ЭКОНОМЕТРИКА

Лекция 1.

§ 1. Введение.

Список рекомендуемой литературы.

Основная.

1. Бородич С.А., Эконометрика. Минск, ООО «Новое знание», 2004.
2. Магнус Я.Р., Катышев П.К., Пересецкий А.Л. Эконометрика. Начальный курс. М.: Дело, 2001.
3. Эконометрика: Учебник / Под ред. И.И. Елисеевой. М.: Финансы и статистика, 2006.
4. Катышев П.К., Магнус Я.Р., Пересецкий А.А. Сборник задач к начальному курсу эконометрики. М.: Дело, 2002.

Дополнительная.

1. Прикладная статистика. Основы эконометрики: Учебник для вузов: В 2-х т. – Т. 2. Айвазян С.А. Основы эконометрики. М: ЮНИТИ-ДАНА, 2001.
2. Замков О.О., Толстопятенко А.В., Черемных Ю.Н. Математические методы в экономике. М. 1997.
3. Харин Ю.С., Малюгин В.И., Харин А. Ю. Эконометрическое моделирование. Минск, БГУ, 2003.

Эконометрика – одна из базовых дисциплин экономического образования во всем мире. Основные составляющие эконометрики – экономическая теория, экономическая статистика и математика.

Существуют различные варианты определения эконометрики:

1) расширенное определение, при котором к эконометрике относят все, что связано с измерениями в экономике;

2) узко инструментально ориентированное определение, при котором под эконометрикой понимают определенный набор математико-статистических средств, позволяющих верифицировать модельные соотношения между анализируемыми экономическими показателями.

Термин эконометрия (эконометрика) был введен в научную литературу в 1930 году норвежским статистиком Рагнаром Фришем для обозначения нового направления научных исследований, возникшего из необходимости научно-обоснованного подтверждения и доказательства концепций и выводов экономической теории результатами количественного анализа рассматриваемых процессов.

Основные задачи эконометрики:

- Построение эконометрических моделей, т.е. представление экономических моделей в математической форме, удобной для проведения эмпирического анализа. Данную проблему принято называть проблемой *спецификации.*
- Оценка параметров построенной модели, делающих выбранную модель наиболее адекватной реальным данным. Этот этап носит название этапа *параметризации.*
- Проверка качества найденных параметров модели и самой модели в целом. Этот этап иногда называют этапом *верификации.*
- Использование построенных моделей для анализа исследуемых экономических показателей, прогнозирования и предсказания, а также для осмысленного проведения экономической политики. В целом эта задача носит название задачи *прогнозирования.*

Изучение экономических процессов (взаимосвязей) осуществляется через математические (эконометрические) модели. Для описания сущности эконометрической модели удобно разбить весь процесс моделирования на шесть основных этапов:

- 1-й этап (постановочный) – определение конечных целей моделирования, набора участвующих в модели факторов и показателей, их роли;
- 2-й этап (априорный) – предмодельный анализ экономической сущности изучаемого явления, формирование и формализация априорной информации, в частности, относящейся к природе и содержанию исходных статистических данных и случайных остаточных составляющих;
- 3-й этап (параметризация) – собственно моделирование, т.е. выбор общего вида модели, в том числе состава и формы входящих в нее связей;
- 4-й этап (информационный) – сбор необходимой статистической информации, т.е. регистрация значений участвующих в модели факторов и показателей на различных временных или пространственных промежутках функционирования изучаемого явления;
- 5-й этап (идентификация модели) – статистический анализ модели и в первую очередь статистическое оценивание неизвестных параметров модели;
- 6-й этап (верификация модели) – сопоставление реальных и модельных данных, проверка адекватности модели, оценка точности модельных данных.

§ 2. Основные типы моделей.

Построение эконометрической модели является центральной проблемой любого эконометрического исследования, поскольку ее “качество” непосредственно определяет достоверность и обоснованность результатов анализа тенденций развития, прогнозов рассматриваемых социально-экономических процессов, а также вытекающих из них выводов, в том числе и по вопросам разработки необходимых управленческих мероприятий.

