Лекция 5

Преобразования переменных в парной регрессии

Что будет в этой лекции

- Ключевые точки (начало координат)
- Кривая или прямая
- Форма криволинейной зависимости
- Вспомогательные экономические показатели (скорость и темп роста, эластичность)
- Уточнение формы (экстремумы, пределы)
- Сравнение функциональных форм

За пределами этой лекции темы: разбиение выборки на части, структурные сдвиги, лаги, сложные формы нелинейных зависимостей (не сводящиеся к линейным)

Этапы построения модели регрессии

Этапы построения модели

1. Выбор теоретических предпосылок

2. Формализация предпосылок

3. Математический вывод зависимости

4. Анализ построенной модели

Почему нужны нелинейные эконометрические модели

Пример: Анализ роста

Теоретический феномен - экономический рост

Анализ предпосылок: прирост пропорционален накопленному потенциалу

Формализация предпосылок:

$$dY = \beta Y \;\Rightarrow \frac{dY}{Y} = \beta \;\Rightarrow\; \ln Y = \alpha + \beta t$$

Интерпретация и анализ: коэффициент регрессии «бета» - годовой темп роста, возможно сопоставление с реальными данными

Альтернативные функциональные формы

Правила выбора формы зависимости

1) Исходить из экономической теории
2) Оценивать формальное качество
3) Дополнительно проверять по нескольким содержательным критериям

Линейные формы: интерпретация регрессии

Линейные формы

$$Y_i = \alpha + \beta X_i + u$$

Интерпретация коэффициента регрессии - предельный эффект независимого фактора

$$\beta = \frac{\Delta Y}{\Delta X}$$

Для полученных оценок уравнения регрессии

$$Y = a + bX$$

$$b = \frac{\Delta Y}{\Delta X} \quad \Delta X = 1 \quad b = \Delta Y \quad a = \overline{Y} - b\overline{X}$$

Коэффициент регрессии показывает прирост результирующей переменной при изменении независимого фактора на единицу

Интерпретация регрессии от времени

Линейные формы

$$Y_i = \alpha + \beta t_i + u$$

Интерпретация коэффициента регрессии от времени - ежегодный прирост зависимой переменной

$$\beta = \frac{\Delta Y}{\Delta t}$$

Вычисление эластичности для линейной регрессии

Линейные формы

$$Y_i = \alpha + \beta t_i + u$$

Вычисление эластичности

$$e = \frac{\Delta Y / Y}{\Delta X / X} = \frac{\Delta Y}{\Delta X} \cdot \frac{X}{Y} = \beta \cdot \frac{X}{Y}$$

Пример:
```
LS HOUS C DPI
GENR E=C(2)*DPI/HOUS
PLOT E
=@MEAN(E)
```

Двойная логарифмическая форма

(Двойная) логарифмическая форма

$$Y_i = \alpha' X_i^{\beta} u'$$

Преобразование переменных, приводящее к линейной регрессии

$$\ln Y_i = \alpha + \beta \ln X_i + u$$

Пример: *GENR LGHOUS=LOG(HOUS)*
GENR LGDPI=LOG(DPI)
LS LGHOUS C LGDPI

Коэффициенты лог. регрессии - эластичности

(Двойная) логарифмическая форма

$$\ln Y_i = \alpha + \beta \ln X_i + u$$

Интерпретация коэффициента регрессии - эластичность результирующей переменной по независимой переменной

$$\frac{dY}{Y} = \beta \frac{dX}{X} \Rightarrow \beta = \frac{dY / Y}{dX / X}$$

Коэффициент при независимой переменной показывает, на сколько процентов возрастает Y при возрастании X на один процент

Двойную логарифмическую следует использовать там, где есть основание предполагать постоянство эластичности

Предельный эффект в логарифмической регрессии

(Двойная) логарифмическая форма

$$\ln Y_i = \alpha + \beta \ln X_i + u$$

Вычисление коэффициента наклона (скорости роста)

$$\frac{dY}{dX} = \beta \frac{Y}{X}$$

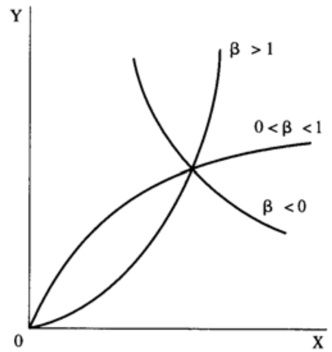
Наклон постоянно меняется с изменением номера наблюдения

