_2=b_{21}y_1+a_{22}x_2 \end{cases}$$

Приведенная форма модели выглядит следующим образом:

$$\begin{cases} y_1=\sigma_{11}x_1+\sigma_{12}x_2 \\ y_2=\sigma_{21}x_1+\sigma_{22}x_2 \end{cases}$$

Выразив эндогенную переменную y_2 из 1-го уравнения через y_1 и x $Y_2=(Y_1-a_{11}x_1)/b_{12}$ и решая получившуюся систему для Y_1, мы получим следующие формулы зависимости между коэффициентами структурной формы модели и приведенной формы модели:

Таким образом:

$$\sigma_{11}=\frac{a_{11}}{1-b_{12}b_{21}};\quad \sigma_{12}=\frac{a_{22}b_{12}}{1-b_{12}b_{21}}$$

Проделав аналогичную замену для экзогенной переменной y_1 (выразим ее через y_2 и x_2 по 2-му уравнению структурной модели) $Y_1=(Y_2-a_{22}x_2)/b_{21}$ получим аналогичные формулы для коэффициентов приведенной модели для 2-го уравнения:

$$\sigma_{21}=\frac{a_{11}b_{21}}{1-b_{12}b_{21}};\quad \sigma_{22}=\frac{a_{22}}{1-b_{12}b_{21}}$$

Эконометрические модели обычно включают в систему не только уравнения, отражающие взаимосвязи между отдельными переменными, но и выражение соответствующие тенденции развития процесса, а также различного вида тождества. Например в линейной зависимости потребления от дохода можно одновременно использовать тождества дохода. Эта модель одна из первых моделей такого вида и была создана в 40-х гг. ХХв.

$$\begin{cases} y=a+by \\ y=c+x \end{cases}$$

где x – инвестиции в основной капитал и в запасы экспорта и импорта

c – потребление

y – доход

a и b – параметры линейной зависимости потребления от дохода.

Оценки параметров такой модели должны учитывать тождество дохода в отличие от параметров обычной линейной регрессии. Эта модель также переводится в приведенную форму модели, а затем решается. Приведенная форма модели позволяет получить значения эндогенных переменных через экзогенные, но аналитически уступает структурной форме модели, т.к. в ней отсутствуют оценки взаимосвязи между эндогенными переменными.

8.2. Условия идентифицируемости модели

Слово идентификация можно выразить, хотя и неточно русскими словами: узнавание, опознание, установление единства. На этом этапе устанавливается: едины или нет приведенные уравнения со структурными; можно ли по коэффициентам приведенных уравнений опознать, вычислить коэффициенты структурных уравнений.

Из курса математики известно, что не любая система уравнений имеет решение. Невозможно, например, решить систему двух уравнений с тремя и более неизвестными. Однако система трех уравнений с двумя переменными может быть решена, но имеет не одно, а по крайней мере три решения, т.е. опять же определенного ответа не имеет. Сходная ситуация имеет место и в вопросе об идентификации системы структурных уравнений по приведенным уравнениям.

Однозначное решение, т.е. точную идентификацию, имеет такая система, в которой число коэффициентов регрессии приведенных уравнений точно равно числу коэффициентов регрессии структурных уравнений.

Модель считается неидентифицированной, если среди уравнений модели есть хотя бы одно неидентифицированное.

Модель считается сверхидентифицировнной, если среди уравнений модели есть хотя бы одно сверхидентифицированное.

Выполнение условия индефицируемости модели определяется

Уравнение сверхидентифицировано, если для некоторых структурных параметров можно получить более одного численного значения.

Уравнение называется неидентифицировнным, если оценки его структурных параметров невозможно найти по коэффициентам приведенной модели.

Если обозначить число эндогенных переменных в j-m уравнение системы через H, а число экзогенных (предопределенных) переменных, которые содержатся в системе, но не входят в данном: уравнение, — через D, то условие идентифицируемости модели может быть записано в виде следующего счетного правила:

$D + 1 = H$ — уравнение идентифицируемо;

$D + 1 < H$ — уравнение неидентифицируемо;

$D + 1 > H$ — уравнение сверхидентифицируемо.

