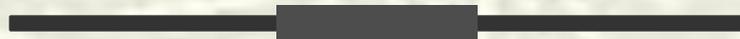


# Преобразование переменных в парной регрессии

Лекция



# Цели лекции

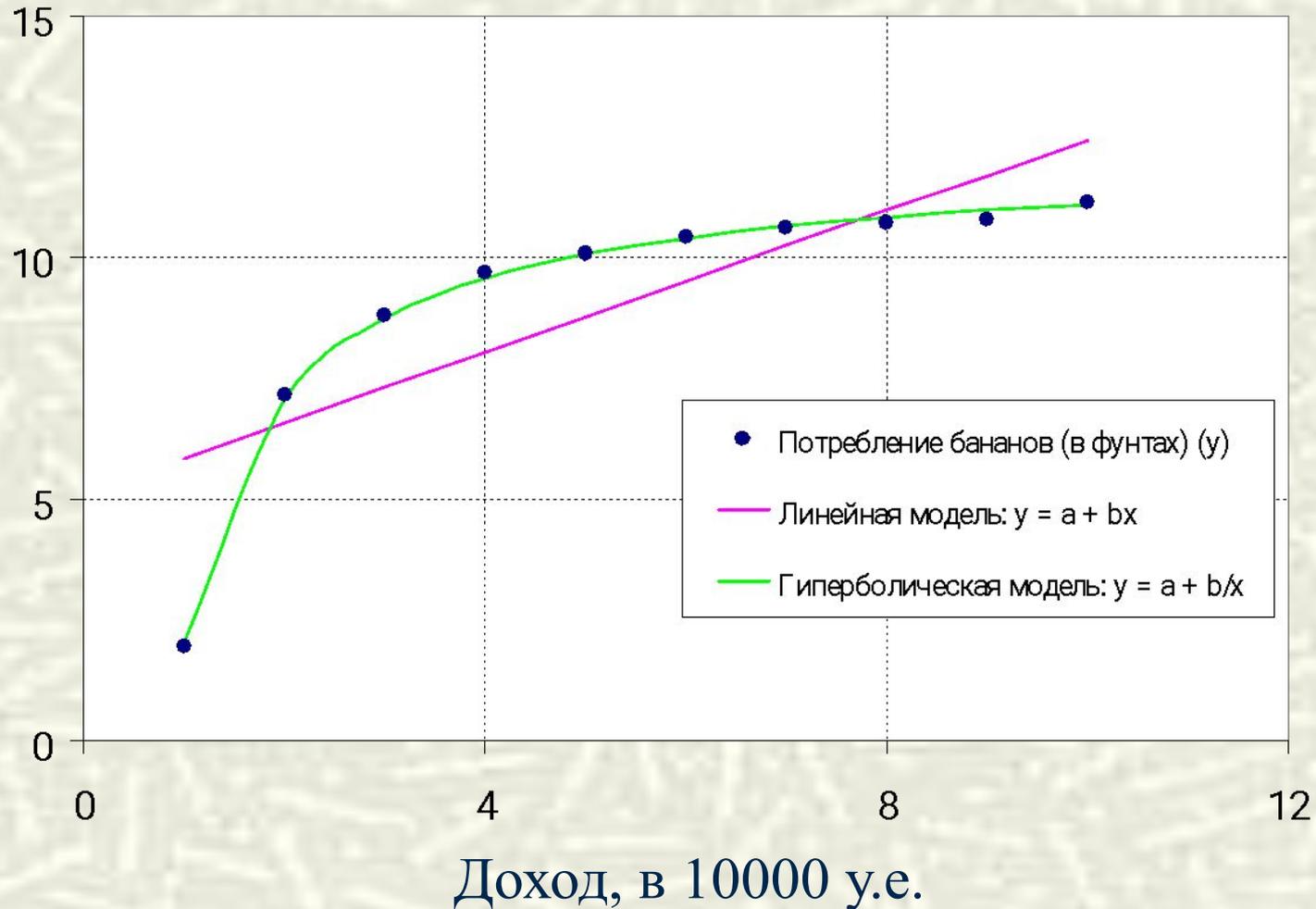
---

- Понять смысл нелинейной регрессии
- Научиться выполнять преобразования переменных
- Экономическая интерпретация регрессионной модели

# Пример нелинейной зависимости

## зависимости

Бананы, в фунтах



# Направления анализа и развития парной линейной регрессии

---

- 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

# Этапы построения модели

---

1. Выбор теоретических предпосылок
2. Формализация предпосылок
3. Построение математической модели
4. Анализ построенной модели

# Производственная функция Кобба-Дугласа

Многие экономические процессы не являются линейными по сути. Их моделирование линейными уравнениями не даст положительного результата.

