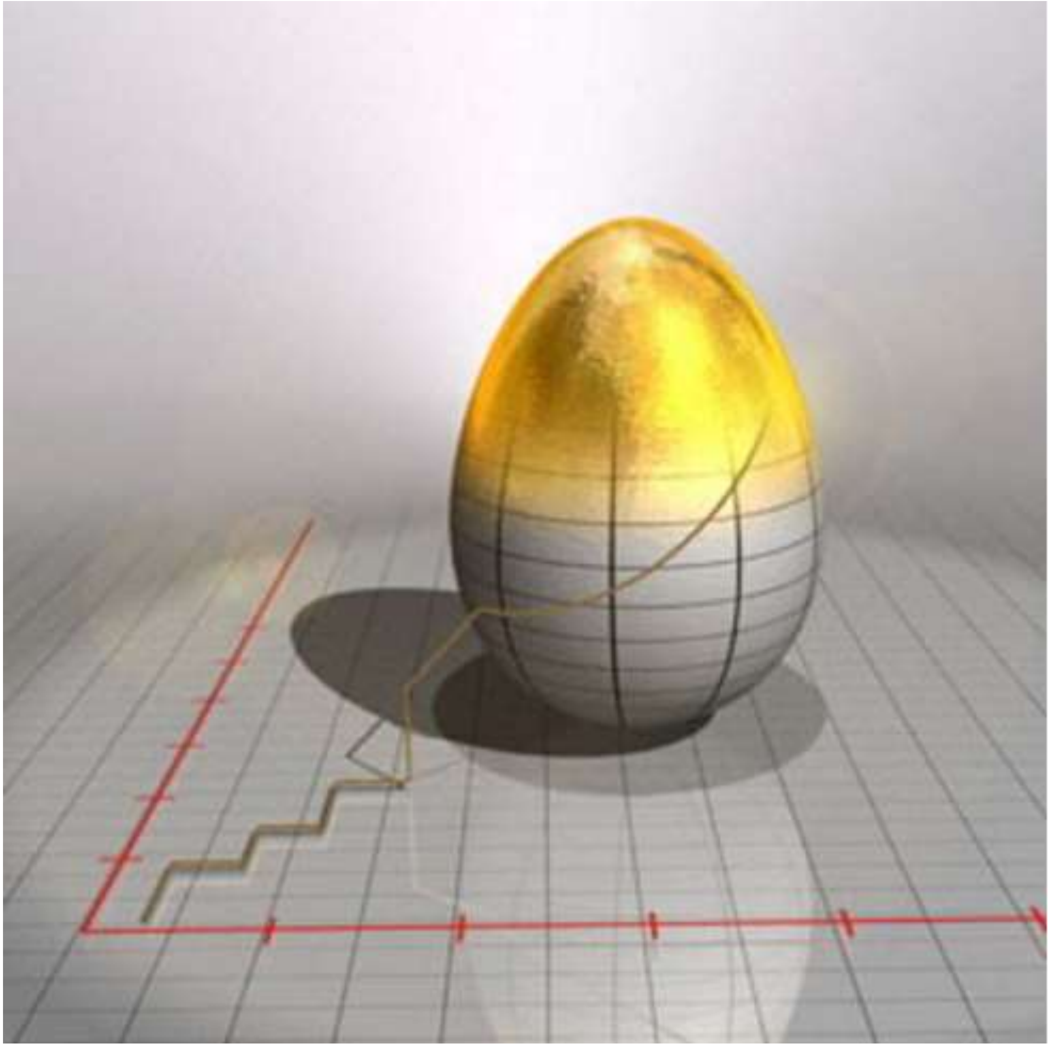
Ярославский государственный университет им. П.Г. Демидова
Факультет информатики и вычислительной техники

Е.М. Спиридонова

Эконометрика: материалы лекций



Ярославль – 2013

Спиридонова Е.М. **Эконометрика: материалы лекций**. Альбом наглядных пособий. − Ярославль: ЯрГУ им. П.Г. Демидова, 2013. – 57 с.

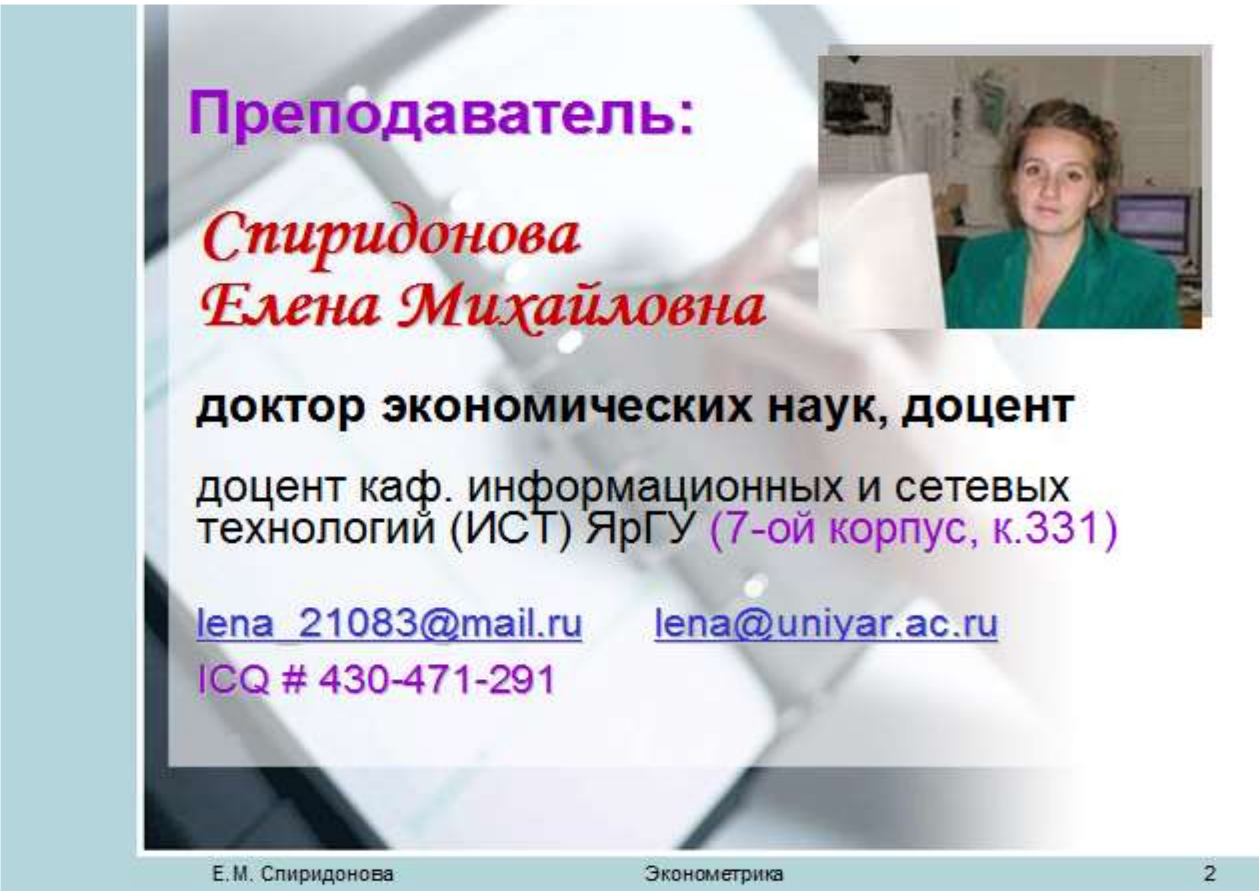
Учебное пособие представляет собой материалы презентаций, сопровождавших лекции по Эконометрике, прочитанные автором на факультете Информатики и вычислительной техники студентам 4-го курса специальности 080801.65 «Прикладная информатика (в экономике)» в 2010-2012 гг. Пособие предназначено для студентов направления 230700 «Прикладная информатика». Оно также может быть полезно студентам других, в первую очередь − экономических, направлений, изучающим Эконометрику.

В пособии в доступной и наглядной форме отображены основные понятия эконометрики, приведены формулы расчета параметров эконометрических моделей, описаны их свойства и возможные проблемы. Особое внимание уделяется оценке качества полученных моделей, их надежности и предсказательной способности.

Содержание:

Введение


ЭКОНОМЕТРИКА
"… искусство рисования кривой линии
от недоказанного предположения
до тривиального заключения"
Питер Кеннеди
("A Guide to Econometrics", 1992)


Преподаватель:
Спиридонова
Елена Михайловна
доктор экономических наук, доцент
доцент каф. информационных и сетевых
технологий (ИСТ) ЯрГУ (7-ой корпус, к.331)
lena_21083@mail.ru lena@uniyar.ac.ru
ICQ # 430-471-291
Е.М. Спиридонова Эконометрика 2

Цели и задачи:

- изучить основные свойства и закономерности отражения реальных процессов макро- и микро-экономики с помощью регрессионных моделей;
- овладеть приемами построения и расчета параметров эконометрических моделей и иметь навыки оценки их качественных характеристик с помощью специальных стат.пакетов;
- научиться прогнозировать развитие социально-экономических процессов, рассчитывать параметры зависимостей, делать выводы относительно качества полученных моделей, их надежности и предсказательной способности.

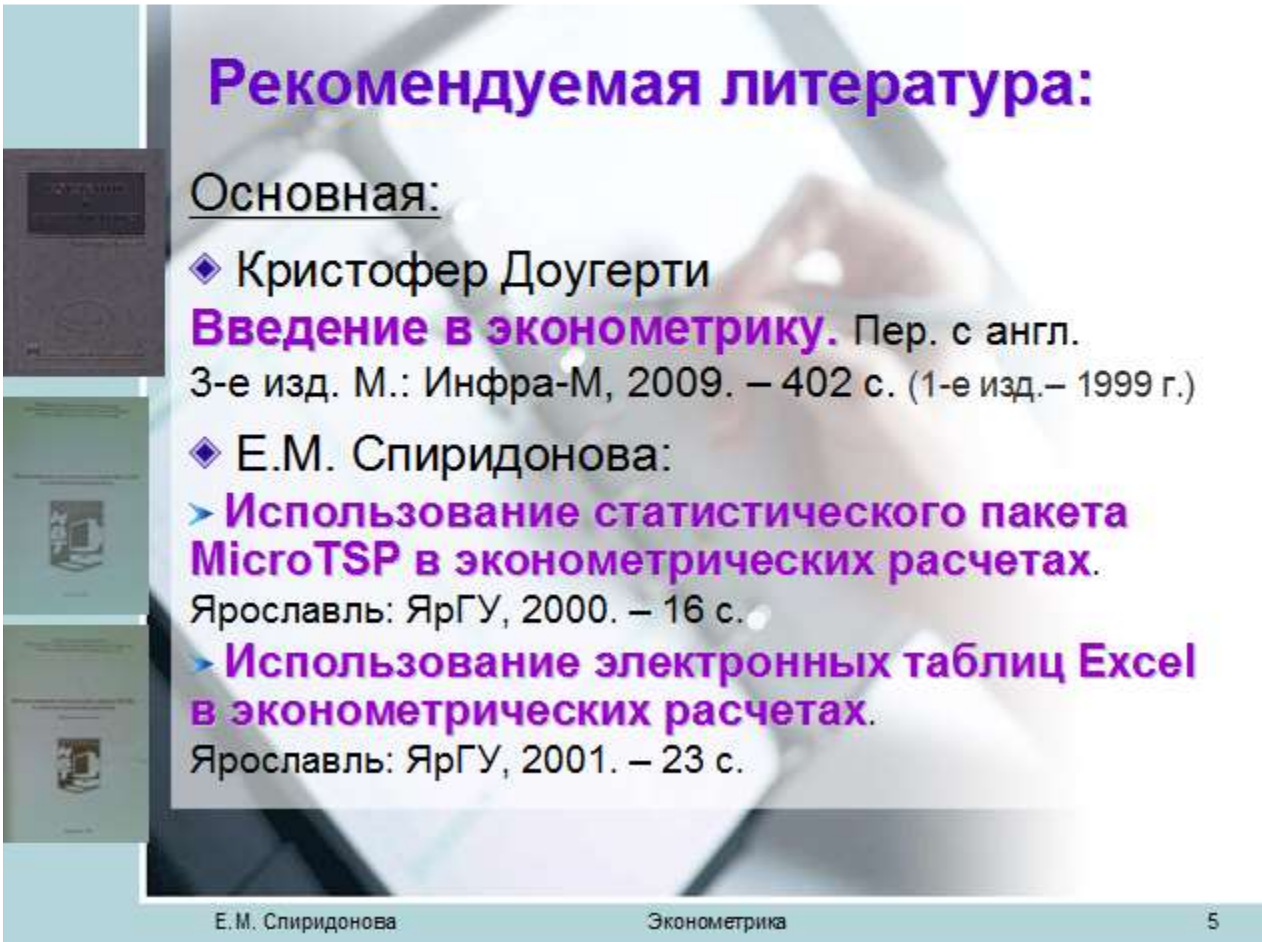
Е.М. Спиридонова Эконометрика 3

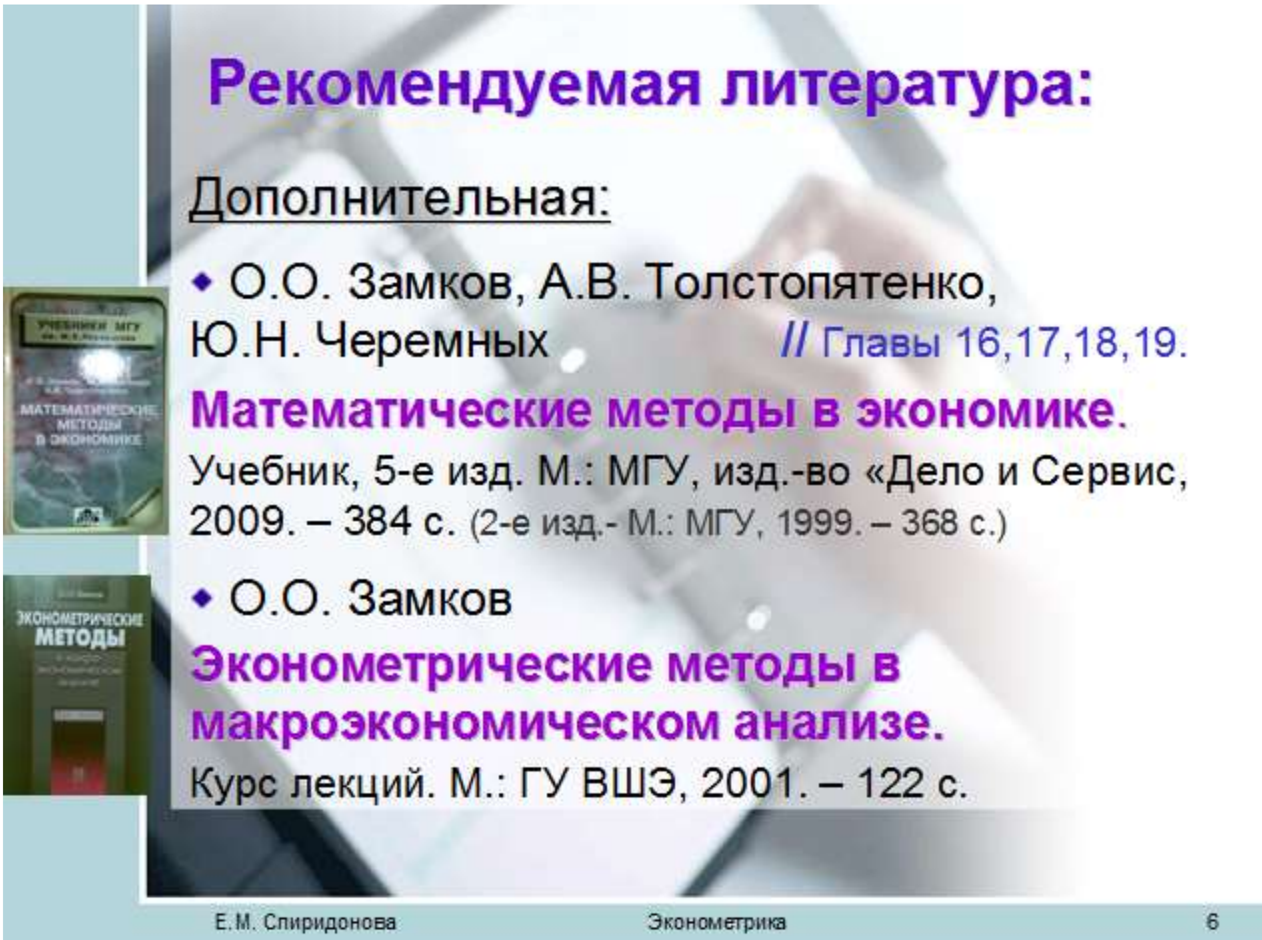
Тематический план:

№	Темы лекционных и практических занятий	Часы	
		л.	пр.
	Вводная лекция: определение, структура и сферы применения эконометрики	2	
1	Математико-статистические основы эконометрики	4	2
2	Регрессионный анализ	10	14
3	Проблемы регрессионного анализа	10	14
4	Продвинутый регрессионный анализ	10	6

1 2 3 4

Е.М. Спиридонова Эконометрика 4


Рекомендуемая литература:
Основная:
Кристофер Доугерти
Введение в эконометрику. Пер. с англ.
3-е изд. М.: Инфра-М, 2009. – 402 с. (1-е изд.– 1999 г.)
Е.М. Спиридонова:
Использование статистического пакета MicroTSP в эконометрических расчетах.
Ярославль: ЯрГУ, 2000. – 16 с.
Использование электронных таблиц Excel в эконометрических расчетах.
Ярославль: ЯрГУ, 2001. – 23 с.
Е.М. Спиридонова
Эконометрика
5


Рекомендуемая литература:
Дополнительная:
О.О. Замков, А.В. Толстопятенко, Ю.Н. Черемных
// Главы 16,17,18,19.
Математические методы в экономике.
Учебник, 5-е изд. М.: МГУ, изд.-во «Дело и Сервис, 2009. – 384 с. (2-е изд.- М.: МГУ, 1999. – 368 с.)
МАТЕМАТИЧЕСКИЕ МЕТОДЫ В ЭКОНОМИКЕ
О.О. Замков
Эконометрические методы в макроэкономическом анализе.
Курс лекций. М.: ГУ ВШЭ, 2001. – 122 с.
ЭКОНОМЕТРИЧЕСКИЕ МЕТОДЫ
Е.М. Спиридонова
Эконометрика
6

Рекомендуемая литература:

Дополнительная:

- С.А. Айвазян, В.С. Мхитарян

Прикладная статистика. Основы эконометрики. В 2-х томах. 2-е изд., испр. – М.: ЮНИТИ-ДАНА, 2001. – т.1: 656 с., т.2: 432 с.

- А.И. Орлов

Эконометрика. Учебник для вузов. 3-е изд. М.: Изд.-во «Экзамен», 2004. – 576 с.

Рекомендуемая литература:

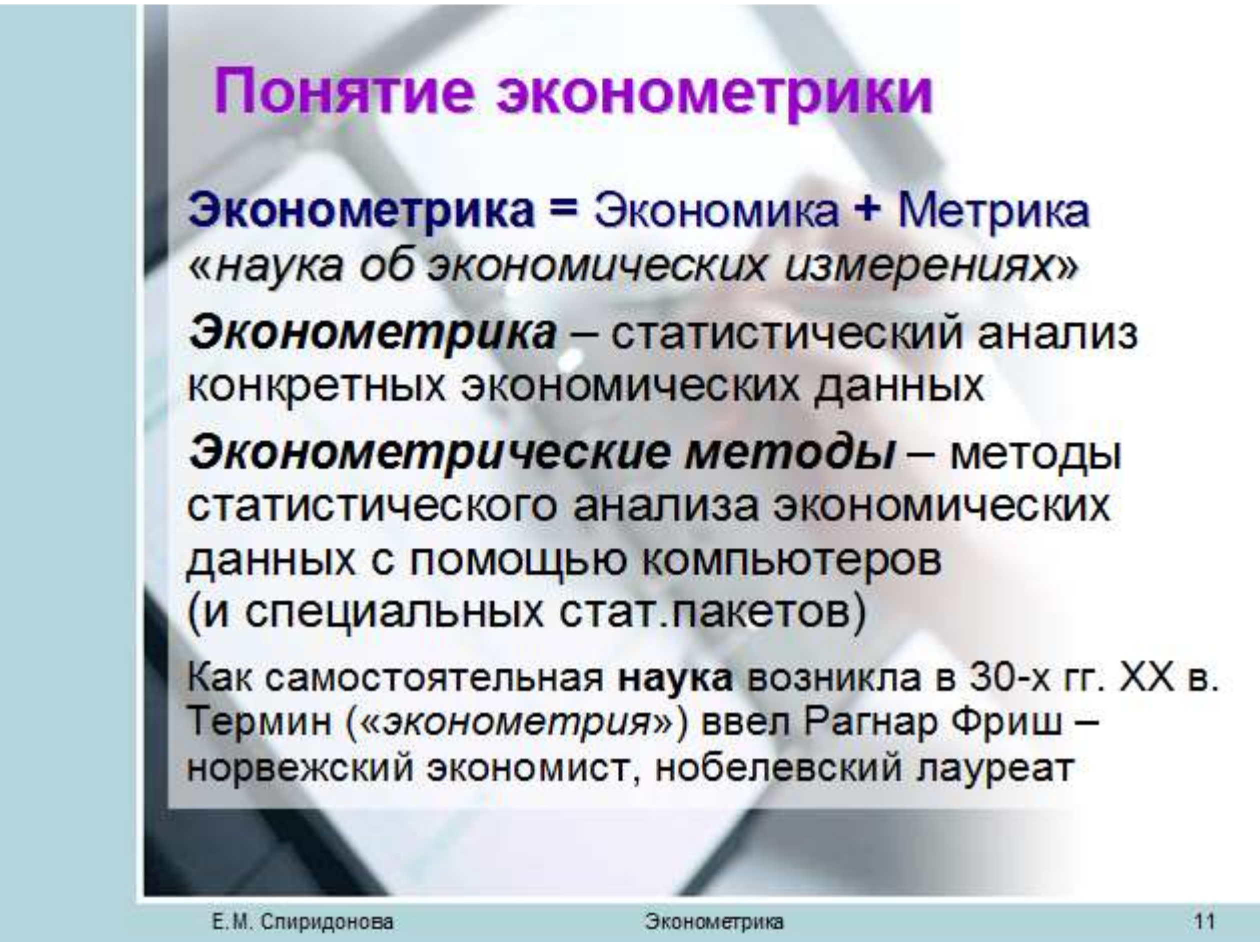
Дополнительная:

- Ю.Н. Тюрин, А.А. Макаров

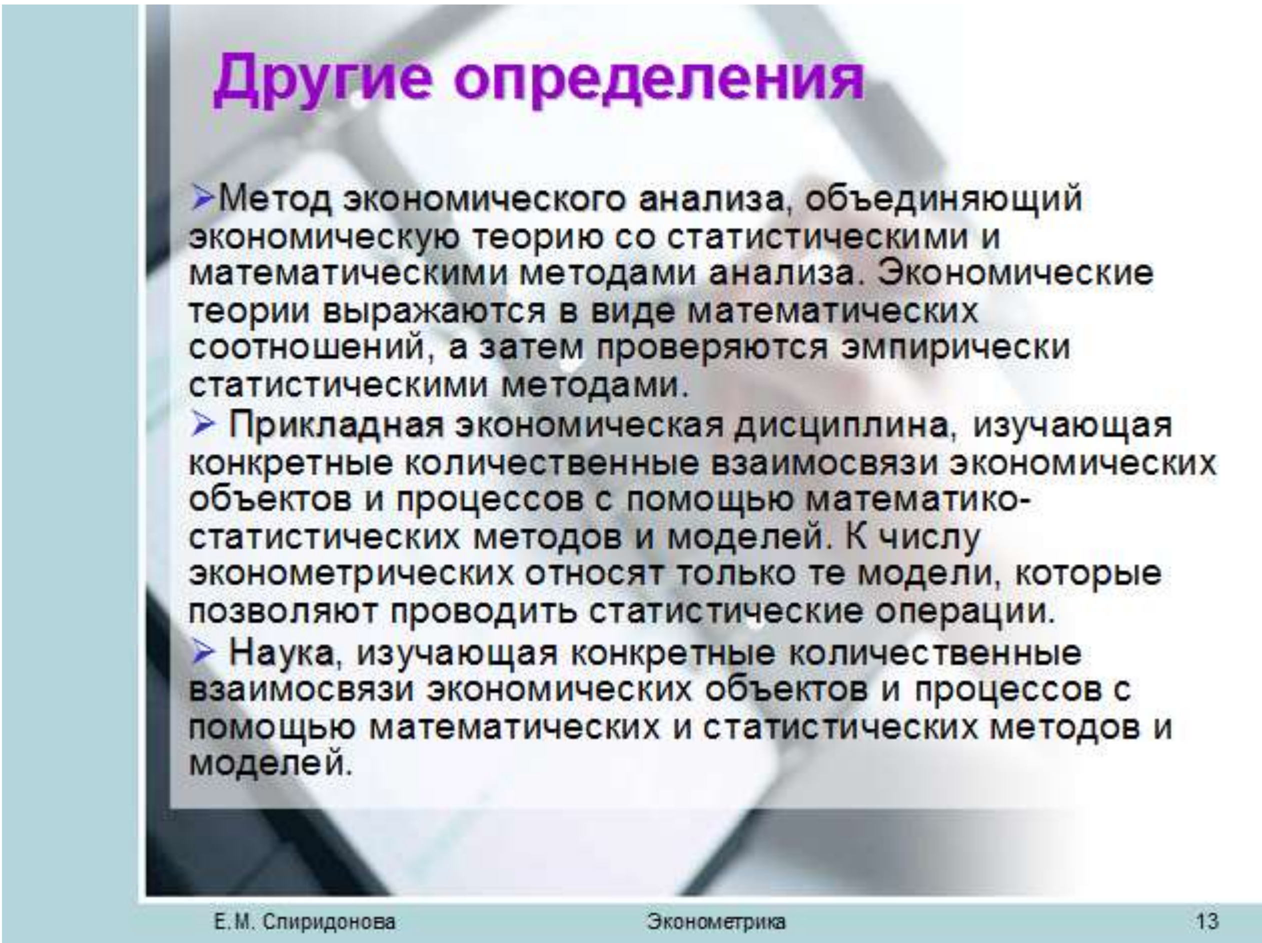
Статистический анализ данных на компьютере. Под ред. В.Э. Фигурнова 3-е изд., перераб. - М.: Инфра-М, 2003. – 544 с. (1-е изд. - М.: Инфра-М, 1998. – 528 с.)

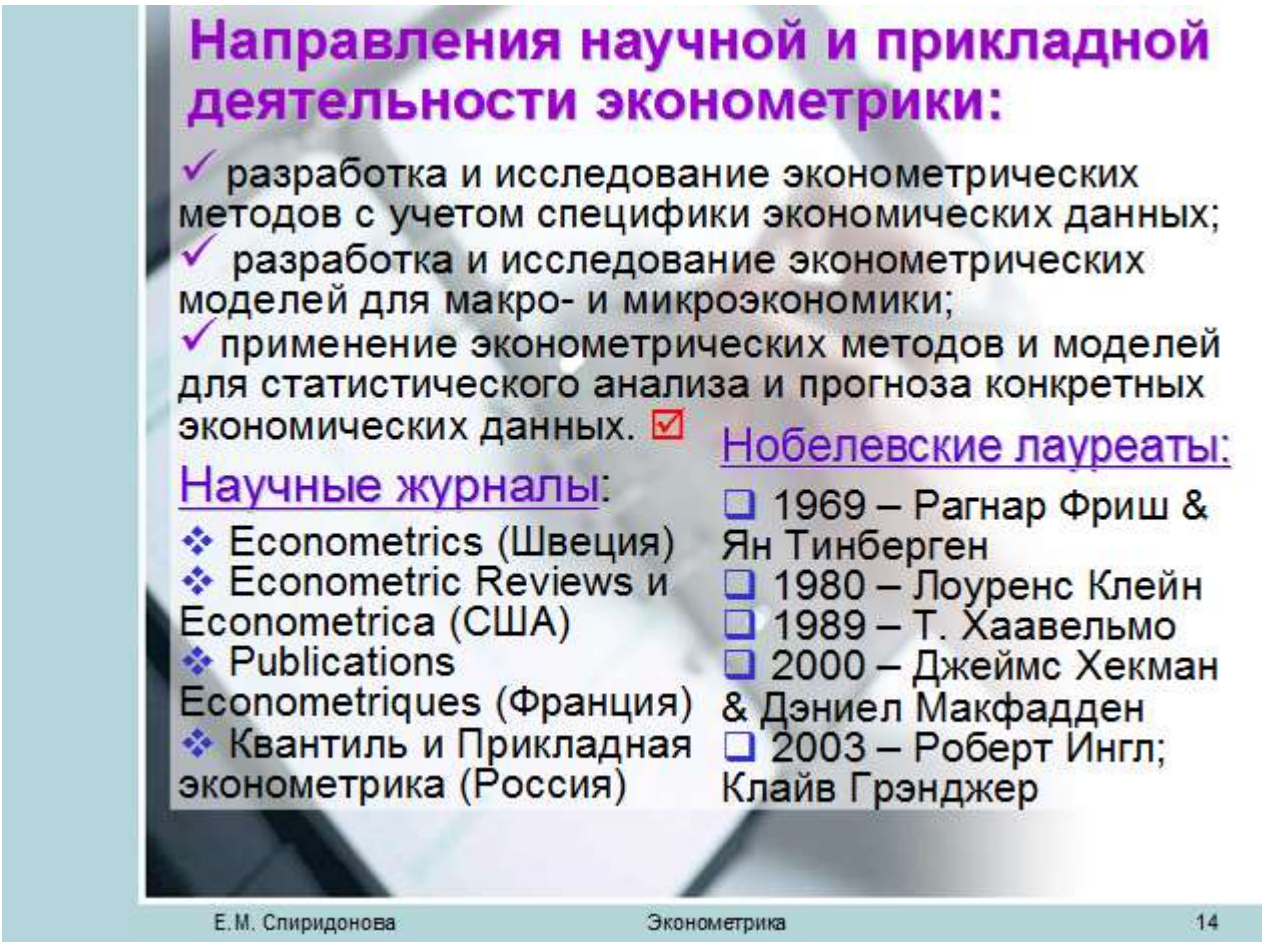
- А.А. Минько

Статистический анализ в MS Excel. М.: Издат.дом «Вильямс», 2004. – 448 с.