Для оценки параметров структурной модели система должна быть идентифицируема или сверхидентифицируема

Дополнительные условия идентификации

Если по условиям задачи один из коэффициентов регрессии заранее известен (например, равен единице), то из проверки идентификации он исключается, не учитывается.

Не подлежит индентификации уравнение, являющееся тождеством, т.е. верным при любых значениях коэффициентов.

Не подлежит идентификации рекуррентная система уравнений, при которой каждая эндогенная переменная зависит от предыдущей по графу связей, но не зависит от последующих эндогенных, т.к. рекуррентная система может быть решена без преобразования структурных уравнений в приведенные.

Легко запомнить простое правило: если все экзогенные переменные входят в уравнения всех эндогенных переменных и последние связаны друг с другом, то система заведомо неиндентифицируемая.

В заключение нужно все же сказать, что же делать, если система не идентифицируемая. Нужно уменьшить число коэффициентов регрессии в структурных уравнениях, т.е. исключить один (или более) экзогенный фактор. Какой из них следует исключить, следует решить, принимая в расчет и содержательное значение фактора, и тесноту его связи с результативным эндогенным признаком. Если без какого- то фактора система вообще теряет смысл, значит, нужно искать другие эндогенные и экзогенные переменные, т.е. другой путь исследования объекта [ ].

8.3. Косвенный метод наименьших квадратов

Косвенный МНК предназначен для оценивания структурных параметров отдельного уравнения системы и может дать результат (без сочетания с другими методами, например, с двухшаговым методом наименьших квадратов) только в применении к точно идентифицируемому уравнению.

Наша цель — оценить структурные параметры t го уравнения системы (. Для анализа t-ro уравнения воспользуемся обшей схемой, которая была рассмотрена в п. при выводе 4-го и 5-го условий идентифицируемости отдельного уравнения системы.

Процедуру статистического оценивания структурных параметров t-го уравнения системы разобьем на 2 этапа.

На 1-м этапе оцениваем с помощью обычного МНК б се параметры я-;; (t = l,2,...,m; j = l,2,...,p) приведенной формы, последовательно (н автономно) решая эту задачу для каждого отдельного уравнения системы (Хорошие свойства получаемых МНК-оценок (их состоятельность, а в случае нормальности распределения и их эффективность) обеспечиваются тем, что в модели выполняются все условия классической линейной модели множественной регрессии.

На 2-м этапе используются соотношения (, связывающие структурные параметры i-го уравнения системы с параметрами кц приведенной формы. В случае точной (или сверх-) идентифицируемости [].

8.3. Двухшаговый и трехшаговый метод наименьших квадратов

(2 МНК) Чтобы обойти главное препятствие в применении обычного МНК к отдельному неидентифицируемому уравнению (в данном примере это было уравнение (17.2), описывающее зависимость спроса ' от цены и дохода— коррелированность играющей роль предиктора эндогенной переменной со случайными остатками на 1-м шаге строится регрессия эндогенной переменной-предиктора по предопределенной переменной. В результате получаем оцененное регрессионное уравнение

На 2-м шаге в правую часть анализируемого уравнения вместо эндогенной переменной вставлялось ее регрессионное выражение через x_{it} в результате чего получали уравнение. После этого к данному уравнению применялся обычный МНК с целью состоятельного оценивания коэффициента с (очевидно, вести речь об отдельном оценивании коэффициентов и с в уравнении в котором, не имеет смысла, т. к. переменные связаны чистой мультиколлинеарностью по построению). Получив оценку параметра мы к имеющимся уже двум соотношениям, связывающим параметры структурной и приведенной форм (вытекающим из общей формулы, добавили еще одно: 4- с После этого, решая систему, мы смогли получить оценки параметров 62 и с неидентифицируемого уравнения. Точно такая же процедура 2МНК предлагается и в общем случае. Правда, в отличие от упрощенной модели спроса-предложения , содержащей единственную предопределенную переменную (что и явилось причиной мультиколлинеарности переменных x' и z, в реалистичных эконометрических моделях мы будем крайне редко сталкиваться с проблемой чистой мультиколлинеарности объясняющих переменных на 2-м шаге процедуры, т. к. состав предопределенных переменных в каждом отдельном уравнении почти никогда не исчерпывает всего их априорного