В эконометрических исследованиях обычно предполагается, что закономерности моделируемого процесса складываются под влиянием ряда других явлений, факторов. Обобщенная форма эконометрической модели, описывающей закономерности развития такого процесса, обозначенного переменной y, в зависимости от уровней, воздействующих на него внешних явлений, факторов x_i, i=1, 2,..., n, может быть представлена следующим уравнением:

$$y_t = f(\alpha, x_t) + \varepsilon_t, \qquad (1.1)$$

где $f(\boldsymbol{\alpha}, \mathbf{x}_t)$ – функциональная зависимость, отражающая вид и структуру взаимосвязей между уровнями переменных y_t и x_{it} в моменты времени t=1, 2,..., T (или на интервалах (t, t+1)); $\mathbf{x}_t = (x_{1t}, x_{2t},..., x_{nt})$ – вектор значений независимых переменных (факторов) в момент t; $\boldsymbol{\alpha}=(\alpha_0, \alpha_1,..., \alpha_n)$ – вектор параметров модели; параметр α_i отражает степень влияния фактора x_i на переменную y на всем рассматриваемом интервале (1, T); α_0 – постоянная модели; ε_t – случайная ошибка модели в момент t, в отношении свойств и характеристик которой, как это будет показано далее, обычно выдвигаются некоторые дополнительные предположения.

Заметим, что в некоторых эконометрических исследованиях значения зависимой переменной y_t и факторов x_{it}, t=1, 2,..., T; i=1, 2,..., n; отражают распределение их уровней на совокупности однородных объектов. В этом случае индекс t выражает порядковый номер объекта (территории, предприятия и т. п.), а модель (1.1) – распределение переменной y на совокупности однородных объектов под влиянием факторов, характеризующих их специфические свойства.

Факторы x_i , i=1, 2,..., n, называют *независимыми*, подчеркивая их независимость от переменной y в смысле отсутствия обратного влияния y на x_i. В связи с этим факторы x_i часто именуют *экзогенными* (внешними или факторными) переменными, а переменную y

– *эндогенной* (внутренней или переменной отклика) переменной модели. Здесь термин "внутренний" подчеркивает также то обстоятельство, что функциональная зависимость $f(\boldsymbol{a}, \boldsymbol{x}_t)$ играет основную роль при определении расчетных значений зависимой переменной $\hat{y}_t$, после того как с использованием того или иного метода будут найдены количественные значения оценок a_i параметров модели α_i, i=0, 1,..., n; $\boldsymbol{a}=(a_0, a_1,..., a_n)$ – вектор оценок параметров модели. Термин "внешний" отражает тот факт, что значения переменных x_{it} определяются вне модели (задаются только в качестве исходных данных).

Эконометрика допускает различные предположения относительно "статистического" содержания внешних переменных x_{it}, в то время как переменная y согласно (1.1) всегда рассматривается как случайная величина.

Три основных варианта статистической интерпретации независимых переменных:

1) независимые переменные x_i (или часть из них) являются *детерминированными величинами* (их значения причинно обусловлены и не зависят от случайных изменений);

2) независимые переменные x_i рассматриваются как *случайные величины*; их значения в каждый момент времени определяются неоднозначно и зависят от случая;

3) независимые переменные x_i – это величины *значения которых x_{it} определены с ошибкой* (ошибкой измерения).

Выражение (1.1) определяет лишь общий вид эконометрической модели. В конкретных эконометрических исследованиях могут использоваться также специальные типы моделей, каждый из которых имеет свои характерные особенности. Эти типы обычно можно классифицировать на основе двух признаков. Во-первых, по виду экзогенных факторов x_i и, во-вторых, по свойствам ошибки модели ε_t.

В *моделях регрессии классического типа* обычно используются факторы, независимые между собой и с ошибкой модели, в предположении, что ошибка модели имеет свойства "белого шума" – процесса с нулевым математическим ожиданием, постоянной конечной дисперсией и нулевой корреляцией между ее разновременными значениями (рядами ε_t и ε_{t-1}, ε_t и ε_{t-2} и т. д., t=1, 2,..., T). Это означает, что в ряду ошибки ε_t отсутствуют автокорреляционные связи.