Понятие эконометрики
Эконометрика = Экономика + Метрика
«наука об экономических измерениях»
Эконометрика – статистический анализ конкретных экономических данных
Эконометрические методы – методы статистического анализа экономических данных с помощью компьютеров (и специальных стат.пакетов)
Как самостоятельная наука возникла в 30-х гг. XX в.
Термин («эконометрия») ввел Рагнар Фриш – норвежский экономист, нобелевский лауреат
Е.М. Спиридонова
Эконометрика
11


Определение эконометрики
Эконометрика – это самостоятельная научная дисциплина, объединяющая совокупность теоретических результатов, приемов, методов и моделей, предназначенных для того, чтобы на базе экономической теории, экономической статистики и математико-статистического инструментария придавать конкретное количественное выражение общим качественным закономерностям, обусловленным экономической теорией. (С.А.Айвазян «Прикладная статистика. Основы эконометрики», т.2)
ЭКОНОМЕТРИКА
Экономическая Теория (модели макро- и микроэкономики)
Социально-Экономическая Статистика (инф.обеспечение)
Теория Вероятностей и Математическая Статистика
Е.М. Спиридонова
Эконометрика
12


Другие определения
➢Метод экономического анализа, объединяющий экономическую теорию со статистическими и математическими методами анализа. Экономические теории выражаются в виде математических соотношений, а затем проверяются эмпирически статистическими методами.
➢ Прикладная экономическая дисциплина, изучающая конкретные количественные взаимосвязи экономических объектов и процессов с помощью математико-статистических методов и моделей. К числу эконометрических относят только те модели, которые позволяют проводить статистические операции.
➢ Наука, изучающая конкретные количественные взаимосвязи экономических объектов и процессов с помощью математических и статистических методов и моделей.
Е.М. Спиридонова
Эконометрика
13


Направления научной и прикладной деятельности эконометрики:
✓ разработка и исследование эконометрических методов с учетом специфики экономических данных;
✓ разработка и исследование эконометрических моделей для макро- и микроэкономики;
✓ применение эконометрических методов и моделей для статистического анализа и прогноза конкретных экономических данных.
Научные журналы:
❖ Econometrics (Швеция)
❖ Econometric Reviews и Econometrica (США)
❖ Publications Econometriques (Франция)
❖ Квантиль и Прикладная эконометрика (Россия)
Нобелевские лауреаты:
❑ 1969 – Рагнар Фриш & Ян Тинберген
❑ 1980 – Лоуренс Клейн
❑ 1989 – Т. Хаавельмо
❑ 2000 – Джеймс Хекман & Дэниел Макфадден
❑ 2003 – Роберт Ингл; Клайв Грэнджер
Е.М. Спиридонова
Эконометрика
14

Сферы применения прикладной эконометрики

- ✓ Управление на макро-, мезо- и микроуровне – анализ экономических показателей, прогнозы развития
- ✓ Стат.методы управления качеством продукции и услуг
- ✓ Анализ рисков различной природы (от экономических до экологических), разработка правил страхования
- ✓ Логистика – анализ и моделирование движения материальных (спрос-поставки-запасы), финансовых и информационных потоков
- ✓ Анализ стабильности (⇒↓доли дефектной продукции) и оптимизация (⇒↑выход полезного продукта) технологических процессов методами экстремального планирования
- ✓ Маркетинг – изучение предпочтений потребителей
- ✓ Построение всевозможных рейтингов и «индексов»
- ✓ и т.д. и т.п.,

в т.ч. – для анализа данных нечисловой природы в гуманитарных науках *(см. работы академика А.Т.Фоменко)*

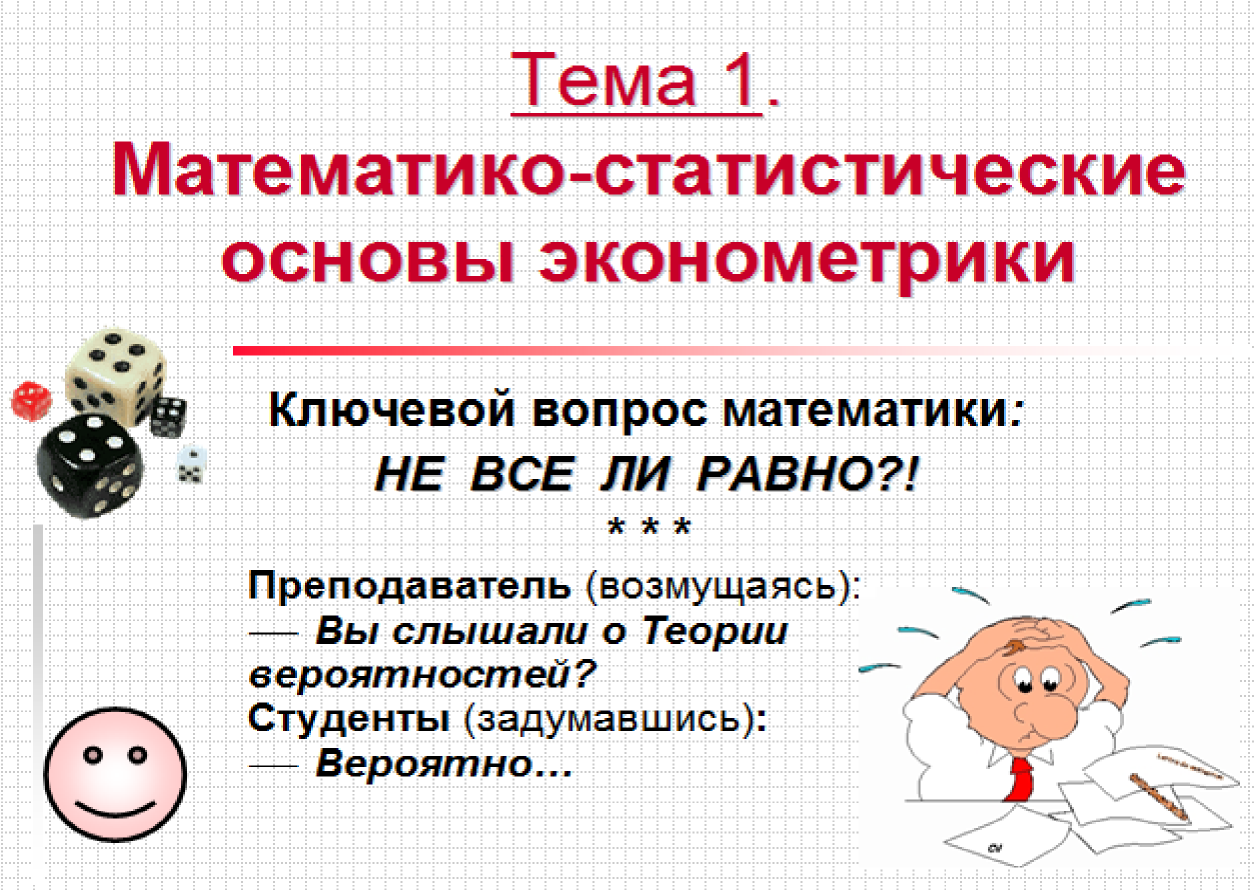
Е.М. Спиридонова | Экономертика | 15

Структура полного курса Эконометрики (*проф. А.И.Орлов*)

1. **Современная эконометрика.** Взаимодействие с прикладной статистикой, мат.статистикой, мат.методами и т.п. Структура и сферы применения. «Точки роста» науки: непараметрика, робастность, бутстреп, интервальные данные, нечисловые данные.
2. **Современная теория измерений.** Шкалы и рейтинги.
3. **Анализ числовых величин и непараметрическая статистика.** Оценивание характеристик распределения и доверительных интервалов для них. *(частично рассм. в ОТС)* Проверка однородности двух выборок.
4. **Многомерный статистический анализ.** ☑ Корреляция и (линейная) регрессия. Дисперсионный анализ. *(част. - в ОТС)* Факторный анализ; кластерный и дискриминантный анализ.
5. **Анализ и прогноз временных рядов.** ☑ Тренды. Авторегрессия. Системы эконометрических уравнений.
6. **Статистика нечисловых данных.**
7. **Статистика интервальных данных.**
8. **Устойчивость статистических процедур.**
9. **Эконометрические методы экспертных исследований.**
10. **Эконометрические методы управления качеством.**

Е.М. Спиридонова | Эконометрика | 16

ТЕМА 1. Математико-статистические основы эконометрики



§1. Случайные переменные и оценки

Основные определения из ТВиМС

1. ***Случайная переменная* (*сл.п.*)** – любая переменная, значение которой не может быть точно предсказано
   - ***Дискретная сл.п.*** – имеет определенный набор возможных значений
   - ***Непрерывная сл.п.*** – может принимать любое значение из определенного диапазона

Тема 1 Эконометрика 3

§1. Случайные переменные и оценки

2. ***Мат.ожидание дискретной сл.п.:*** $$E(x)=\sum_{i=1}^{n} x_i p_i \equiv \mu$$

Мат.ожидание функции: $$E(g(x))=\sum g(x_i)p_i$$

Правила расчета мат.ожидания:

$$E(x+y+z)=E(x)+E(y)+E(z)$$

$$c\equiv const \Rightarrow E(cx)=c\cdot E(x);\ E(c)=c$$

Тема 1 Эконометрика 4

§1. Случайные переменные и оценки

3. *Теоретич. дисперсия дискретной сл.п.:*

$$pop.\mathrm{var}(x) = E\left((x-\mu)^2\right) = \sum_{i=1}^{n}(x_i-\mu)^2 p_i \equiv \sigma^2$$

$$\sigma^2 = E(x^2) - \left(E(x)\right)^2 = E(x^2) - \mu^2$$

4. *Независимость случайных переменных:*

$$E\left(g(x)\cdot q(y)\right) = E\left(g(x)\right)\cdot E\left(q(y)\right)$$

$$\left\langle \text{или } E(x\cdot y) = E(x)\cdot E(y)\right\rangle \Leftrightarrow X \text{ и } Y - \text{независимы}$$

Тема 1 Эконометрика 5

§1. Случайные переменные и оценки

5. *Плотность распределения f(x)*
(англ.: «*функция плотности вероятности*»)
- только для ***непрерывных сл.п.!***

Мат.ожидание: $\mu = \int x\cdot f(x)\cdot dx$

Дисперсия: $\sigma^2 = \int (x-\mu)^2\cdot f(x)\cdot dx$

Тема 1 Эконометрика 6

§1. Случайные переменные и оценки

6. ***Постоянная и случайная составляющие сл.п.:*** $x = \mu + u$

u – чисто случайная составляющая; мат.ожидание и дисперсия:

$$E(u) = 0\,; \quad pop.\mathrm{var}\,(u) = E(u^2) = \sigma_u^2$$

$$\Rightarrow \quad E(x) = \mu\,; \quad pop.\mathrm{var}\,(x) = \sigma_u^2$$

§1. Случайные переменные и оценки

7. ***Оценки и способы оценивания:***

$$\text{Матожидание } \mu \quad - \quad \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \ (\text{ср.арифм.})$$

$$\text{Дисперсия } \sigma^2 \quad - \quad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

Оценки – тоже сл.п.: $\bar{x} = \mu + \bar{u}$

§1. Случайные переменные и оценки

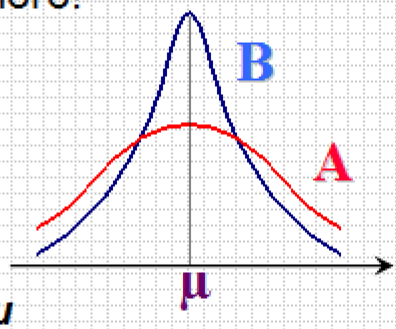
8. Свойства оценок:

A. ***Несмещенность*** – мат.ожидание оценки равно соотв. характеристике генеральной совокупности.

$\bar{x}$ – несмещ.оц. μ, S^2 – несмещ.оц. σ^2

Несмещенных оценок для μ и σ^2 – много!

B. ***Эффективность*** – несмещ.оц.***B*** считается более эффективной, чем несмещ.оц.***A***, если функция плотности ***B*** – более сжата.



$\bar{x}$ – самая эфф. из всех несмещ.оц. μ

§1. Случайные переменные и оценки

8. Свойства оценок:

A vs **B**? Какая оценка лучше – несмещенная или эффективная (т.е. с минимальной дисперсией)? Выбор – осуществляется по «функции потерь».

C. ***Состоятельность*** – «предел по вероятности» оценки равен соответствующей характеристике генеральной совокупности.

$\bar{x}$ - состоятельная оценка: $plim\, \bar{x} = \mu$

В большинстве случаев несмещенная оценка является состоятельной, но есть «искусственные» контрпримеры!

§2. Анализ вариации и взаимосвязи

1. **Дисперсия** (для характеристики вариации):
   - ***теоретическая* – *pop.var(x)***
   - ***выборочная* – *Var(x)***

$$pop.\mathrm{var}(x) = E\left(\left(x - E(x)\right)^2\right)$$

$$Var(x) = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$$

Var(x) - смещенная, но состоятельная оценка ***pop.var(x)***

Тема 1 Эконометрика 11

§2. Анализ вариации и взаимосвязи

<u>Свойства дисперсии</u>:

$$y = x + z \Rightarrow Var(y) = Var(x) + Var(z) + 2Cov(x,z)$$

$$y = c \cdot x,\ c \equiv const \quad \Rightarrow \quad Var(y) = c^2 \cdot Var(x)$$

$$y = x + c,\ c \equiv const \quad \Rightarrow \quad Var(y) = Var(x)$$

$$y = c,\ c \equiv const \quad \Rightarrow \quad Var(y) = 0$$

$$Var(x) \equiv Cov(x,x)$$

Тема 1 Эконометрика 12

§2. Анализ вариации и взаимосвязи

2. Ковариация (для хар.-ки направления связи):

- ***теоретическая – pop.cov(x)***
- ***выборочная – Cov(x)***

$$pop.\mathrm{cov}(x,y)=E\big((x-E(x))(y-E(y))\big)$$

$$Cov(x)=\frac{1}{n}\sum_{i=1}^{n}\big((x_i-\bar{x})(y_i-\bar{y})\big)=\frac{1}{n}\sum_{i=1}^{n}x_i y_i-\bar{x}\cdot\bar{y}$$

Cov(x) - смещенная, но состоятельная оценка ***pop.cov(x)***

§2. Анализ вариации и взаимосвязи

Свойства ковариации:

$$y=w+z\Rightarrow Cov(x,y)=Cov(x,w)+Cov(x,z)$$

$$y=c\cdot z,\ c\equiv const\quad\Rightarrow\quad Cov(x,y)=c\cdot Cov(x,z)$$

$$y=c,\ c\equiv const\quad\Rightarrow\quad Cov(x,y)=0$$

$$y=a+b\cdot z\quad\Rightarrow\quad Cov(x,y)=b\cdot Cov(x,z)$$

Недостаток: ковариация имеет размерность и зависит от единиц измерения $\Rightarrow$ НЕ характеризует СИЛУ связи.