(3 МНК) Трехшаговый метод наименьших квадратов был предложен впервые Зельнером и Тейлом в качестве метода статистического оценивания одновременно всех уравнении модели с учетом возможной взаимной коррелированности регрессионных остатков различных уравнений зтой системы. Этот метод оказывается более эффективным, чем 2 МНК, если случайные остатки в' различных уравнении системы (взаимно коррелированы, т.е. если их ковариационная матрица отлична от

диагональной. Хотя отметим, что и в этой ситуации 2 МНК-оценки структурных параметров системы остаются состоятельными.

В трехшаговом методе наименьших квадратов (ЗМНК) сохранены первые два шага 2 МНК. Однако полученные в результате этих двух шагов, автономно для каждого отдельного (i-го) уравнения, оценки структурных параметров и с определенными соотношениями, не являются окончательными, а пересчитывающей на 3-м шаге следующим образом: используются для подсчета выборочной ковариационной матрицы случайных остатков , а последняя, в свою очередь, используется для одновременного вычисления оценок всех структурных параметров системы с помощью обобщенного метода наименьших (то в рамках соответствующим образом построенной обобщенной модели множественной регрессии) [ ].

Вопросы для самоконтроля

1. Назовите основные виды систем эконометрических уравнений.
2. Каким способом оцениваются параметры одновременных уравнений?
3. Что понимают под идентификацией структурной модели?
4. Перечислите модели на основе систем эконометрических уравнений
5. Как строится структурная модель спроса и предложения?

СПИСОК ЛИТЕРАТУРЫ

Основная

3. *Балдин, К. В.* Эконометрика: учебник / К. В Балдин., В. Н Башлыков., Н. А. Брызгалов; ред. В. Б. Уткин. - М. : Дашков и К, 2009. - 564 с. - ISBN 978-5-394-00431-5.

4. *Буравлев, Александр Иванович.* Эконометрика: учебное пособие для обучающихся высших учебных заведений, обучающихся по специальности "Статистика" и другим экономическим специальностям / А. Буравлев. - Москва: БИНОМ Лаборатория знаний, 2012. – 163 с. - ISBN: 978-5-9963-0741-8

Дополнительная

2. *Елисеева, И. И* Эконометрика : учебник / И. И. Елисеева, С.В. Курышева, Ю. В. Нерадовская и др.] ; под ред. И. И. Елисеевой. - М. :Проспект, 2010. - 288 с.- ISBN 978-5-392-00922-076 с.

Лекция 10

ПАНЕЛЬНЫЕ ДАННЫЕ

10.1. Введение

Панельные данные *(Panel data)* состоят из наблюдений одних и тех же экономических единиц или объектов (индивидуумы, домашние хозяйства, фирмы, регионы, страны и т. п.), которые осуществляются в последовательные периоды времени. Примерами могут служить ежегодные бюджетные обследования одних и тех же домашних хозяйств, ежеквартальные данные о деятельности фиксированного множества предприятий, ежегодные социально- экономические показатели определенной группы стран. Таким образом, панельные данные сочетают в себе как данные пространственного типа (*cross-sectional data),* так и данные типа временных рядов (*time-series data):* в каждый момент времени имеются данные пространственного типа по экономическим единицам, и для каждого такого объекта соответствующие ему данные образуют один или несколько временных рядов. Благодаря специальной структуре панельные данные позволяют строить более гибкие и содержательные модели и получать ответы на вопросы, которые недоступны только в рамках, например, моделей, основанных на пространственных данных.