В *моделях с лаговыми независимыми переменными* в качестве факторов используются разновременные значения хотя бы одной из переменных x, т. е. значения x_{it}, $x_{i,t-1}$, $x_{i,t-2}$,... Аналогично в моделях с лаговыми зависимыми переменными в качестве экзогенных факторов рассматриваются значения переменной y в прошедшие моменты времени, т. е. значения y_{t-1}, y_{t-2},...

Кроме того, модели могут различаться и по свойствам их ошибок. *Модели с непостоянной дисперсией распределения ошибки* на различных участках интервала t=1, 2,..., T; *с наличием автокорреляционных связей между соседними значениями* ε_t *и* ε_{t-1} и т. д. Наконец, рассматриваются модели, в которых *ошибка характеризуется корреляционными связями с экзогенными переменными* x_{it}, как это имеет место, например, в системах эконометрических моделей, а также обладать другими специфическими свойствами.

Специальные типы эконометрических *моделей временных рядов (стационарных и нестационарных)*, получившие широкое применение в исследованиях динамики финансовых показателей и ряда других явлений, описывают процессы, тенденции которых предопределены их внутренними закономерностями.

Состав переменных x_i и форма функциональной зависимости f могут отражать либо экономическую концепцию, лежащую в основе взаимосвязи между зависимой и независимыми переменными, либо эмпирические (т. е. выявленные в ходе конкретных исследований) взаимосвязи между ними.

Исходными данными, необходимыми для построения эконометрической модели, являются известные наборы (массивы) значений зависимой переменной y и независимых факторов x_i. При этом могут использоваться два принципиально различных типа исходных информационных массивов – *статический и динамический*. Статический массив представляет собой значения *результирующей* (*зависимой, объясняемой* и т.п.) *переменной* y и влияющих на нее *факторов* (*независимых, объясняющих переменных*) x_i, имевших место у объектов однородной совокупности в определенный период времени. Примером таких объектов являются однотипные промышленные предприятия (заводы одной отраслевой направленности). В качестве y в практических исследованиях часто рассматриваются показатели производительности труда, объемов выпускаемой продукции и некоторые другие. В качестве x_i – влияющие на уровень этих показателей факторы – объемы используемых фондов, численность и квалификация рабочей силы и т.п.

Приведем другой пример статической информации, характерной для социальных исследований. В качестве y рассматриваются показатели заболеваемости (смертности) населения в регионах страны. Их уровень в каждом из регионов определяют значения независимых факторов, отражающих достигнутый материальный уровень жизни, климатические условия, состояние окружающей среды и т. п. В этом случае необходимая для построения эконометрической модели информация собирается по совокупности регионов страны за фиксированный промежуток времени (год).

Эконометрическая модель, использующая динамическую информацию, связывает значения некоторой зависимой переменной y_t в моменты времени t со значениями независимых переменных (факторов) x_{it}, рассматриваемых в те же моменты времени (или в предшествующие) t=1,2,..., T. Такая информация может отражать, например, уровни производительности труда на одном из заводов и определяющих их значения факторов в последовательные периоды времени.

Исходная информация для построения эконометрических моделей может быть и смешанного типа. Например, она может выражать уровни интересующих переменных по группе заводов за ряд лет.

При формировании исходной информации для эконометрической модели чрезвычайно важной проблемой является *выбор показателей*, адекватных сущности исследуемых явлений. И здесь следует обратить внимание на определенную подмену понятий, которая обычно происходит на первом этапе построения модели при переходе от содержательного анализа явлений к формированию отражающих их уровни количественных характеристик (показателей).

В ходе содержательного анализа явление часто рассматривается на качественном уровне. При этом специалисты оперируют достаточно обобщенными понятиями, например, заболеваемость, уровень медицинского обслуживания, качество и уровень жизни, климат, качество рабочей силы и т. п. В этой связи, заметим, что часто эконометрическая модель строится именно для выражения закономерности, существующей между явлениями. Однако при построении модели используется исходная информация, наборы показателей, которые выражают эти явления, их свойства, тенденции в виде количественных характеристик. Вследствие этого желательно, чтобы такое «выражение» было в некотором смысле как можно более “точным”.