§2. Анализ вариации и взаимосвязи

3. Корреляция (характеризует силу связи):

3.1. В случае **парной** зависимости ***Y(X)*** исп. ***(парный) линейный коэфф.-т корреляции***

3.2. В случае **множественной** зависимости ***Y(X₁, X₂, …, X_k)*** исп.:

- ***множественный коэфф.-т корреляции***
- ***частные коэфф.-ты корреляции***

Тема 1 Эконометрика 15

§2. Анализ вариации и взаимосвязи

3.1. *Парный линейный к.-т корреляции:*

- ***теоретический – ρ***
- ***выборочный – r***

$$\rho_{xy} = \frac{pop.\mathrm{cov}(x,y)}{\sqrt{pop.\mathrm{var}(x)\cdot pop.\mathrm{var}(y)}}$$

$$r_{xy} = \frac{\frac{n}{n-1}Cov(x,y)}{\sqrt{\frac{n}{n-1}Var(x)\frac{n}{n-1}Var(y)}} = \frac{Cov(x,y)}{\sqrt{Var(x)Var(y)}}$$

$\boldsymbol{r_{xy}}$ – НЕсмещенная (!) и состоятельная оценка $\boldsymbol{\rho_{xy}}$

Тема 1 Эконометрика 16

§2. Анализ вариации и взаимосвязи

Расчетные формулы:

$$r_{xy} = \frac{\frac{1}{n}\sum(x_i - \bar{x})(y_i - \bar{y})}{S_x \cdot S_y} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \cdot \sum(y_i - \bar{y})^2}}$$

$$\bar{x} = \frac{1}{n}\sum x_i \quad \bar{y} = \frac{1}{n}\sum y_i \quad S_x = \sqrt{\frac{1}{n}\sum(x_i - \bar{x})^2} \quad S_y = \sqrt{\frac{1}{n}\sum(y_i - \bar{y})^2}$$

$$r_{xy} = \frac{n\sum xy - \sum x \sum y}{\sqrt{\left(n\sum x^2 - (\sum x)^2\right)\cdot\left(n\sum y^2 - (\sum y)^2\right)}}$$

Тема 1 Эконометрика 17

§2. Анализ вариации и взаимосвязи

Характеристики коэффициента корреляции:

- $-1 \le \rho \le +1$
- Если $\rho < 0 \Rightarrow$ связь обратная,
- если $\rho > 0 \Rightarrow$ связь прямая.
- Чем ближе $|\rho|$ к **0**, тем связь слабее,
- чем ближе $|\rho|$ к **1**, тем связь сильнее.

$\rho = 0 \not\Leftrightarrow r = 0$

Если X и Y – независимы $\Rightarrow \rho = 0$.

Если $r = 0 \not\Rightarrow X$ и Y – независимы.

Если $Y = a + bX$, то: при $b > 0 \Rightarrow r = +1$; при $b < 0 \Rightarrow r = -1$

Тема 1 Эконометрика 18

§2. Анализ вариации и взаимосвязи

3.2. Оценки *множественного* (*R*) и *частного* (*r*) *коэффициентов корреляции.*

Рассм. совокупность ***m*** факторов: $X_1, X_2, \ldots, X_m$

Строится матрица парных линейных к.-тов корреляции:

$$R_{i\,(от\ всех\ ост.)} = \sqrt{1 - \frac{|\hat{R}|}{\hat{R}_{ii}}} \qquad \hat{R} = \begin{pmatrix} 1 & r_{12} & \ldots & r_{1m} \\ r_{12} & 1 & \ldots & r_{2m} \\ \ldots & \ldots & \ldots & \ldots \\ r_{1m} & r_{2m} & \ldots & 1 \end{pmatrix}$$

$$r_{ij\,(искл.\ все\ ост.)} = -\frac{\hat{R}_{ij}}{\sqrt{\hat{R}_{ii}\hat{R}_{jj}}}$$

M_{ij} – минор – определитель матрицы, получ. из $\hat{R}$ вычеркив. *i*-стр. и *j*-стлб.

где $\hat{R}_{ij} = (-1)^{i+j} M_{ij}$ $\quad$ $\hat{R}_{ij}$ – алгебраическое дополнение элемента r_{ij}

§2. Анализ вариации и взаимосвязи

Для 3-хмерной совокупности факторов {*X*,*Y*,*Z*}, где *Y*=*f*(*X*,*Z*):

$$R_{y(x,z)} = \sqrt{1 - \frac{|\hat{R}|}{\hat{R}_{yy}}} = \sqrt{\frac{r_{yx}^2 + r_{yz}^2 - 2 \cdot r_{yx} \cdot r_{yz} \cdot r_{xz}}{1 - r_{xz}^2}}$$

$$r_{yx/z} = -\frac{\hat{R}_{yx}}{\sqrt{\hat{R}_{yy}\hat{R}_{xx}}} = \frac{r_{yx} - r_{xz} \cdot r_{yz}}{\sqrt{(1 - r_{xz})^2 (1 - r_{yz})^2}}$$

Характеристики: **множеств. 0 ≤ *R* ≤ 1** – сила связи;
парные и **частные (–1) ≤ *r* ≤ +1** – сила и направление.

§2. Анализ вариации и взаимосвязи

Проверка значимости коэффициентов корреляции:

- Для ***парных*** и ***частных*** -

***t*-статистика** (*распределение Стьюдента*)

$$t=\frac{r}{\sqrt{1-r^2}}\sqrt{n-l-2}$$

где n – объем наблюдений,
l – порядок частного коэфф.-та
(= число исключенных факторов).
Для парных коэфф.-тов $l=0$

Если $t>t_{крит}$ (*α; ч.с.с.=n-l-2*), то с вероятностью **P=1–*α*** гипотеза об отсутствии связи (H_0: $\rho=0$) отвергается, связь – не случайна.

Тема 1 Эконометрика 21

§2. Анализ вариации и взаимосвязи

Проверка значимости коэффициентов корреляции:

- Для ***множественного*** -

***F*-статистика** (*распределение Фишера-Снедекора*)

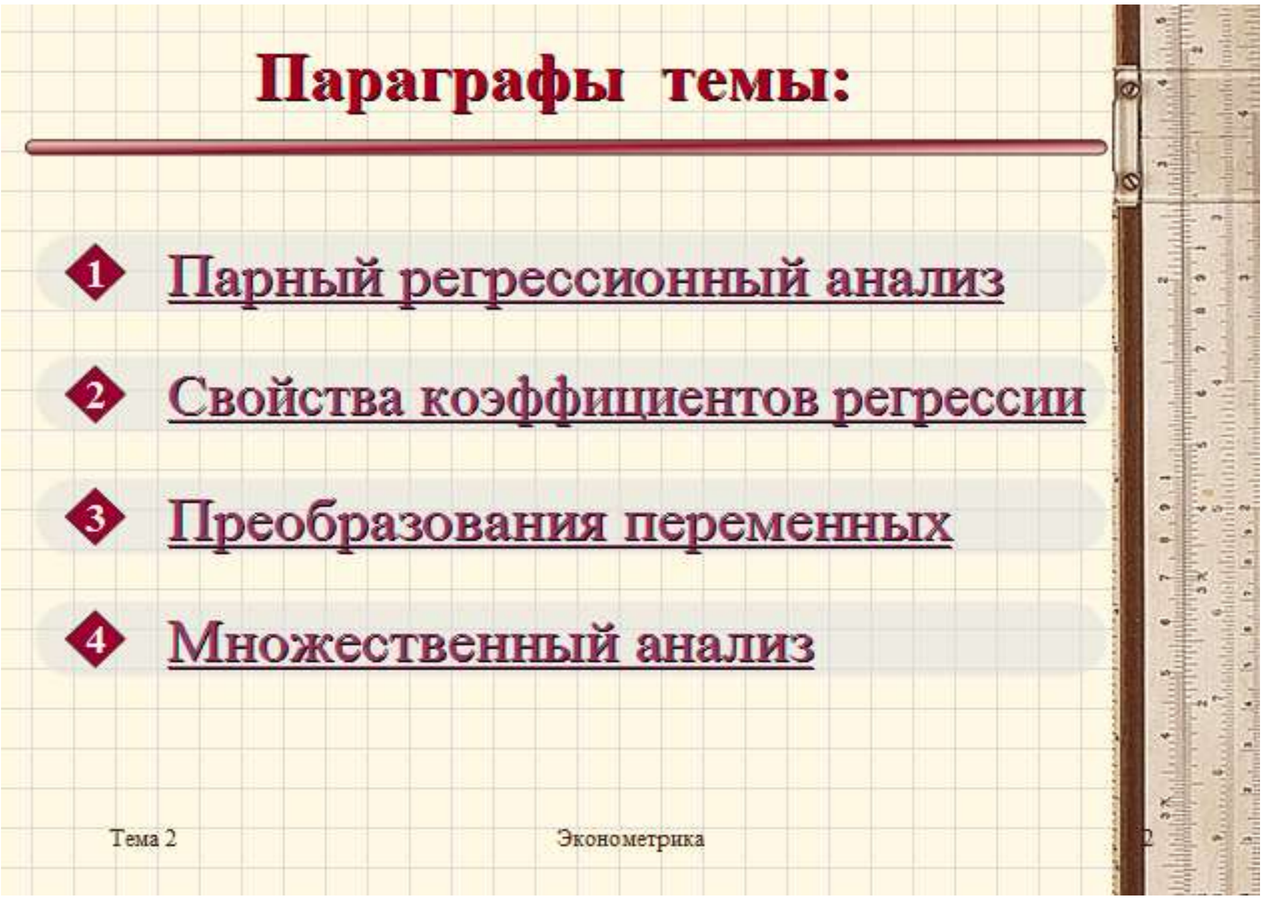
$$F=\frac{R^2_{y(x_1,\ldots,x_k)}}{k}\div\frac{1-R^2_{y(x_1,\ldots,x_k)}}{n-k-1}$$

где n – объем наблюдений,
k – число факторных признаков.

Если $F>F_{крит}$ (*α; ч.с.с.$_1$=k, ч.с.с.$_2$=n-k-1*), то с вероятностью **P=1–*α*** гипотеза об отсутствии связи (H_0: $\rho=0$) отвергается, связь – не случайна.

Тема 1 Эконометрика 22

ТЕМА 2. Регрессионный анализ



§1. Парный регрессионный анализ

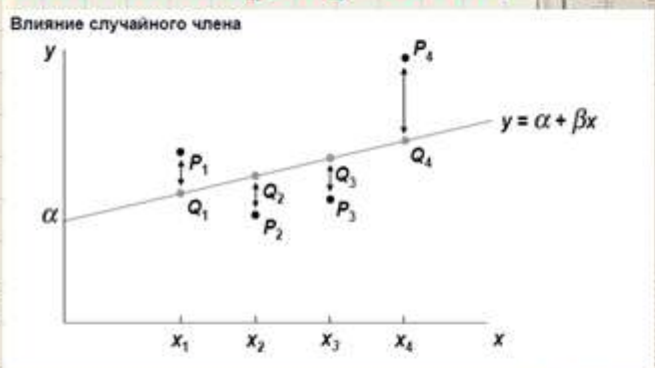
План лекции:

- 1. Модель парной линейной регрессии
- 2. Оценка парной линейной регрессии методом наименьших квадратов
- 3. Качество оценки: коэффициент детерминации

1. Модель парной линейной регрессии

$$y = \alpha + \beta x + u$$

где: **Y** – зависимая (результативная) переменная,
X – независимая (факторная) переменная;
α и **β** – параметры регрессии;
u – ***случайный член***.



Причины существования u:

- Невключение объясняющих переменных
- Агрегирование переменных
- Неправильная структура модели
- Неправильная функциональная зависимость
- Ошибки измерения

2. Оценка парной линейной регрессии методом наименьших квадратов

Уравнение: $\hat{y} = a + b\,x$

(**a** – оценка α, **b** – оценка β)

Отклонения (ошибки): $e_i = y_i - \hat{y}_i = y_i - (a + b\,x_i)$

Метод наименьших квадратов (МНК):

$$S = \sum e_i^2 = \sum (y_i - a - b\,x_i)^2 \to min$$

$$\begin{cases} \dfrac{\partial S}{\partial a} = 0 \\ \dfrac{\partial S}{\partial b} = 0 \end{cases} \Rightarrow \begin{cases} na + b\sum x_i - \sum y_i = 0 \\ a\sum x_i - b\sum x_i^2 - \sum x_i y_i = 0 \end{cases} \Rightarrow \begin{cases} a = \dfrac{\sum y_i - b\sum x_i}{n} \\ b = \dfrac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - (\sum x_i)^2} \end{cases}$$

2. Оценка парной линейной регрессии Интерпретация моделей

➢ Уравнение **регрессии**: $\hat{y} = a + b\,x$

✓ Коэффициент **b** показывает на сколько своих единиц в среднем изменится **Y** при увеличении **X** на одну свою единицу;

✓ Коэффициент **a** может не иметь смысла (определяет положение прямой на плоскости).

➢ Уравнение **динамики**: $\hat{y} = a + b\,t$

где **t – временной параметр** (порядковый номер периода)

✓ Коэффициент **b** показывает на сколько своих единиц изменится **Y** в среднем за период;

✓ Если **t = {1, 2,…, n}**, то коэффициент **a** показывает оценочное значение **Y** до начала исследуемого периода; в других случаях коэффициент **a** может не иметь смысла.

2. Оценка парной линейной регрессии

Средства построения уравнений

В специальных статистических пакетах:

- **STATISTICA**
- **Micro TSP (EViews)**

В электронных таблицах **MS Excel**:

- **стандартные средства:**
  - f_x: Статистические. **ЛИНЕЙН**
  - **график:** Диаграмма. **Точечная**. Добавить линию тренда. ☑ Показать уравнение
- специальные (статистические) средства:

меню **Сервис. Надстройки. Пакет анализа**

Тема 2 Эконометрика

3. Качество оценивания: коэффициент детерминации

$\hat{y}_i = a + b\,x_i$ ← «теоретические» значения

фактические значения → $y_i = \hat{y}_i + e_i$ (e_i – «остатки»)

Правило сложения дисперсий: $Var(y) = Var(\hat{y}) + Var(e)$

Коэффициент детерминации

$$R^2 = \frac{Var(\hat{y})}{Var(y)} = 1 - \frac{Var(e)}{Var(y)} = 1 - \frac{\sum e_i^2}{\sum (y_i - \bar{y})^2} \qquad 0 \le R^2 \le 1$$

Характеризует «объясняющую способность» модели!

Интерпретация: вариация Y на $\{R^2\}$ обуславливается вариацией X, а на $\{1-R^2\}$ – прочими факторами.