В частности, возникает возможность учитывать и анализировать ийдйбидуальныё отличйя между экономическими единицами.

ми, что нельзя сделать в рамках стандартных регрессионных моделей.

Классический пример такой ситуации, вошедший в большинство книг, посвященных панельным данным, привел (Ben-Porah, 1973). Предположим, что ежегодное исследование рынка труда показало, что процент работающих замужних женщин равен 50%. Как можно интерпретировать этот факт? Возможны две крайние точки зрения. Согласно первой, полученный результат означает, что каждая замужняя женщина имеет шанс 50% работать в течение года. Согласно второй, результаты исследования показывают, что 50% всех замужних женщин работают полный рабочий день, а остальные 50% вообще не работают. Ясно, что прогноз состояния рынка труда будет существенно разный в зависимости от того, какая из ситуаций имеет место. Более или менее адекватное представление о реальном положении можно получить, если проследить историю некоторого числа индивидуумов в течение определенного периода времени, т. е. в рамках панельных данных.

Можно привести еще несколько примеров, показывающих, что панельные данные дают возможность учесть эффекты, которые невозможно проследить, оставаясь в рамках обычных моделей. Так, при изучении величины ВВП надушу населения имеется возможность для какой-либо страны в каждый период времени наблюдать уровень инфляции, объем инвестиций, денежную массу и т. п. Но кроме этого существуют факторы, которые либо не наблюдаемы, либо нельзя представить в численной форме, но которые могут оказывать существенное влияние на исследуемый показатель: географическое положение, история, культурные традиции и т. д. При этом действие этих факторов можно считать постоянным (т. е. не зависящим от времени) для каждой национальной экономики. Имея лишь пространственные данные для нескольких стран, можно определить влияние обычных экономических факторов на величину ВВП на душу населения, но нельзя выявить индивидуальные различия между странами. При наличии наблюдений за одними и теми же странами в течение нескольких

На микроуровне одна из традиционных задач — объяснение расходов домашних хозяйств на тот или иной товар, например, средства личной гигиены. Можно собрать пространственные данные, включив в них экономические и социально-демографические характеристики семьи и т.п., и получить значимую зависимость расходов на изучаемый товар от семейного дохода. Однако панельные данные могут показать, что доход не всегда оказывает существенное влияние, а разницу в расходах следует объяснять семейными традициями, уровнем культуры и другими факторами, не всегда поддающимися измерению и наблюдению.

Часто индивидуальные факторы коррелированы с другими объясняющими переменными. Так, например, общий уровень культуры семьи и уровень ее дохода естественно считать связанными. В рамках моделей регрессии это означает, что индивидуальный фактор является существенной переменной в модели и ее исключение приводит к смещенным оценкам остальных параметров (см. п. 4.4). Иными словами, модели с панельными данными позволяют получать более точные оценки параметров.

В то же время, поскольку панельные данные содержат наблюдения за одними и теми же объектами в разные периоды времени, предположение о взаимной независимости этих наблюдений становится нереалистичным, поэтому анализ этих моделей может потребовать применения более тонких (по сравнению с обычным методом наименьших квадратов) методов оценивания.

В настоящее время существует много баз панельных данных как на микро-, так и на макроуровне. Например, в National Longitudinal Survey of Labor Market Experience (NLS), http://www.bls.gov/nls/home.htm. содержатся данные по рынку труда в США, в Michigan Panel Study of Income Dynamics (PSID), http://www.isr.umich.edu/src/psid/, — результаты бюджетных обследований домашних хозяйств в США. В последние годы проводится работа по сбору панельных данных в России (Russia Longitudinal Monitoring Survey, http://www.cpc.unc.edu/rlms/), пред-