Для традиционных направлений исследований проблема обоснования состава показателей обычно считается решенной. Например, в исследованиях производительности труда, в макроэкономическом анализе обычно рассматриваются уже устоявшиеся наборы показателей, значения которых публикуются в статистических сборниках, научных отчетах и т. п. Их примером являются выработка на одного работающего как показатель, выражающий явление “производительность труда”, объемы ВВП (показатель результативности экономики), объем основных фондов (показатель уровня материальной обеспеченности производственного процесса, экономики) и т.д.

Вместе с тем, в ряде областей эконометрических исследований такие системы показателей не могут быть сформированы столь однозначно. Часто одно и то же явление может быть выражено альтернативными вариантами показателей. Например, общий показатель заболеваемости в регионе за год может быть выражен суммарным числом заболеваний населения в течение этого периода времени. С другой стороны, в качестве меры заболеваемости может выступать и показатель, выраженный в виде суммарного количества дней продолжительности болезней.

Рассматривая проблему выбора конкретного вида функциональной зависимости $f(\boldsymbol{\alpha}, \boldsymbol{x}_t)$ из выражения (1.1) заметим, что в практике эконометрических исследований используется достаточно широкий круг функциональных зависимостей:

1) линейная

$$y_t = \alpha_0 + \alpha_1 x_{1t} + ... + \alpha_n x_{nt} + \varepsilon_t, \qquad (1.2)$$

2) правая полулогарифмическая

$$y_t = \alpha_0 + \alpha_1 \ln x_{1t} + ... + \alpha_n \ln x_{nt} + \varepsilon_t, \qquad (1.3)$$

3) степенная

$$y_t = \alpha_0 \cdot x_{1t}^{\alpha_1} \cdot x_{2t}^{\alpha_2} \cdot ... \cdot x_{nt}^{\alpha_n} + \varepsilon_t \ , \qquad (1.4)$$

4) гиперболическая

$$y_t = \alpha_0 + \alpha_1 / x_{1t} + ... + \alpha_n / x_{nt} + \varepsilon_t, \qquad (1.5)$$

5) логарифмическая гиперболическая

$$\ln y_t = \alpha_0 + \alpha_1 / x_{1t} + ... + \alpha_n / x_{nt} + \varepsilon_t, \qquad (1.6)$$

6) обратная линейная (функция Торнквиста)

$$1/y_t = \alpha_0 + \alpha_1 / x_{1t} + ... + \alpha_n / x_{nt} + \varepsilon_t, \qquad (1.7)$$

7) функция с постоянной эластичностью замены

$$y_t = \left[\alpha_1 x_{1t}^{-\rho} + ... + \alpha_n x_{nt}^{-\rho}\right]^{-\lambda/\rho} \cdot \varepsilon_t, \qquad (1.8)$$

где λ и ρ – также параметры функции.

8) экспоненциальная функция

$$y_t = \alpha_0 + \alpha_1 \cdot e^{\beta_1 \cdot x_{1t}} + ... + \alpha_n \cdot e^{\beta_n \cdot x_{nt}} + \varepsilon_t, \qquad (1.9)$$

где $\beta_1, ..., \beta_n$ – также параметры функции.

На практике могут встретиться и комбинации рассмотренных выше зависимостей. Например,

$$y_t = \alpha_0 + \alpha_1 / x_{1t} + \alpha_2 \ln x_{2t} + \alpha_3 x_{3t}^{\alpha_4} + ... + \varepsilon_t$$

и т. п.

Здесь необходимо отметить, что большинство функций $f(\boldsymbol{\alpha}, \boldsymbol{x}_t)$ с помощью определенного набора преобразований могут быть приведены к линейной форме (1.2). Например, если y и x_i связаны зависимостью $y \sim 1/x_i$ (выражение (1.5)), то, введя переменные $u_i = 1/x_i$, получим выражение (1.2) с точностью до преобразования исходных факторов.