Тема 2 Эконометрика

§2. Свойства коэффициентов регрессии. Проверка гипотез

План лекции:

- 1. Постоянные и случайные составляющие коэффициентов регрессии
- 2. Предположения о случайном члене
- 3. Надежность коэффициентов регрессии
- 4. Проверка гипотез: t-тесты
- 5. F-тест на качество оценивания
- 6. Взаимосвязь между критериями

Тема 2 Эконометрика 9

1. Постоянные и случайные составляющие коэффициентов регрессии

$$Cov(x,y) = Cov(x,[\alpha + \beta x + u]) = Cov(x,\alpha) +$$
$$+ Cov(x,\beta x) + Cov(x,u) = \beta Var(x) + Cov(x,u)$$

Коэффициент угла наклона *b*

$$b = \frac{Cov(x,y)}{Var(x)} = \beta + \frac{Cov(x,u)}{Var(x)}$$

постоянная составляющая

случайная составляющая

Аналогично: **a = α + ...**

Истинные значение ***α*** и ***β*** неизвестны.
Чтобы определить, насколько хороши оценки ***a*** и ***b***, использ. метод Монте-Карло.

Тема 2 Эконометрика 10

2. Предположения о случайном члене

Чтобы регрессионный анализ по МНК давал лучшие результаты необходимо, чтобы случайный член ***u*** удовлетворял **условиям Гаусса-Маркова**:

- **1-е условие Г.-М.**: $M(u)=0$ (при наличии в уравнении свободного члена ***a*** выполняется автоматически)
- **2-е условие Г.-М.**: $D(u)=\sigma^2$ (дисперсия случайного члена постоянна – «*гомоскедастичность*»)
- **3-е условие Г.-М.**: $Cov(u_i,u_j)=0$ ***для всех*** $i\neq j$ (остатки во всех наблюдениях статистически независимы)
- **4-е условие Г.-М.**: $Cov(x,u)=0$ (объясняющие переменные и случайный член – независимы)
- **предположение о нормальности**: $u \sim N(0,\sigma)$ основано на **ЦПТ** (см. **ТВиМС**); проблем не возникает

Тема 2 Эконометрика

3. Надежность коэффициентов регрессии

Необходимо определить:

- являются ли оценки к.-тов регрессии несмещенными?
- какова их точность (эффективны ли они)?

Несмещенность

$$E(b)=E\left(\beta+\frac{Cov(x,u)}{Var(x)}\right)=\beta+E\left(\frac{Cov(x,u)}{Var(x)}\right)$$

$$\text{если } X - \text{не случ., то } E\big(Cov(x,u)\big)=0 \quad \Rightarrow \quad E(b)=\beta$$

$$E(a)=E(\bar{y})-E(b\cdot\bar{x})=E(\alpha+\beta\cdot\bar{x})-\bar{x}\cdot E(b)=(\alpha+\beta\cdot\bar{x})-\bar{x}\cdot\beta=\alpha$$

✓ если выполнено 4-е условие Гаусса-Маркова, то ***b*** – **несмещенная оценка *β***

✓ если выполнены 1-е и 4-е условия Гаусса-Маркова, то ***a*** – **несмещенная оценка α**

Тема 2 Эконометрика 12

3. Надежность коэффициентов регрессии

Эффективность

Док.-во: 1983, Дж.Томас (Thomas)

$$pop.\mathrm{var}\,(a) = \frac{\sigma_u^2}{n}\left(1+\frac{\overline{x}^2}{Var(x)}\right) \qquad pop.\mathrm{var}\,(b) = \frac{\sigma_u^2}{n\cdot Var(x)}$$

σ_u^2 *– неизвестна, её несмещенная оценка*: $S_u^2 = \dfrac{n}{n-2}Var(e)$

$\Rightarrow$ **стандартные ошибки** (оценки станд.отклонений):

$$S.e.(a) = \sqrt{\frac{S_u^2}{n}\left(1+\frac{\overline{x}^2}{Var(x)}\right)} \qquad S.e.(b) = \sqrt{\frac{S_u^2}{n\cdot Var(x)}}$$

Теорема Гаусса-Маркова: Если условия Гаусса-Маркова выполнены, то коэффициенты регрессии, построенной по МНК, будут **наилучшими линейными несмещенными оценками** (best linear unbiased estimators - BLUE)

Тема 2 Эконометрика 13

4. Проверка гипотез: t-тесты

Общий случай

Нулевая гипотеза H_0: $\beta=\beta_0$ (альтернативная H_1: $\beta\neq\beta_0$)

Если H_0 верна, то $\beta \sim N\left(E=\beta_0;\ S.d.=\sqrt{\dfrac{\sigma_u^2}{n\cdot Var(x)}}\right)$

Если $\beta < (\beta_0 - z\cdot S.d.)$ или $\beta > (\beta_0 + z\cdot S.d.)$, то **гипотеза H_0 отвергается** с вероятностью, соответствующей **z** (пок.-ль функции Лапласа)

P=0,95 ⇔ z=1,96
P=0,99 ⇔ z=2,58

Тема 2 Эконометрика 14

4. Проверка гипотез: t-тесты

Общий случай

На практике стандартное отклонение коэффициента регрессии ***S.d.(b)*** – неизвестно, его оценивают с помощью «стандартной ошибки» – ***S.d.(b)*** (рассч. в стат.пакетах).

Соотв. ***t*-статистика** примет вид: $t = \dfrac{b - \beta_0}{S.e.(b)}$

***t*-статистика** имеет **распределение Стьюдента** (*t*-распр.)
Зависит от числа степеней свободы ***чсс***=***n-k-1***,
где ***n*** – объем наблюдений, ***k*** – число факторов.
(При ***чсс***$\geq$20 распр.Стьюдента практически совп. с норм.распр.)

Т.о., гипотезу **H_0: $\beta=\beta_0$**:

- не следует отвергать, если $-t_{крит} \leq \dfrac{b - \beta_0}{S.e.(b)} \leq t_{крит}$
- следует отвергнуть с вероятностью ***P=1-α***, если | ***t*** |>$t_{крит}$ для определенного ***чсс*** и заданного уровня доверия ***α***

4. Проверка гипотез: t-тесты

Доверительные интервалы

b – «**точечная оценка**» ***β***.

Вероятность совпадения ***b*** с истинным значением ***β*** - невелика.

Строят **доверительный интервал**, включающий значение ***β*** с высокой, заранее определенной вероятностью:

$$b - t_{крит} \cdot S.e.(b) \leq \beta \leq b + t_{крит} \cdot S.e.(b)$$

где: ***β*** – теоретическое значение коэффициента регрессии,
b – полученная в результате расчетов оценка, ***S.e.(b)*** – его «стандартная ошибка», $t_{крит}$ – соотв. «критическое значение»
(см. таблицу t-распределения Стьюдента)

Любое гипотетическое значение ***β***, соответствующее интервалу, не будет отвергаться полученным результатом оценивания.

99%-ный интервал (для *α*=0,01) всегда шире 95%-ного (для *α*=0,05)!

4. Проверка гипотез: t-тесты

Проверка значимости

Проверка того, что коэффициент регрессии значимо отличается от 0.

Нулевая гипотеза $\mathbf{H_0}$: $\boldsymbol{\beta=0}$ (альтернативная $\mathbf{H_1}$: $\boldsymbol{\beta\neq 0}$)

Тогда расчет ***t*-статистики** можно «упростить»: $t = \dfrac{|b|}{S.e.(b)}$

Выводы о значимости коэффициента:

- Если $t > t_{крит}$ **(см. таблицу t-распределения Стьюдента)** для определенного ***чсс*** и заданного уровня доверия $\boldsymbol{\alpha}$, то гипотезу $\mathbf{H_0}$ (о незначимости коэффициента регрессии) следует отвергнуть с вероятностью $\boldsymbol{P=1-\alpha}$; коэффициент (с той же вероятностью) – значимый.
- Если $t < t_{крит}$, то гипотезу $\mathbf{H_0}$ отвергать не следует; коэффициент не является значимым.

Тема 2 Эконометрика 27

4. Проверка гипотез: t-тесты

«Лирическое отступление» в ТВиМС

Обычно проверяют НЕСКОЛЬКО уровней значимости: $\boldsymbol{\alpha}$={0,01; 0,05; 0,1} ⇔ **P**={99%; 95%, 90%}.

- ✓ Если гипотеза отвергается для всех $\boldsymbol{\alpha}$, вывод делают с большей вероятностью (**P**=99%).
- ✓ Если гипотеза НЕ отвергается для всех $\boldsymbol{\alpha}$, вывод делают с меньшей вероятностью (**P**=90%).
- ✓ Результаты двух тестов представляют в том случае, если с **P**=95% $\mathbf{H_0}$ следует отвергнуть, а с **P**=99% – нет.

- ❖ **Ошибка I рода**: отвергается истинная гипотеза
- ❖ **Ошибка II рода**: не отвергается ложная гипотеза

Гипотеза: подсудимый невиновен
Ошибка I рода: осудили невиновного
Ошибка II рода: оправдали виновного

Тема 2 Эконометрика 28

4. Проверка гипотез: t-тесты

Односторонний t-критерий

В экономических моделях часто заранее известен **знак** коэффициента угла наклона β:

- > 0 – при прямой зависимости,
- < 0 – при обратной зависимости.

В этих случаях для проверки значимости можно использовать «**односторонние t-тесты**»: нулевая гипотеза H_0: $\beta=0$; альтернативная H_1: $\beta>0$ ($\beta<0$).

уровень доверия α	двусторонний **t**-тест	0,2	0,1	0,05	0,02	0,01
	односторонний **t**-тест	0,1	0,05	0,025	0,01	0,005

Т.о. использование одностороннего критерия (где это допустимо!) позволяет получить бОльшую мощность для любого уровня доверия.
$\Rightarrow \downarrow$ вероятность принятия ложной H_0 ($\downarrow$ вероятность ошибок II рода)

5. F-тест на качество оценивания

Необх. узнать, отражает ли полученное значение к.-та детерминации R^2 истинную зависимость, или оно случайно? Таблиц для R^2 – нет! Исп. «косвенный подход» − **F-тест** (осн. на анализе дисперсий):

$$F=\frac{ESS}{k}\div\frac{RSS}{n-k-1}=\frac{R^2}{k}\div\frac{1-R^2}{n-k-1}$$

где: n *– объем выборки (количество наблюдений),*
k *– количество факторных переменных* (X_j),

$ESS=\sum(\hat{y}-\bar{y})^2$ *– "объясненная" сумма квадратов,*

$RSS=\sum(y-\hat{y})^2=\sum e^2$ *– сумма квадратов остатков*

$$\left[Var(y)=Var(\hat{y})+Var(e)\ \Rightarrow\ TSS=ESS+RSS=\sum(y-\bar{y})^2\right]$$

Теория: А.Муд, Ф.Грейбилл (Mood, Graybill); 1963

Если $\mathbf{F} > \mathbf{F}_{\text{крит}}$ (**см.** таблицы F-распределения Фишера) для определенного уровня доверия α (0,01 или 0,05) и чисел степеней свободы $\text{чсс}_1=k$ и $\text{чсс}_2=(n-k-1)$, то с вероятностью $P=(1-\alpha)$ гипотезу H_0: $\rho=0$ (о случайности связи) следует отвергнуть; зависимость в целом значима.

6. Взаимосвязь между критериями в парном регрессионном анализе

В парном регрессионном анализе (и только в нем!):

$$R^2 = \frac{Var(\hat{y})}{Var(y)} = \frac{Var(a+bx)}{Var(y)} = \frac{b^2 \cdot Var(x)}{Var(y)} = \left(\frac{Cov(x,y)}{Var(x)}\right)^2 \frac{Var(x)}{Var(y)} = \left(\frac{Cov(x,y)}{\sqrt{Var(x)\cdot Var(y)}}\right)^2 = r_{xy}^2$$

t-статистика для парного линейного к.-та корреляции (**см. тему 1, § 2**):

$$t(r_{xy}) = r_{xy}\sqrt{\frac{n-2}{1-r_{xy}^2}} \quad \Rightarrow \quad \left(t(r_{xy})\right)^2 = \frac{r_{xy}^2(n-2)}{1-r_{xy}^2}$$

F-статистика (**см. вопр.5**):

$$F = \frac{R^2}{k} \div \frac{1-R^2}{n-k-1}$$

$$\Rightarrow F = \left(t(r_{xy})\right)^2$$

Стандартная ошибка для коэффициента угла наклона (**см. вопр.3**): :

$$S.e.(b) = \sqrt{\frac{Var\ (e)}{(n-2)\cdot Var\ (x)}} = \sqrt{\frac{Var\ (e)}{(n-2)\cdot \frac{Var\ (\hat{y})}{b^2}}} = b\sqrt{\frac{1-R^2}{(n-2)\cdot R^2}} = \sqrt{\frac{1-r_{xy}^2}{(n-2)\cdot r_{xy}^2}}$$

t-статистика для коэффициента угла наклона (**см. вопр.4**):

$$t(b) = \frac{b}{S.e.(b)} = \frac{b}{b\cdot\frac{\sqrt{1-r_{xy}^2}}{r_{xy}\sqrt{n-2}}} = \frac{r_{xy}\sqrt{n-2}}{\sqrt{1-r_{xy}^2}} = t(r_{xy})$$

$$\Rightarrow F = \left(t(b)\right)^2$$

Тема 2 Эконометрика

§3. Преобразования переменных

План лекции:

- 1. Базисная процедура
- 2. Логарифмические преобразования
- 3. Требования к случайному члену
- 4. Нелинейная регрессия
- 5. Выбор функции: специальные тесты

Тема 2 Эконометрика

1. Базисная процедура
Линейность модели

Модель может быть линейной в двух смыслах:

1. относительно переменных (**y** и **x**)
2. относительно параметров регрессии (**α** и **β**)

$$y = \alpha + \beta \cdot x \quad \langle 1 \rangle$$

$$y = \alpha + \frac{\beta}{x} \quad \langle 2 \rangle$$

$$y = \alpha + \beta\sqrt{x} \quad \langle 3 \rangle$$

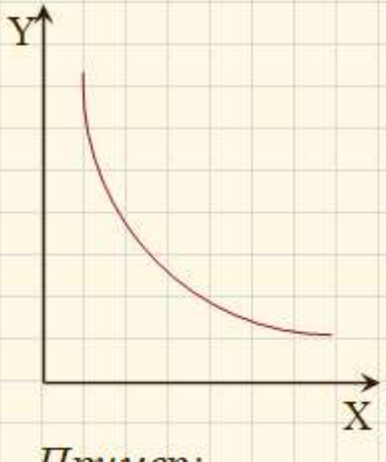
$$y = \alpha \cdot x^{\beta} \quad \langle 4 \rangle$$

$$y = \alpha \cdot \beta^{x} \quad \langle 5 \rangle$$

$\langle 1 \rangle$ - линейна в обоих смыслах,
$\langle 2 \rangle$ и $\langle 3 \rangle$ - **2**-ой тип линейности,
$\langle 4 \rangle$ и $\langle 5 \rangle$ - НЕлинейные.