11.2. Обозначения и основные модели

При работе с реальными панельными данными всегда возникает проблема, какую модель (обычная регрессия, фиксированный или случайный эффект) следует выбрать. На содержательном уровне разницу между моделями можно интерпретировать следующим образом. Обычная модель предполагает, что у экономических единиц нет индивидуальных различий, и в некоторых простых ситуациях такое предположение оправданно. В модели с фиксированным эффектом считается, что каждая экономическая единица «уникальна» и не может рассматриваться как результат случайного выбора из некоторой генеральной совокупности. Такой подход вполне справедлив, когда речь идет о странах, крупных регионах, отраслях промышленности, больших предприятиях. Если же объекты попали в панель «случайно» в результате выборки из большой совокупности, то приемлемой является модель со случайным эффектом. Примером могут служить небольшие фирмы, домашние хозяйства, индивидуумы. Следует, однако, подчеркнуть, что и в подобных ситуациях (особенно для небольшого числа экономических единиц) может возникнуть вопрос о наличии индивидуальных различий, и тогда модель с фиксированным эффектом представляется более предпочтительной.

Сделаем еще одно важное замечание. Модель со случайным эффектом предполагает, что ошибки а* некоррелированы с регрессорами Xjt, т. е. индивидуальный эффект не связан с объясняющими переменными Xjt- Это условие выполняется далеко не всегда, даже для выборок из большой совокупности. Так, в

риведенном выше примере оценивания производственной функции (уравнение (13.5)) индивидуальный эффект, связанный с качеством управления, коррелирует с производственными факторами Xjt: при прочих равных условиях на предприятии с более высоким качеством управления производственные издержки ниже. В некоторых учебниках по панельным данным именно наличие или отсутствие корреляции между индивидуальным эффектом и регрессорами Xjt рассматривается как ключевое различие между моделями с фиксированным и случайным эффектами (см., например, (Johnston and DiNardo, 1997), (Hsiao, 1986)).

Заметим, что независимо от того, коррелированы индивидуальные эффекты с другими объясняющими переменными или нет, оценки с фиксированным эффектом являются несмещенными и состоятельными. Поэтому в любой ситуации модель с фиксированным эффектом дает приемлемые оценки. Однако при отсутствии корреляции эти оценки будут неэффективными, что может быть весьма важным при выборках небольшого объема. В то же время, если в модели присутствует фиксированный эффект, то оценки с помощью случайного эффекта будут несостоятельными. Таким образом, качество оценок существенно зависит от правильной спецификации модели.

Подробнее о выборе моделей с фиксированным или случайным эффектом можно прочесть в книге (Searle, Casella, McCul-loch, 1992).

На практике оценивание с помощью разных моделей часто приводит к существенно разным результатам.

Статистические тестыПомимо содержательных соображений существуют статистические тесты, позволяющие частично решать проблему выбора модели с помощью стандартной техники проверки гипотез. В рассмотренных выше моделях существует определенная иерархия:

1) обычная модель регрессии (13.2) есть частный случай модели с фиксированным эффектом (13.3), когда в последней оц = О, г = 1> ...> п;

Выбор модели

2) обычная модель регрессии (13.2) есть частный случай модели со случайным эффектом (13.14), когда в последней отсутствуют ошибки щ или (что эквивалентно) когда ст2 = 0;

3) модель со случайным эффектом есть частный случай модели с фиксированным эффектом, когда в последней отсутствует корреляция между а* и xjt-

Вопросы для самоконтроля

1. Сформулируйте определение термина "панельные данные"
2. Охарактеризуйте переменные модель со случайными эффектами
3. Охарактеризуйте модель с фиксированными эффектами.
4. Сформулируйте принцип построения динамической модели.