В практических исследованиях часто, используя преобразование $u_i = \ln x_i$ и $z = \ln y$, степенную модель (1.4) преобразуют к линейному виду, связывающему логарифмы переменных y и x_i. Однако заметим, что в данном случае с точки зрения математики такое преобразование не совсем корректно из-за “аддитивности” ошибки в выражении (1.4). Вследствие этого значения коэффициентов линейной относительно логарифмов переменных модели нельзя в общем случае полагать равными соответствующим значениям степенного аналога.

На примере линейной эконометрической модели покажем еще одну возможную форму представления моделей такого типа – моделей, в которых отсутствует свободный коэффициент α_0. В общем случае такая модель представляется в следующем виде:

$$\overset{\circ}{y}_t = \alpha_1 \overset{\circ}{x}_{1t} + \alpha_2 \overset{\circ}{x}_{2t} + \ldots \alpha_n \overset{\circ}{x}_{nt} + \varepsilon_t. \qquad (1.10)$$

Найдем взаимосвязи между переменными y_t и $\overset{\circ}{y}_t$, x_{it} и $\overset{\circ}{x}_{it}$ и определим, чему равен коэффициент α_0. Для этого просуммируем по индексу t правую и левую части модели (1.2). Получим

$$\sum_t y_t = T \cdot \alpha_0 + \alpha_1 \sum_t x_{1t} + \ldots \alpha_n \sum_t x_{nt} + \sum_t \varepsilon_t.$$

Поскольку $\sum \varepsilon_t = 0$, что отражает свойство равенства нулю математического ожидания ошибки ($M[\varepsilon_t]$= $\sum \varepsilon_t / T = 0$), то, разделив правые и левые части этого выражения на T, получим

$$\frac{\sum_t y_t}{T} = \bar{y} = \alpha_0 + \alpha_1 \bar{x}_1 + \ldots + \alpha_n \bar{x}_n,$$

откуда следует, что

$$\alpha_0 = \bar{y} - \alpha_1 \bar{x}_1 - \ldots - \alpha_n \bar{x}_n. \qquad (1.11)$$

Вычтем α_0 из уравнения (1.2). Получим для всех t

$$y_t - \bar{y} = \alpha_1 (x_{1t} - \bar{x}_1) + \ldots + \alpha_n (x_{nt} - \bar{x}_n) + \varepsilon_t. \qquad (1.12)$$

Из (1.12) непосредственно вытекает, что

$$\overset{\circ}{y}_t = y_t - \bar{y}, \; \overset{\circ}{x}_{it} = x_{it} - \bar{x}_i. \qquad (1.13)$$

Операция (1.13) определяет так называемые *центрированные переменные* и называется *операцией центрирования*. Отметим, что для центрированных переменных справедливы следующие очевидные соотношения:

$$\sum_t \overset{\circ}{y}_t = 0 \;, \; \sum_t \overset{\circ}{x}_{it} = 0 \;. \qquad (1.14)$$

Использование центрированных переменных иногда значительно упрощает процедуры получения некоторых результатов, делает более наглядной их интерпретацию.

При этом следует помнить, что исходная информация для такой модели (вектор $\overset{\circ}{\boldsymbol{y}}$ и матрица $\overset{\circ}{\boldsymbol{X}}$) получается путем вычитания из каждого элемента каждого столбца соответствующего среднего (по столбцу) значения.

Как было отмечено выше, конкретный вид функциональной зависимости $f(\boldsymbol{\alpha}, \boldsymbol{x}_t)$ может выражать какую-либо содержательную концепцию, отражающую предполагаемый характер взаимосвязей между процессами y_t и x_{it}, i=1, 2,..., n.