Для регрессионного анализа принципиальное значение имеет **2**-й тип линейности; нелинейность **1**-го типа можно «обойти» заменой переменных.

1. Базисная процедура
Замена переменных



Гипербола: **y = a + b / x**

Замена: **z = 1 / x →**

Линейная ф-я: **y = a + b z**

Аналогично:

y = a + b √x

y = a + b $\mathbf{x^3}$

и т.п.

Пример:

Y – спрос, X – цена

2. Логарифмические преобразования

Свойства логарифмов

$$y = \alpha \cdot x^{\beta} \quad \langle 4 \rangle$$

$$y = \alpha \cdot \beta^{x} \quad \langle 5 \rangle$$

Степенные и показательные функции – нелинейны, но их можно линеаризовать!

ЛОГАРИФМ: $x = \log_a b \;\Leftrightarrow\; a^x = b$

Правила вычисления:

$$\log(X \cdot Z) = \log X + \log Z; \quad X, Z > 0$$

$$\log\left(\frac{X}{Z}\right) = \log X + \log Z; \quad X, Z > 0$$

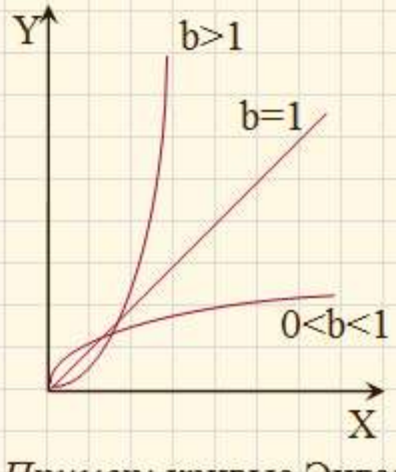
$$\log(X^{C}) = C \cdot \log X; \quad X > 0$$

Основание логарифма может быть любым (***a>0, a≠1***); в эконометрике использ. **натуральные** логарифмы (***ln***).

Ранее использовались **десятичные** логарифмы (для них есть таблицы).

2. Логарифмические преобразования

Степенная функция



Степенная ф-я: $\mathbf{y = a \cdot x^b}$

Логарифмируем:

$\mathbf{\ln y = \ln a + b \ln x} \;\rightarrow$

Замена: $\tilde{y} = \ln y$; $z = \ln x \rightarrow$

Линейная ф-я: $\tilde{y} = \tilde{a} + b\, z$

Эластичность степенной ф-и:

$$\varepsilon = \frac{\partial Y / Y}{\partial X / X} = abX^{b-1}\,\frac{X}{aX^{b}} = b$$

Пример: кривые Энгеля

Y – спрос, X – доход

(***Экономическая теория***)

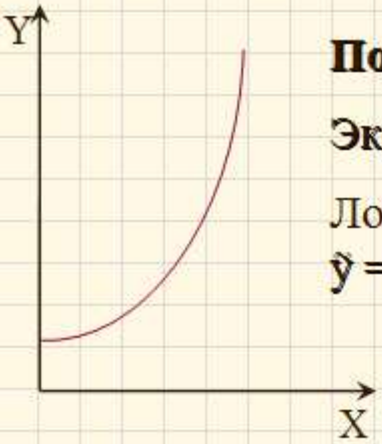
0<b<1 – товары 1-ой необходимости

b=1 – относительной необходимости

b>1 – товары роскоши

2. Логарифмические преобразования

Экспоненциальная функция



Показательная ф-я: $y = a \cdot b^X$

Экспоненциальная: $y = a \cdot e^{bx}$

Логарифмируем: $\ln y = \ln a + b\,x \rightarrow$

$\tilde{y} = \ln y \rightarrow$ Линейная ф-я: $\tilde{y} = \tilde{a} + b\,x$

Интерпретация:

b – средний темп прироста явления во времени

Пример: X – время, Y – явление, имеющее постоянный прирост во времени *(инфляция)*

Большинство функций м. б. линеаризованы!

3. Требования к случайному члену

Случайный член ***u*** должен:

- удовлетворять условиям Гаусса-Маркова,
- присутствовать в преобразованном уравнении в виде слагаемого (**+*u***).

1). Линейные (во 2-м смысле) уравнения вида ⟨1⟩, ⟨2⟩, ⟨3⟩ должны иметь **аддитивный** случайный член:

⟨2⟩: $y=\alpha+\beta/x+\boldsymbol{u} \Rightarrow y=\alpha+\beta\cdot z+\boldsymbol{u},\ z=1/x$

2). Нелинейные (в обоих смыслах) уравнения вида ⟨4⟩, ⟨5⟩ должны иметь **мультипликативный** случайный член:

⟨4⟩: $y=\alpha\cdot x^{\beta}\cdot\boldsymbol{v} \Rightarrow \tilde{y}=\tilde{a}+\beta\cdot z+\boldsymbol{u},\ z=\ln x,\ \tilde{y}=\ln y,\ \tilde{a}=\ln \alpha$

Случ.чл. ***v*** имеет логарифмически нормальное распределение.

3). Если случайный член не соотв. требованиям 1-2, модель не м.б. линеаризована $\Rightarrow$ исп. спец. методы

4. Нелинейная регрессия

Рассм. функция вида: $y=\alpha+\beta \cdot x^{\gamma}+u$

Ф-я – нелинейная и не м.б. линеаризована. $\Rightarrow$ Обычный МНК не годится. Тем не менее, прим. принцип минимизации суммы квадратов отклонений.

Процедура оценивания параметров:

1). Принимаются некоторые правдоподобные значения параметров (α_0, β_0, γ_0). По графику, напр.

2). Вычисляются «предсказанные» значения $\hat{y}$ по фактическим значениям **x** с использ. α_0, β_0, γ_0.

3). Вычисляются остатки e_i для всех наблюдений и рассчитывается сумма квадратов остатков:

$$S=\sum_{i=1}^{n} e_i^2=\sum_{i=1}^{n}\left(y_i-\hat{y}(x_i)\right)^2$$

Тема 2 Эконометрика 29

4. Нелинейная регрессия

Процедура оценивания параметров:

4). Вносятся небольшие изменения в одну или несколько оценок параметров: α_1, β_1, γ_1.

5). Вычисляются «новые» значения $\hat{y}$ и остатки e_i

6). Рассч. соотв. сумма квадратов остатков S_1.

Если $S_1 < S_0$, то новые оценки (α_1, β_1, γ_1) считаются лучше предыдущих (α_0, β_0, γ_0) и их следует использовать в качестве новой «отправной точки».

7). Шаги **4 –> 5 –> 6** повторяют до тех пор, пока **S** не перестанет уменьшаться. Делают вывод, что **S минимизирована**, и «конечные» оценки параметров являются оценками по МНК.

Итерационные процедуры оценивания нелинейных регрессий реализованы в некоторых стат.пакетах.

В пакете **MicroTSP**: **>nls**

Тема 2 Эконометрика 30

5. Выбор функции: спец. тесты
Общие соображения

Сравнение различных моделей:

1). Содержательный анализ (выбор – из знаний теории и практики анализа). Например, функции спроса – степенные, а не линейные.

См. Экономическую теорию

2). Формальный анализ (строят все возможные функции и выбирают лучшую):

- **R^2**, **F**- и **t**-статистики (в общем случае – чем больше, тем лучше)
- при примерно одинаковых характеристиках – **RSS** (чем меньше) или специальные тесты

Тема 2 Эконометрика

5. Выбор функции: спец. тесты
Тесты Бокса-Кокса

Идея метода: переменная **$(Y^{\lambda}-1)/\lambda$** при изменении **λ** переходит от линейной функции к логарифму: Box, Cox, 1964

$$\frac{Y^{\lambda}-1}{\lambda} \to \begin{cases} \lambda = 1 & \to Y \\ \lambda \to 0 & \to \ln Y \end{cases}$$

Процедура:

1. Преобр. завис.перем.**Y**: $Y_i^{*} = \dfrac{Y_i}{Y_0}, \quad Y_0 = \sqrt[n]{\prod_{i=1}^{n} Y_i}$

2. Рассч. новые перем. **Y** и **X** при **λ** от **1** до **0**: $\widetilde{Y}_i = \dfrac{\left(Y_i^{*\lambda}-1\right)}{\lambda}, \quad \widetilde{X}_i = \dfrac{\left(X_i^{\lambda}-1\right)}{\lambda}$

3. Рассч. новые регрессии для всех **λ** от **1** до **0**: $\widetilde{Y}_i = \alpha^j + \beta^j \cdot X_i + u_i; \; RSS^j$

4. Опред. из всех RSS^j минимальное значение RSS^{MIN}

5. Выбир. «крайнюю» регрессию, кот. ближе к RSS^{MIN}

Тема 2 Эконометрика

5. Выбор функции: спец. тесты

Тест Заребки

Zarembka, 1968

Упрощенный вариант теста Бокса-Кокса; позволяет сравнить линейную и логарифмическую модели и оценить значимость различий.

Рассм. 2 модели:

$$\langle 1\rangle \quad Y = a_1 + b_1 X\,; \qquad R_1^2,\, RSS_1$$

$$\langle 2\rangle \quad \ln Y = a_2 + b_2 \ln X\,; \quad R_2^2,\, RSS_2$$

$$R_1^2 \approx R_2^2$$

RSS зд. сравнивать нельзя! (**RSS_2** всегда < **RSS_1**)

⟨1⟩ **VS** ⟨2⟩?

Тема 2 Эконометрика

5. Выбор функции: спец. тесты

Тест Заребки

Процедура:

1. Преобр. завис.перем.**Y**: $Y_i^* = \dfrac{Y_i}{Y_0}$, $\quad Y_0 = \sqrt[n]{\prod_{i=1}^{n} Y_i}$

2. Оцен. новые регрессии (с **Y*** вместо **Y**):

$$\langle 3\rangle \quad Y^* = a_3 + b_3 X\,; \quad RSS_3$$

$$\langle 4\rangle \quad \ln Y^* = a_4 + b_4 \ln X\,; \ RSS_4$$

Теперь **RSS** сопоставимы!

- $RSS_3 < RSS_4 \;\Rightarrow\;$ следует использовать **⟨1⟩**
- $RSS_4 < RSS_3 \;\Rightarrow\;$ следует использовать **⟨2⟩**

3. Опред., значимо ли различие: $\chi^2 = \dfrac{n}{2} \cdot \left| \ln\left(\dfrac{RSS_3}{RSS_4}\right) \right|$

Если $\chi^2 > \chi^2_{\text{крит}}$ (**чсс**=1, α), то с вероятностью **P=1-α** разница в качестве оценивания – значима.

Тема 2 Эконометрика

§4. Множественный регрессионный анализ

План лекции:

- 1. Линейная модель с несколькими независимыми переменными
- 2. Нелинейные модели
- 3. Свойства коэффициентов
- 4. Качество оценивания
- 5. Проблема мультиколлинеарности

Тема 2 Эконометрика 35

1. Линейная модель с двумя переменными

«Классическая» линейная регрессия:

$$y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + u_i$$

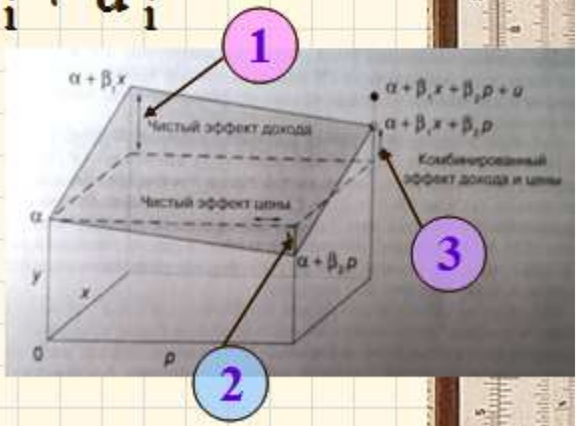
Пример: зависимость спроса (Y) от дохода (X) и цен товара (P)

Истинная зависимость:

$$Y = \alpha + \beta_1 X + \beta_2 P + u$$

Оценка – уравнение плоскости:

$$Y = a + b_1 X + b_2 P$$



Чистые эффекты: (1) – дохода (>0), (2) – цены (<0); комбинированный эффект дохода и цены (3).

Тема 2 Эконометрика 36

1. Линейная модель

Оценка по МНК

$$МНК: \quad S = \sum e_i^2 = \sum \left(y_i - (a + b_1 x_{1i} + b_2 x_{2i})\right)^2 \to \min$$

$$\begin{cases} \dfrac{\partial S}{\partial a} = -2\sum \left(y_i - a - b_1 x_{1i} - b_2 x_{2i}\right) = 0 \\ \dfrac{\partial S}{\partial b_1} = -2\sum x_{1i}\left(y_i - a - b_1 x_{1i} - b_2 x_{2i}\right) = 0 \\ \dfrac{\partial S}{\partial b_2} = -2\sum x_{2i}\left(y_i - a - b_1 x_{1i} - b_2 x_{2i}\right) = 0 \end{cases} \Rightarrow \text{правило Крамера}$$

$$\begin{cases} b_1 = \dfrac{\Delta_1}{\Delta} = \dfrac{Cov(x_1, y)\cdot Var(x_2) - Cov(x_2, y)\cdot Cov(x_1, x_2)}{Var(x_1)\cdot Var(x_2) - \left(Cov(x_1, x_2)\right)^2} \quad \langle * \rangle \\ b_2 = \dfrac{\Delta_2}{\Delta} = \dfrac{Cov(x_2, y)\cdot Var(x_1) - Cov(x_1, y)\cdot Cov(x_1, x_2)}{Var(x_1)\cdot Var(x_2) - \left(Cov(x_1, x_2)\right)^2} \\ a = \bar{y} - b_1\cdot\bar{x}_1 - b_2\cdot\bar{x}_2 \end{cases}$$

Тема 2 Эконометрика 37

1. Линейная модель

Проявление множественных связей

Множественный регрессионный анализ позволяет разграничить влияние независимых переменных, допуская при этом возможность их коррелированности.

Рассм. линейную регрессию с двумя переменными:

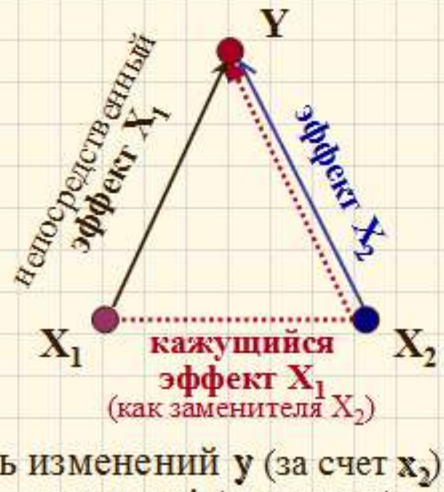
$$\mathbf{y = \alpha + \beta_1 x_1 + \beta_2 x_2 + u}$$

Пусть $\beta_1 > 0$ и $\beta_2 > 0$, а x_1 и x_2 – положительно коррелир.