Дважды логарифмические кривые

(Двойная) логарифмическая форма

$$\ln Y_i = \alpha + \beta \ln X_i + u$$

Формы зависимости, выражаемые логарифмической формой зависимости



Полулогарифмические формы

Полулогарифмические формы

1) Линейно-логарифмическая форма (логарифм при независимой переменной)

2) Логарифмически-линейная форма (логарифм при зависимой переменной)

Линейно-логарифмическая форма

Линейно-логарифмическая форма

$$Y_i = \alpha + \beta \ln X_i + u$$

Интерпретация коэффициента регрессии

$$dY = \beta \frac{dX}{X} \Rightarrow \beta \cdot = \frac{dY}{dX / X} \Rightarrow \beta / 100 = \frac{dY}{(dX / X) \cdot 100}$$

Коэффициент при независимой переменной показывает на сколько единиц возрастает Y при возрастании X на один процент

При интерпретации коэффициент следует делить на 100

Если X увеличится на 1%, то прирост Y составит $\beta / 100$ единиц (в которых измеряется Y).

Область применения линейно-логарифмической формы

Линейно-логарифмическая форма

$$Y_i = \alpha + \beta \ln X_i + u$$

Расчет эластичности

$$dY = \beta \frac{dX}{X} \Rightarrow e = \frac{dY}{dX} \cdot \frac{X}{Y} = \beta \frac{1}{X} \cdot \frac{X}{Y} = \beta \frac{1}{Y}$$

Эластичность убывает с ростом Y

Это указывает на класс зависимостей, где стоит применять эту функциональную форму

Логарифм при X снижает влияние роста X *(степень влияния X снижается с ростом X)*. Моделирование эффектов насыщения на уровне скорости роста: «возрастание с убывающей скоростью»

Пример: кривые Энгеля

Логарифмически-линейная форма

Логарифмически-линейная форма

$$\ln Y_i = \alpha + \beta X_i + u$$

Интерпретация коэффициента регрессии

$$\frac{dY}{Y} = \beta dX \Rightarrow \frac{\beta}{dX} \cdot = \frac{dY}{Y} \Rightarrow (dX = 1) \Rightarrow \beta \cdot 100\% = \frac{dY}{Y} 100\%$$

Коэффициент при независимой переменной показывает на сколько процентов возрастает Y при возрастании X на одну единицу

При интерпретации коэффициент следует умножать на 100

Область применения логарифмически-линейной формы

Логарифмически-линейная форма

$$\ln Y_i = \alpha + \beta X_i + u$$

Расчет эластичности

$$\frac{dY}{Y} = \beta dX \Rightarrow \frac{dY}{dX} = \beta Y \Rightarrow e = \frac{dY}{dX}\frac{X}{Y} = \beta\frac{YX}{Y} = \beta X$$

Эластичность растет с ростом Y

Это указывает на класс зависимостей, где стоит применять эту функциональную форму

Моделирование эффектов насыщения на уровне скорости роста: «возрастание с возрастающей скоростью»

Примеры: кривые Энгеля для товаров роскоши и товаров, спрос на которые проявляется при большом доходе, моделирование оплаты труда (процентная надбавка за стаж или опыт).

Лог.-линейная форма как модель темпа роста

Логарифмически-линейная форма по времени

$$\ln Y_i = \alpha + \beta t_i + u_i$$

Вид уравнения

$$Y_i = e^{\alpha} e^{\beta t_i} e^{u_i}$$

Интерпретация

$$\frac{dY}{Y} = \beta t$$

Коэффициент при переменной времени выражает темп прироста (часто говорят «темп роста»). Он показывает на сколько процентов (если умножить его на 100) возрастает Y ежегодно

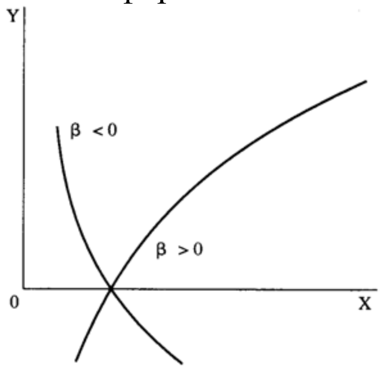
Эту функциональную форму удобно использовать для моделирования процессов экономического роста

Логарифмически-линейные кривые

Логарифмически-линейная форма

$$\ln Y_i = \alpha + \beta t_i + u_i$$

Формы зависимости, выражаемые логарифмически-линейной формой зависимости



Обратные зависимости

Обратные зависимости

$$Y_i = \alpha + \beta_1 \frac{1}{X_i} + u_i$$

Вычисление эластичности

$$e = \frac{\Delta Y / Y}{\Delta X / X} = \beta \cdot \left(-\frac{1}{XY}\right)$$