В основе использования степенной функции (1.4), например, обычно лежит концептуальное *допущение о постоянстве частной эластичности выпуска у по каждому ресурсу (фактору)* x_i. Напомним, что частная эластичность в точке t показывает, на сколько процентов изменится зависимая переменная y_t при изменении фактора x_{ti} на 1% при условии постоянства значений остальных факторов в этой точке. Частная эластичность определяется следующим выражением:

$$Э_{it} = \frac{\partial\ y_t}{\partial\ x_{it}} \cdot \frac{x_{it}}{y_t}. \quad (1.15)$$

Подставим вместо y_t в правую часть выражения (1.15) функцию $\alpha_0 x_{1t}^{\alpha_1} \dots\ x_{nt}^{\alpha_n}$. Учитывая, что $\frac{\partial y_t}{\partial x_{it}} = \alpha_i \cdot \alpha_0 \cdot x_{1t}^{\alpha_1} \dots x_{it}^{\alpha_i - 1} \dots x_{nt}^{\alpha_n}$, получим

$$Э_i = \alpha_i. \quad (1.16)$$

Таким образом, коэффициент модели (1.4) α_i сразу определяет значение эластичности y по фактору x_i на всем интервале $(1,T)$.

Удобство экономической интерпретации параметров модели (1.4), относительная простота ее записи и послужили причиной ее широкого использования особенно в макроэкономических исследованиях.

Особую известность получили различные модификации *двухфакторной функции Кобба-Дугласа*

$$y_t = \alpha_0 x_{1t}^{\alpha_1} x_{2t}^{\alpha_2} + \varepsilon_t, \quad (1.17)$$

которые обычно применяется в макроэкономических исследованиях при анализе взаимосвязи между объемом полученного валового внутреннего продукта (y) и используемыми ресурсами (x_1 – основные фонды и x_2 – затраты живого труда). Между собой эти модификации, в основном, различаются ограничениями, накладываемыми на значения коэффициентов α_1 и α_2, а также способом выражения и содержательной интерпретацией коэффициента α_0. Например, “классический” вариант функции (1.17) предполагает, что значения α_1 и α_2 удовлетворяют следующим ограничениям: $\alpha_1+\alpha_2=1$; α_1, $\alpha_2 \geq 0$. В других вариантах этой функции дополнительно вводят сомножитель $e^{\rho t}$, выражающий влияние на валовый продукт временного фактора, характеризующего научно-технический прогресс и т. п.

Функция (1.8) обычно используется в *предположении о постоянстве эластичности замещения одного фактора другим*. Например, если речь идет о замене фактора “труд”(L) фактором “капитал” (K), то значение коэффициента эластичности замещения показывает, на сколько процентов измениться капиталовооруженность (K/L) при изменении предельной нормы замещения труда капиталом ($N_{KL} = -dK/dL$) на 1% при условии, что зависимая переменная не изменится. Значения всех других факторов при этом предполагаются также неизменными. В общем случае, эластичность замещения i-го фактора j-м определяется выражением:

$$\acute{Y}_{ji} = \left(\frac{\partial N_{ji}}{\partial x_j / x_i} \cdot \frac{x_j / x_i}{N_{ji}} \right)^{-1}, \quad \dots\dots\dots\dots\dots\dots(1.18)$$

Предельная норма замещения i-го фактора j-м N_{ji} показывает количество j-го фактора, которое требуется для замены одной единицы i-го фактора при сохранении постоянных уровня зависимой переменной и значений остальных независимых переменных.

Проводя расчеты по формуле (1.18) для функции (1.8), получим, что для всех i и j и для всех значений t=1,2,..., T эластичность замещения прироста одного фактора соответствующим изменением другого является постоянной:

$$Э_{ij} = \frac{1}{1+\rho}. \quad (1.19)$$

Во многих практических исследованиях столь строгие теоретические концепции, предварительные допущения о содержательных сторонах взаимодействия между явлениями отступают на второй план. Для них главным является построение уравнения, достаточно точно выражающего взаимосвязи, адекватные тенденциям изменений переменных y и x_i, i=1, 2,...,n; на временном интервале (1,T). Более того, часто именно “удачная” форма уравнения эконометрической модели кладется в основу разрабатываемой теоретической концепции, которая затем находит свое применение в последующем анализе. Очевидно, что наиболее “подходящая” форма обеспечивает наилучшее приближение теоретических (расчетных) значений $\hat{y}_t = f(\boldsymbol{a}, \boldsymbol{x}_t)$ к действительным значениям y_t. Обычно выбор такой формы осуществляется на основе графического анализа тенденций развития соответствующих процессов.