Тогда при $\uparrow x_1$ «дв.эффект»:

1). $\uparrow y$, т.к. $\beta_1 > 0$

2). $\uparrow x_2 \Rightarrow \uparrow y$, т.к. $\beta_2 > 0$



Если x_2 не вкл. в модель, то часть изменений y (за счет x_2) будет приписана $x_1 \Rightarrow$ оценка β_1 – смещена! (завышена)

Тема 2 Эконометрика 38

1. Линейная модель
Интерпретация коэффициентов

Предп., что перем. $\mathbf{x_1}$ удалось разложить: $\mathbf{x_1 = z + x^*}$
где $\mathbf{z}$ – составл., способная замещать $\mathbf{x_2}$; $\mathbf{x^*}$ - ост. часть.
Тогда парная регрессия $\mathbf{y\,(x^*)}$ даст неискаженную оценку влияния $\mathbf{x_1}$ на $\mathbf{y}$. Коэффициент угла наклона для $\mathbf{y\,(x^*)}$:

$$b^* = \frac{Cov(x^*, y)}{Var(x^*)} = \frac{Cov(x_1, y) - Cov(z, y)}{Var(x_1) + Var(z) - 2Cov(x_1, z)}$$

$$z = c + d \cdot x_2 \;\Rightarrow \text{исп. свойства } Cov \text{ и } Var \Rightarrow b^* \equiv b_1 \;\langle * \rangle$$

Интерпретация: $\beta_1 = \dfrac{\Delta y}{\Delta x_1}, \quad x_2 = const$

к-т $\mathbf{b_1}$ (при переменной $\mathbf{x_1}$) выражает предельный прирост зависимой переменной $\mathbf{y}$ (т.е. на сколько своих единиц в среднем изменится $\mathbf{y}$ при увеличении $\mathbf{x_1}$ на одну свою единицу) при условии постоянства $\mathbf{x_2}$

1. Линейная модель
с несколькими переменными

$$\mathbf{y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + u}$$

где: $\mathbf{Y}$ – зависимая (результативная) переменная,
$\mathbf{X_j}$ – независимые (факторные) переменные;
$\boldsymbol{\beta_j}$ – параметры регрессии; $\mathbf{u}$ – случайный член.

МНК: $\mathbf{\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_k x_k}$

Интерпретация:
коэффициент $\mathbf{b_j}$ показ. на сколько своих единиц в среднем изменится $\mathbf{Y}$ при увеличении $\mathbf{X_j}$ на одну свою единицу и при неизменности прочих факторов.

☝ Значения $\mathbf{b_j}$ более точные, чем $\mathbf{b}$ в соотв.парных регрессиях!

1. Линейная модель

Расчет параметров

МНК: система из (k+1) уравнений.

В матричной форме: $\mathbf{b} = (X^T X)^{-1} X^T Y$

где: **X** – матрица размера (**k+1**) стл и **n** стр (1-ый стл – из 1);
Y – вектор-столбец размера **n**; **b** – вектор размера (**k+1**).

Расчет в MS Excel: f_x. Статистические. **ЛИНЕЙН**

b_k	b_{k-1}	…	b_1	b_0
S.e.(b_k)	S.e.(b_{k-1})	…	S.e.(b_1)	S.e.(b_0)
R^2	S.e.(y)			
F-stat	ч.с.с.			
ESS	RSS			

2. Нелинейные модели

Линеаризация

Как и в парных моделях – заменой и\или логарифмированием переменных (см. § 3).

Например, функция спроса: $Y = \alpha X^{\beta} P^{\gamma} v$,
где **Y** – спрос, **X** – доход, **P** – цена, **v** – случ.член.

Логарифмируем: $\ln Y = \ln \alpha + \beta \ln X + \gamma \ln P + u$

Обычный МНК: $\ln Y = \tilde{a} + b \ln X + c \ln P$

Потенцируем: $Y = e^{\tilde{a}} X^b P^c$, где **b** и **c** – оценки коэф.-тов эластичности спроса по доходу и цене.

2. Нелинейные модели

Производственная функция

Производственная функция (ПФ) Кобба-Дугласа.

Рассм. показатели реального объема выпуска (Y), капитальных затрат (K) и затрат труда (L).

$$Y = A \cdot K^{\alpha} \cdot L^{1-\alpha}$$

Пол Дуглас, Чарльз Кобб (Cobb, Douglas), 1929

Чтобы получ. оценку $\boldsymbol{\alpha}$, ур.-ие дел. на $\boldsymbol{L}$ и логарифмир.:

$$\left(\frac{Y}{L}\right) = A \cdot \left(\frac{K}{L}\right)^{\alpha} \cdot (v) \qquad \lg\left(\frac{Y}{L}\right) = \lg A + \alpha \cdot \lg\left(\frac{K}{L}\right) + (u)$$

По данным макроэкономики США за 1899-1922 гг.:

$$\lg\left(\frac{Y}{L}\right) = 0{,}02 + 0{,}25 \cdot \lg\left(\frac{K}{L}\right) \qquad R^2 = 0{,}63;\ F = 38{,}0$$

2. Нелинейные модели

Общий вид ПФ

$$Y = A \cdot K^{\alpha} \cdot L^{\beta} \cdot (v)$$

Эластичность выпуска продукции:

- по капиталу: $$\frac{\partial Y / \partial K}{Y / K} = \frac{A \cdot \left(\alpha \cdot K^{\alpha-1}\right) \cdot L^{\beta}}{A \cdot K^{\alpha-1} \cdot L^{\beta}} = \alpha$$
- по труду: $$\frac{\partial Y / \partial L}{Y / L} = \frac{A \cdot K^{\alpha} \cdot \left(\beta \cdot L^{\beta-1}\right)}{A \cdot K^{\alpha} \cdot L^{\beta-1}} = \beta$$

Эффект от масштаба:

- **(α+β)>1** – возрастающий
- **(α+β)=1** – постоянный
- **(α+β)<1** – убывающий

$$w = \beta \frac{Y}{L} \qquad \rho = \alpha \frac{Y}{K}$$

Если рынок факторов имеет конкурентный характер, то:

- ✓ ставка зарплаты (**w**) будет = предельному продукту труда, а доля труда в выпуске = $\boldsymbol{\beta}$
- ✓ аналогично, норма прибыли – $\boldsymbol{\rho}$, а доля прибыли = $\boldsymbol{\alpha}$

3. Свойства коэффициентов

Условия Гаусса-Маркова

Условия **1**, **2**, **3** – идентичны случаю парной регрессии (**см. § 3, вопр.2**); **4**-е – «обобщается»: ***pop.cov(x$_j$,u)=0*** (распр. случ.члена не завис. от распр. ∀ объясн.переменной).

Добавл. 2 «практических» требования:

1. нужно иметь достаточное количество данных: **n ≥ k+1**
2. между независимыми переменными не должно быть строгой линейной зависимости.

Несмещенность $y_i=\alpha+\beta_1 x_{1i}+\beta_2 x_{2i}+u_i$

$$\langle * \rangle \Rightarrow E(b_1)=E\left(\beta_1+\frac{Cov(x_1,u)\cdot Var(x_2)-Cov(x_2,u)\cdot Cov(x_1,x_2)}{Var(x_1)\cdot Var(x_1)+\left(Cov(x_1,x_2)\right)^2}\right)$$

если вып. 4 – е усл. Г. – М., то $E(Cov(x_1,u))=0$ *и* $E(Cov(x_2,u))=0$

$\Rightarrow E(b_1)=\beta_1$ *Аналогично* $E(b_2)=\beta_2$

Тема 2 Эконометрика

3. Свойства коэффициентов

Надежность и точность

Эффективность $y_i=\alpha+\beta_1 x_{1i}+\beta_2 x_{2i}+u_i$

$$pop.\mathrm{var}(b_1)=\frac{\sigma_u^2}{n\cdot Var(x_1)}\cdot\frac{1}{1-(r_{x_1,x_2})^2}$$

несмещ. оценка σ_u^2: $S_u^2=\frac{n}{n-3}Var(e)$

$$\Rightarrow S.e.(b_1)=\sqrt{\frac{Var(e)}{(n-3)\cdot Var(x_1)}\cdot\frac{1}{1-(r_{x_1,x_2})^2}} \qquad S.e.(b_2)=\sqrt{\frac{Var(e)}{(n-3)\cdot Var(x_2)}\cdot\frac{1}{1-(r_{x_1,x_2})^2}}$$

Если выполняются условия Гаусса-Маркова, МНК-оценки являются наилучшими линейными несмещенными оценками!

Для случая **k** факторных переменных закономерности сохраняются (формулы имеют более сложный вид).

К.-ты **b$_j$** и их станд.ошибки S.e.(**b$_j$**) – рассч. в стат.пакетах.

Тема 2 Эконометрика

4. Качество оценивания

Коэффициент детерминации

Как и в парном анализе, характ. «объясняющую способность»:

$$R^2 = \frac{Var(\hat{y})}{Var(y)} = 1 - \frac{\Sigma e_i^2}{\Sigma (y_i - \overline{y})^2}; \qquad 0 \le R^2 \le 1$$

$$где \quad e_i = y_i - \hat{y}(x_{1_i}, \ldots, x_{k_i})$$

При включении новых переменных в модель **R^2**, как правило, увелич.!

Интерпретация: **R^2** – доля (%) вариации переменной **Y**, обусловленной вариацией (всех) факторных признаков **X_j**

Скорректированный R^2 (adjusted **$\breve{R}^2$**) компенсирует автоматический сдвиг вверх при добавлении переменных:

$$\breve{R}^2 = R^2 - \frac{k}{n-k-1}(1 - R^2)$$

$\uparrow \breve{R}^2 \not\Rightarrow$ **улучш.качества**

4. Качество оценивания

F-статистика и t-статистики

Как и в парном анализе, характ. значимость модели в целом и отдельных коэффициентов соотв.

$$H_0 : \beta_1 = \beta_2 = \ldots = \beta_k = 0 \qquad H_0 : \beta_j = 0$$

$$F = \frac{ESS}{k} \div \frac{RSS}{n-k-1} = \frac{R^2}{k} \div \frac{1-R^2}{n-k-1} \qquad t = \frac{|b_j|}{S.e.(b_j)}$$

✓Если **$F > F_{крит}$** для уровня доверия **α** и **$чсс_1=k$** и **$чсс_2=(n-k-1)$**, то с вероятностью **$P=(1-\alpha)$** модель значима.

✓Если **$t > t_{крит}$** для уровня доверия **α** и **чсс=(n-k-1)**, то с вероятностью **$P=(1-\alpha)$** коэффициент значим.

Связь между F- и t-критериями: если хотя бы один коэффициент в модели значим, **F-stat** покажет значимость модели в целом. Обратное неверно!

4. Качество оценивания

Вклад группы переменных

Пусть оценена модель с ***k*** переменными. Добавл. еще несколько переменных, и теперь в модели ***m*** переменных. Для оценки значимости совместного вклада новых ***(m-k)*** переменных исп. F-статистика:

$$F = \frac{\text{улучшение качества}}{\text{использованное чсс}} \div \frac{RSS \text{ лучшей модели}}{\text{чсс лучшей модели}}$$

$$F = \frac{RSS_k - RSS_m}{m-k} \div \frac{RSS_m}{n-m-1} \qquad (k < m)$$

Если $\mathbf{F > F_{крит}}$ для уровня доверия $\boldsymbol{\alpha}$ и $\mathbf{чсс_1=(m-k)}$ и $\mathbf{чсс_2=(n-m-1)}$, то с вероятностью $\mathbf{P=(1-\alpha)}$ вклад новых переменных – значим.

Если в модель добавл. ОДНА переменная, то результаты **F**-теста будут идентичны результатам (двустороннего) **t**-теста для этой переменной.

Тема 2 Эконометрика

4. Качество оценивания

Как сравнивать.

Можно сравнивать (по одной выборке!):

- Модели разной функциональной формы с одинаковым набором переменных;
- Модели одной формы с разным количеством факторных переменных.

Чем выше $\mathbf{R^2}$, **F-stat** и **t-stat**, тем лучше.

(если примерно одинаковы, исп. и др. характеристики)

Как правило, при включении дополнительных переменных, формальные хар.-ки улучшаются.

Самое главное – значения коэффициентов регрессии не должны противоречить логике! ☺

Тема 2 Эконометрика

5. Проблема мультиколлинеарности
Причины: «кто виноват?» ☺

Мультиколлинеарность (термин ввел Р.Фриш):
• **Строгая** (perfect) – наличие линейной функциональной связи между независимыми переменными. Построение регрессии невозможно.
• **Нестрогая** (imperfect) – наличие сильной корреляции между независимыми переменными. Затрудняет работу, но не препятствует получению правильных выводов.

Возможные причины:
• Ошибочное включение в модель двух (и более) линейно зависимых переменных (выявить и оставить одну!)
• Характерна для регрессий временных рядов (если переменные имеют ярко выраженные тренды).
• Особенности конкретной выборки

В парных моделях вида У(Х) – НЕТ по опред.! ☺

Тема 2 Эконометрика 51

5. Проблема мультиколлинеарности
Последствия: «чем грозит?» ☺

Нестрогая линейная зависимость между объясняющими переменными ⇨ Оценки коэффициентов – ненадежные!

Мультиколлинеарность (**М**.)– нормальное явление: практически любая модель ее содержит (т.к. корреляционные связи между переменными есть всегда)

Однозначных критериев **М**. – НЕТ!

М. может проявиться: в незначимости коэф.-тов (т.к. стандартные ошибки ↑) при «важных» переменных, в искажении значений коэф.-тов (изредка – вплоть до смены знака) при затронутых **М**. переменных и др.

Если явных признаков нет – «бороться» с **М**. не надо!

Тема 2 Эконометрика 52

5. Проблема мультиколлинеарности
Устранение: «что делать?» ☺

Методы смягчения М.:

1. ↑степень выполнения условий надежности:

- ✓ Увеличить число наблюдений **n**:
  - Для динамических выборок: год→квартал →месяц
  - Для перекрестных – на этапе планирования исследования - см. **ОТС**, тема 7).
- ✓ Максимизировать дисперсию независимых наблюдений. Напр., «расслоение» выборки по факторному признаку - см. **ОТС**; работы Л.Киша (Kish, 1965), К.Мозера и Г.Калтона (Moser, Kalton, 1979).
- ✓ Сократить дисперсию случайного члена u: если в модели отсутств. важные переменные – включить!
- ✓ Сделать «новую» выборку, в кот. переменные слабо коррелир. между собой (на практике – малореально!)

Тема 2 Эконометрика 33

5. Проблема мультиколлинеарности
Устранение: «что делать?» ☺

2. Использование «внешней информации»:

✓ Внешние эмпирические оценки.

Напр., по перекрестной выборке получ. оценку к.-та эластичности по доходу и вкл. ее в динамическую модель. (Есть проблемы несоответствия кратко- и долгосрочных показателей. См. работы Э.Ку и Дж.Мейера (Kuh, Meyer, 1957))

✓ Теоретические ограничения на величину коэффициента или связь между ними.

Напр., ограничение на эффект от масштаба в **ПФ Кобба-Дугласа** ($\alpha+\beta=1$). Подробнее будет рассм. в теме 3, §1)

3. **Метод главных компонент** – «подбор» линейных комбинаций переменных т.о., чтобы новые переменные не коррелир. между собой.