С ростом X зависимая переменная приближается к некоторому числу (моделирование эффекта насыщения)

Пример: Моделирование потребления товаров первой необходимости (быстрое достижение насыщения)

Сводка формул для нелинейных моделей регрессии

Сводка результатов для альтернативных функциональных форм в парной регрессии

Функциональная форма	Уравнение (без учета других факторов)	Наклон $\frac{\Delta Y}{\Delta X}$	Эластичность $\frac{\Delta Y}{\Delta X}\cdot\frac{X}{Y}$
Линейная	$Y_i = \alpha + \beta X_i + u$	β	$\beta\left(\frac{X_i}{Y_i}\right)$
Двойная логарифмическая	$\ln Y_i = \alpha + \beta \ln X_i + u$	$\beta\left(\frac{Y_i}{X_i}\right)$	β
Линейно-логарифмическая (lnX)	$Y_i = \alpha + \beta \ln X_i + u$	$\beta\left(\frac{1}{X_i}\right)$	$\beta\left(\frac{1}{Y_i}\right)$
Логарифмически-линейная (lnY)	$\ln Y_i = \alpha + \beta X_i + u$	βY_i	βX_i
Обратная	$Y_i = \alpha + \beta \frac{1}{X_i} + u$	$-\beta\left(\frac{1}{X_i^2}\right)$	$-\beta\left(\frac{1}{X_i Y_i}\right)$

1

Сравнение различных моделей

1) Содержательный анализ

2) Формальный анализ

* Метод Зарембки

* Преобразование Бокса-Кокса

1

Идея метода Зарембки

Метод Зарембки

Применим для выбора из двух форм моделей (несравнимых непосредственно), в одной из которых зависимая переменная входит с логарифмом, а в другой - нет

Метод позволяет сравнить линейную и логарифмическую регрессии и оценить значимость наблюдаемых различий

1

Алгоритм метода Зарембки

Метод Зарембки

1. Вычисляем среднее геометрическое значений зависимой переменной и все ее значения делятся на это среднее

$$Y_i^* = Y_i / \sqrt[n]{Y_1 Y_2 ... Y_n} = Y_i / e^{\frac{1}{n}(\ln Y_1 + \ln Y_2 + ... + \ln Y_n)}$$

2. Рассчитываются линейная и логарифмическая регрессии и сравниваются значения их суммы квадратов остатков (*SSR*)

$$Y_i^* = \alpha_1 + \beta_1 X_i + u_i \ ; SSR_1 \qquad \ln Y_i^* = \alpha_2 + \beta_2 \ln X_i + u_i \ ; SSR_2$$

3. Вычисляем χ^2-статистику для оценки значимости различий

$$\chi^2 = \left(\frac{n}{2}\right)\cdot\left|\ \ln\frac{SSR1}{SSR2}\right|$$

4. Сравниваем с критическим значением χ^2-распределения с одной степенью свободы, различия значимы, если $\chi^2 > \chi^2_{крит.}$

1

Метод Бокса-Кокса

Метод Бокса-Кокса (Box-Cox)

Идея метода: переменная

$$\frac{Y^{\lambda}-1}{\lambda}$$

при λ=1 превращается в линейную функцию

$$\frac{Y^{1}-1}{1}$$

а при λ→0 переходит в логарифм

$$\frac{Y^{\lambda}-1}{\lambda}\to \ln Y$$

Плавно изменяя λ, можно постепенно перейти от линейной регрессии к логарифмической, все время сравнивая качество

Алгоритм метода Бокса-Кокса

Метод Бокса-Кокса (Box-Cox)

1. Преобразуем зависимую переменную по методу Зарембки

$$Y_i^{*}=Y_i/\sqrt[n]{Y_1Y_2\ldots Y_n}=Y_i/e^{\frac{1}{n}(\ln Y_1+\ln Y_2+\ldots+\ln Y_n)}$$

2. Рассчитываем новые переменные (преобразование Бокса-Кокса) при значениях λ от 1 до 0.

$Y_{i(B-C)}=(Y_i^{*\lambda}-1)/\lambda$ и $X_{i(B-C)}=(X_i^{\lambda}-1)/\lambda$

3. Рассчитываем регрессии для новых переменных при значениях λ от 1 до 0. $Y_{i(B-C)}=\alpha+\beta X_{i(B-C)}+u_i$

4. Выбираем минимальное значение суммы квадратов остатков (SSR), выбираем одну из крайних регрессий, к которой ближе точка минимума

Конец лекции