СПИСОК ЛИТЕРАТУРЫ

Основная

1. *Уткин, В.Б.* Эконометрика [Текст]: учебник / под ред. проф. В.Б. Уткина – М. : Издательско-торговая корпорация «Дашков и К°», 2007. – 564 с. –ISBN 978-5-91131-346-3

Дополнительная

1. *Елисеева, И. И* Эконометрика : учебник / И. И. Елисеева, С.В. Курышева, Ю. В. Нерадовская и др.] ; под ред. И. И. Елисеевой. - М. :Проспект, 2010. - 288 с.- ISBN 978-5-392-00922-076 с.
2. . *Буравлев, Александр Иванович.* Эконометрика: учебное пособие для обучающихся высших учебных заведений, обучающихся по специальности "Статистика" и другим экономическим специальностям / А. Буравлев. - Москва: БИНОМ Лаборатория знаний, 2012. – 163 с. - ISBN: 978-5-9963-0741-8

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Балдин К.В*. Эконометрика [Текст]: учебник/ К.В. Балдин [и др.].— Электрон. текстовые данные.— М.: Дашков и К, 2011.— с.ISBN: 978-5-394-01221-1

2. *Буравлёв А.И*. Эконометрика [Текст]: учебное пособие/ Буравлёв А.И.— М.: БИНОМ. Лаборатория знаний, 2012.— 168 с ISBN: 978-5-9963-1047-0 гриф УМО

3. *Валентинов В.А*. Эконометрика [Текст]: учебное пособие/ Валентинов В.А.— М.: Дашков и К, 2010.— 436 с.ISBN: 978-5-394-00682-1

4. *Валентинов, В. А.* Эконометрика: практикум -2-е изд. – М. : Дашков и К, 2009. - 436 с. – ISBN 978-5-394-00428-5

5. *Доугерти, Кристофер*. Введение в эконометрику [Текст]: учебник: пер. с англ : учебник / Кристофер Доугерти. - М. : Инфра-М, 2009. - 465 с. - ISBN: 978-5-16-003640-3

6. *Елисеева, И. И* Эконометрика : учебник / И. И. Елисеева, С.В. Курышева, Ю. В. Нерадовская и др.] ; под ред. И. И. Елисеевой. - М. :Проспект, 2010. - 288 с.- ISBN 978-5-392-00922-076 с.

7. *Кремер Н.Ш*. Эконометрика [Текст]: учебник/ Кремер Н.Ш., Путко Б.А.— М.: ЮНИТИ-ДАНА, 2010.— 328 с.ISBN: 978-5-238-01720-4.

8. *Мхитарян В.С.* Эконометрика [Текст]: учебное пособие/ Мхитарян В.С., Архипова М.Ю., Сиротин В.П— М.: Евразийский открытый институт, 2012.— 224 с.ISBN: 978-5-374-00053-5

9. *Новиков А.И.* Эконометрика [Текст]: учебное пособие/ Новиков А.И.— Электрон. текстовые данные.— М.: Дашков и К, 2013.— 224 с.ISBN: 978-5-394-01683-7 – Режим доступа: http://e.lanbook.com/books/element.php?pl1_id=5670, регистрация.

10. *Орлов А.И*. Эконометрика [Электронный ресурс]/ Орлов А.И.— Электрон. текстовые данные.— М.: Интернет-Университет Информационных Технологий (ИНТУИТ), 2009.— 623 с.ISSN: 2227-8397

11. *Тихомиров, Н.П.* Методы эконометрики и многомерного статистического анализа [Текст]: учебник / Н.П. Тихомиров, Т.М. Тихомирова, О.С. Ушмаев. - М: Экономика, 2011 г. - 647 с. - ISBN: 978-5-282-03080-8.

12. *Уткин, В.Б.* Эконометрика [Текст]: учебник / под ред. проф. В.Б. Уткина – М. : Издательско-торговая корпорация «Дашков и К°», 2007. – 564 с. –ISBN 978-5-91131-346-3

13. *Яковлева А.В*. Эконометрика [Текст]: учебное пособие/ Яковлева А.В. — Саратов: Ай Пи Эр Медиа, 2011.— с.ISSN: 2227-8397.

СОДЕРЖАНИЕ