«+» позв. полностью ликвидировать **М**.

«–» новые переменные лишены экономического смысла

Тема 2 Эконометрика 34

ТЕМА 3. Проблемы регрессионного анализа

Тема 3. Проблемы регрессионного анализа

Все, что может испортиться, – портится!

1-й закон Чизхолма

Все, что не может испортиться, – портится тоже!

Следствие 1-го закона Чизхолма

Параграфы темы:

1. Спецификация модели
2. Гетероскедастичность
3. Автокорреляция
4. Ошибки измерения

Тема 3 Эконометрика 2

1. Спецификация модели

Влияние «лишних» переменных

Оценки коэффициентов – несмещенные, но (в больш. случаев) неэффективные.

Для «лишних» переменных $b_j \to 0$.

<u>Метод пошагового регрессионного анализа</u>: удалять из модели ПО ОДНОЙ самой незначимой переменной, до тех пор, пока не останутся только значимые.

Тема 3 Эконометрика 3

1. Спецификация модели

Невключение важной переменной

Оценки коэффициентов – смещаются, S.e.(bj) и t-статистики не корректны!

Пусть истинная модель: $\mathbf{y=\beta_0+\beta_1 x_1+\beta_2 x_2+u}$

Исследователь считает, что x_2 – не важна, и не вкл. ее в модель. Оцен. модель: $\mathbf{\hat{y} = b_0 + b_1 x_1}$

Пусть $\mathbf{\beta_1>0}$, $\mathbf{\beta_2>0}$, $\mathbf{x_1}$ и $\mathbf{x_2}$ – положительно коррелир.

Тогда: $\nearrow\mathbf{x_1} \Rightarrow$ 1) $\nearrow\mathbf{y}$, т.к. $\beta_1>0$; 2) $\nearrow\mathbf{x_2} \to \nearrow\mathbf{y}$, т.к. $\beta_2>0$.

Т.е. **y** получ. «ускорение», его изменения преувелич.

$\Rightarrow$ $\mathbf{b_1}$ - <u>завышенная</u> оценка $\boldsymbol{\beta_1}$

Тема 3 Эконометрика 4

1. Спецификация модели

Невключение важной переменной

Т.о. при невключении в модель переменной x_2 переменная x_1 играет двойную роль: отражает свое прямое влияние и «замещает» недостающую x_2.

Изв.знаки $\boldsymbol{\beta_2}$ и $\mathbf{r_{X1,X2}}$ ⇨ опред. направление смещения.

В рез.-те «эффекта замещения» искажаются значения коэф.-тов детерминации $\mathbf{R^2}$ (чаще – завышаются).

Если бы $\mathbf{r_{X1,X2}}=0$, то $\mathbf{R^2_{Y(X1,X2)} = R^2_{Y(X1)} + R^2_{Y(X2)}}$
[на практике обычно: $\mathbf{R^2_{Y(X1,X2)} < (R^2_{Y(X1)} + R^2_{Y(X2)})}$]

Аналогично для случая **k** переменных.

Тема 3 Эконометрика 5

1. Спецификация модели

Замещающие переменные

Если переменная – важная и д.б. вкл. в модель, НО:
- не поддается прямой количественной оценке (уровень образования, качество жизни),
- или по ней трудно найти достоверные данные,

то в модель следует вкл. ее ***заменитель*** *(proxy)*.

Преимущества:
- ✓ качество модели ↗ (по сравн. с отсутствием),
- ✓ оценки др. коэф.-тов – несмещенные,
- ✓ косвенная информация о замещенной переменной.

Пример: Время **T** как замещающая перем. для фактора НТП.
Производственная ф-я Кобба-Дугласа: $\mathbf{Y=A\cdot K^{\alpha}\cdot L^{\beta}\cdot e^{\gamma T}}$

На практике м.б. непреднамеренное использование.

Тема 3 Эконометрика 6

1. Спецификация модели

Лаговые переменные

Лаг – задержка во времени.
Иногда необх. учитывать.
(Например, при изучении спроса на дорогостоящие товары)
Часто встреч. в моделях временных рядов.

Модель с лаговой структурой:

$$\mathbf{Y_t = \beta_0 + \beta_1 X1_{t-s} + \beta_2 X2_{t-s} + \ldots + u_t}$$

где **s** – период запаздывания.

<u>Достоинства</u>: ↘ проблему автокорреляции *(вопр.4)*.
<u>Недостатки</u>: выборка сокращается на **s** первых периодов.

Тема 3 Эконометрика 7

1. Спецификация модели

Линейные ограничения

Если параметры модели линейно-зависимы, то на них ввод. «линейное ограничение».
⇨ ↘ кол.-во объясняющих переменных ⇨
1. ↗ эффективность оценок коэффициентов
2. ↘ проблема мультиколлинеарности (вопр.2)

<u>Пример</u>: Производственная ф-я Кобба-Дугласа.
Постоянный эффект от масштаба: **α + β = 1** ⇨
Модель с 2-мя перем. → модель с одной перем.

$$Y = A \cdot K^{\alpha} \cdot L^{\beta} \quad \Rightarrow \text{делим на } L \Rightarrow \quad \frac{Y}{L} = A \cdot \left(\frac{K}{L}\right)^{\alpha}$$

Тема 3 Эконометрика 8

1. Спецификация модели

Проверка линейного ограничения

Линейное ограничение может быть введено:

- на размер определенного коэффициента (H_0: $\beta_j = \beta^*$),
- на связь между ними (как в пред.примере: H_0: $\beta = -\alpha$)

Пусть модель без ограничений содержит ***k*** переменных. Вводится ***l*** ограничений (если это ограничения на связь между переменными, то можно сократить на ***l*** переменных).

Проверка значимости ограничения с пом. F-статистики:

$$F = \frac{RSS_R - RSS_U}{l} \div \frac{RSS_U}{n-k-1} \qquad \text{где } RSS_R - RSS \text{ модели с ограничениями,} \quad RSS_U - \text{без ограничений;} \quad RSS_U < RSS_R$$

Если $\mathbf{F>F_{крит}}$ ($\boldsymbol{\alpha}$; **$\mathbf{чсс_1}$=*l*, $\mathbf{чсс_2}$=*n-k-1***), то с **P**=(1-**α**) качество модели без ограничений значимо лучше, чем без ограничений ⇨ ограничение вводить не следует. Если $\mathbf{F<F_{крит}}$, то ограничение не отбрасывается.

1. Спецификация модели

Проверка линейного ограничения

<u>Пример</u>. Производственная функция Кобба-Дугласа. Общий вид: $Y=A\cdot K^{\alpha}\cdot L^{\beta}$. Ограничение: $\alpha+\beta=1$.

По данным экономики США, 1899-1922 гг.:

- без огр.: $\ln Y = -0{,}18 + 0{,}23\cdot\ln K + 0{,}81\cdot\ln L$
  $R^2=0{,}96$ F=35,8 RSS=0,0710
- с огр.: $\ln(Y/L) = 0{,}02 + 0{,}25\cdot\ln(K/L)$
  $R^2=0{,}63$ F=38,0 RSS=0,0716

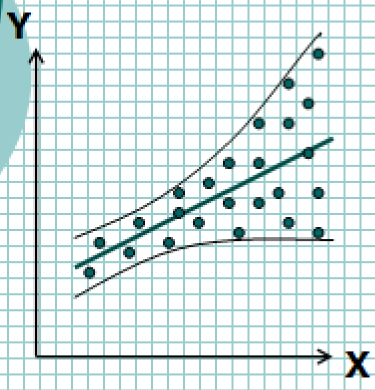
$$F = \frac{RSS_R - RSS_U}{l} \div \frac{RSS_U}{n-k-1} = \frac{0{,}0716-0{,}0710}{1} \div \frac{0{,}0710}{24-2-1} = 0{,}18$$

$F_{крит}$ (α=0,05; $чсс_1$=*l*=1, $чсс_2$=*n-k-1*=21) =4,32.

$F<F_{крит}$ ⇨ ограничение не отбрасывается.

2. Гетероскедастичность

Причины и последствия



Разброс вокруг линии регрессии – неодинаков.

Встречается чаще в регрессиях временных рядов из-за разницы масштаба единиц измерения

1) оценки коэфф.-тов не эффективны;
2) ошибки S.e.(bj) ↘ ⇨ t-статистики ↗.

Тема 3 Эконометрика 11

2. Гетероскедастичность

Обнаружение и устранение

1. Тест ранговой корреляции Спирмена.

$$\rho_{x,e} = 1 - \frac{6\cdot\sum(p_x - p_{|e|})^2}{n\cdot(n^2-1)}; \quad t(\rho) = \rho\cdot\sqrt{n-1} \quad ? \quad t_{норм}.$$

Если t<1,96 (2,58), то с вер. P=95% (99%) в модели присутствует гетероскедастичность.

2. Тест Голдфелда-Квандта (распр. в посл.время).
Наблюдения упоряд. по величине X и дел. на 3 части: 1-ая и 3-я по n′=[n/3]+1, средняя – отбрасывается. Для обеих строятся регрессии и рассчитыв. RSS.
$F=RSS_3/RSS_1$. Если $F > F_{крит}(\alpha;\ чсс_1=чсс_2=n')$, то с вер. $P=(1-\alpha)$ модель гетероскедастична.

Устранение: масштабирование модели (дел. на перем., пропорциональную разбросу)

Тема 3 Эконометрика 12

2. Гетероскедастичность
Пример наличия в перекрестной выборке

Рассм. зависимость между «гос.расходами на образование» (Y) и ВВП (X): $Y=\alpha+\beta X+u$
В выборке – и большие и маленькие страны.

<u>Итоги</u>: на образование тратится (6±3)% от ВВП. Очевидно, что отклонение на 1 п.п. от среднего (6%) при большом ВВП больше, чем при малом.

⇨ Гетероскедастичность! ⇨ Масштабирование.

1) Величина отклонения пропорц. размеру ВВП:
$(Y/X)=\alpha(1/X)+\beta+(u/X)$ [линейная → гипербола]

2) ИЛИ - численности населения (N):
$(Y/N)=\alpha(1/N)+\beta(X/N)+(u/N)$ [1перем → 2перем]

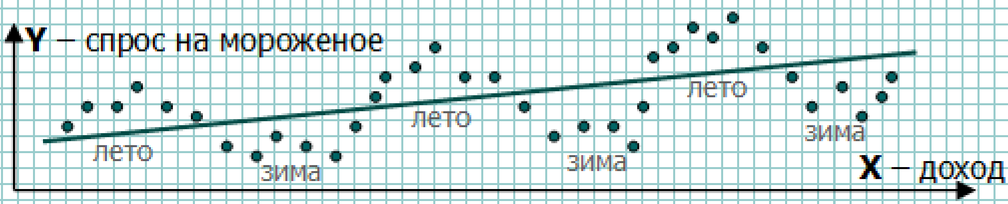
Тема 3 Эконометрика 13

3. Автокорреляция
Причины и последствия

Зависимость между остатками в различных наблюдениях. Последствия (похожи на гет.-ть):

1) оценки коэфф.-тов не эффективны;
2) S.e.(bj) (в больш.) ↘ ⇨ t-статистики ↗.

Встречается ТОЛЬКО в регрессиях временных рядов из-за сезонности, цикличности и т.п.
Чаще – положительная автокорреляция. **Пример:**

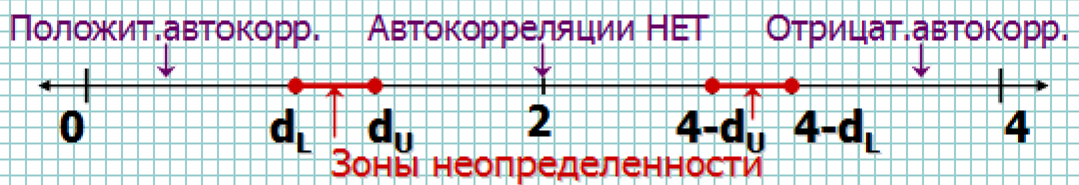


Тема 3 Эконометрика 14

3. Автокорреляция
Обнаружение

Обнаружение автокорреляции 1-го порядка (зависимость между остатками соседних наблюдений) – статистика Дарбина-Уотсона (не рассч. в MS Excel!)

$$DW = \frac{\sum_{t=2}^{T}(e_t - e_{t-1})^2}{\sum_{t=1}^{T} e_t^2}; \quad \left\{ \begin{array}{l} DW \approx 2\cdot(1-\hat{\rho}_{(e_t;e_{t-1})}) \\ где\ \hat{\rho}\ -\ к.т\ корр.\ остатков \end{array} \right\}$$



Крит. знач. $\mathbf{d_L}$ и $\mathbf{d_U}$ – по табл. Дарбина-Уотсона для α={0,01;0,05}, числа наблюд. ***n*** и кол.-ва перем. ***k***

Тема 3 Эконометрика 15

3. Автокорреляция
Устранение

Для устранения автокорреляции 1-го порядка исп. авторегрессионная схема 1-го порядка.

Суть: $\mathbf{Y_t - \rho\, Y_{t-1}}$, где $\boldsymbol{\rho}$ – к.-т корреляции остатков

На практике исп. метод Кокрана-Оркатта - итеративная процедура расчета $\boldsymbol{\rho}$ и одновременной корректировки коэфф.-тов (реал. в **TSP**).

Статистика DW и метод К.-О. не работают в случаях:
- автокорреляции более высоких порядков (др.тесты);
- в моделях с лаговой зависимой пер.: $\mathbf{Y_t(X_t, Z_t, \ldots Y_{t-1})}$ (зд. нужна ***h***-статистика Дарбина!);
- при несоответствии формы зависимости.

Тема 3 Эконометрика 16

4. Ошибки измерения

Причины, последствия и устранение

Причины возникновения:

○ Ошибки «по вине опрашиваемого» - умышленные или нет;
○ Объясн. переменная не м.б. точно измерена;
○ Исп. «несовершенные» замещающие перем.

В результате объясняющая переменная **X** и случайный член **u** одномоментно коррелированы

Последствия: оценки коэффициентов МНК – смещенные (заниженные).

Устранение: метод инструментальных переменных (МИП) → состоятельные оц.

Тема 3 Эконометрика 17

4. Ошибки измерения

Пример: Ф.-я потребления Фридмена

Доход: $Y_t=Y_t^P+Y_t^T$. Потребление: $C_t=C_t^P+C_t^T$.
(P – постоянная, T – переменная части)
Фридмен (1957 г.) выдв. гипотезу, что $C_t^P = \beta Y_t^P$
Постоянный доход – субъективен, не измеряется непосредственно. Фактические значения Y_t и C_t рассм. как Y_t^P и C_t^P с ошибками измерения.
Оцен. «обычную» ф.-ю потребления: $C = a + b Y$
⇨получ. заниженная оценка мультипликатора β (причем, ↘ тем больше, чем > вариация Y_t^T), что и было доказано Фридменом на факт.данных.
МИП: данные за 2 последовательных периода. Тогда Y_2 – инстр.переменная для Y_1 и наоборот.
$b_{МИП} > b_{Мнк}$ (что доказал ***Ливиатан***, 1963 г.)

Тема 3 Эконометрика 18