Пример. Производственная функция Кобба – Дугласа

$$Y = AK^{\alpha}L^{\beta}$$

$Y$  – объем выпуска;  $K$ ,  $L$  – затраты капитала и труда;  $\alpha$ ,  $\beta$  – параметры модели.

# Анализ экономического роста

Анализ теоретических 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  $\beta$  – годовой темп роста, возможно сопоставление с реальными данными

# Классы нелинейных регрессий

---

Различают *два класса* нелинейных регрессий:

1. Регрессии, нелинейные относительно переменных, но линейные по оцениваемым параметрам.
2. Регрессии, нелинейные по оцениваемым параметрам.

Регрессии, нелинейные относительно объясняющих переменных, всегда сводятся к линейным моделям.

# Альтернативные функциональные формы: правила выбора

## Правила выбора формы зависимости:

- 1. Исходить из экономической теории.*
- 2. Оценивать формальное качество модели.*
- 3. Дополнительно проверять по нескольким содержательным критериям.*
- 4. Ответить на вопросы, возникающие при анализе модели:*
  - каковы признаки качественной модели;
  - какие ошибки спецификации встречаются и каковы их последствия;
  - как обнаружить ошибку спецификации;
  - каким образом можно исправить ошибку спецификации и перейти к более качественной модели.

# Линейная форма

---

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

Интерпретация коэффициента регрессии

$\beta$  – предельный эффект независимого фактора

$$\beta = Y'_X = \frac{dY}{dX} = \frac{\Delta Y}{\Delta X}$$

# Линейная форма

Для полученных оценок  $a$ ,  $b$  уравнения регрессии:

$$Y = a + bX$$

$$b = \frac{\Delta Y}{\Delta X} \quad \Delta X = 1 \quad \Rightarrow \quad b = \Delta Y$$

$$a = \bar{Y} - b\bar{X} \quad \Rightarrow \quad a = \bar{Y} (\bar{X} = 0)$$

# Линейная форма

---

Коэффициент регрессии  $b$  показывает прирост зависимой переменной при изменении объясняющей переменной на единицу.

Коэффициент регрессии  $b$  – угловой коэффициент линии регрессии

Коэффициент регрессии  $a$  – среднее значение зависимой переменной при нулевом значении объясняющей переменной

# Линейная форма от времени

$$Y_i = a + bt_i$$

Интерпретация коэффициента регрессии от времени – ежегодный (ежемесячный и т.д.) прирост зависимой переменной

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

# Моделирование эластичности

Независимо от вида математической связи между  $Y$  и  $X$  *эластичность* равна:

$$L = \frac{dY / Y}{dX / X} = \frac{dY / dX}{Y / X} \approx \frac{\Delta Y / Y}{\Delta X / X} = \frac{\Delta Y}{\Delta X} \cdot \frac{X}{Y}$$

Эластичность  $y$  по  $x$  рассчитывается как относительное изменение  $y$  на единицу относительного изменения  $x$ .

# Пример расчета эластичности

Рассмотрим кривую Энгеля:  $Y = \alpha X^\beta$

где  $Y$  – спрос на товар,  $X$  – доход. Имеем:

$$\text{Эластичность} = \frac{dY / dX}{Y / X} = \frac{\alpha \beta X^{\beta-1}}{\alpha X^{\beta-1}} = \beta,$$

Например для модели  $Y = 0,01X^{0,3}$  эластичность спроса по доходу равна 0,3. Иными словами, изменение дохода ( $X$ ) на 1% вызывает изменение спроса ( $Y$ ) на 0,3%

# Эластичность – переменная величина

---

Эластичность не всегда бывает постоянной для различных значений  $X$  и  $Y$

Например, для линейной модели

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

# Средний коэффициент эластичности

Средний коэффициент эластичности  $\bar{L}$  показывает, на сколько процентов в среднем по совокупности изменится результат  $Y$  от своей средней величины при изменении фактора  $X$  на 1% от своего среднего значения

$$\bar{L} = f'(\bar{X}) \frac{\bar{X}}{\bar{Y}}$$

# Логарифмическая форма

---

$$Y_i = \alpha X_i^\beta \varepsilon_i'$$

Прологарифмировав обе части уравнения,  
получим

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

# Логарифмическая форма

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

Интерпретация коэффициента регрессии  $\beta$  – эластичность зависимой переменной по объясняющей переменной

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

Коэффициент при объясняющей переменной показывает, на сколько процентов возрастает  $Y$  при возрастании  $X$  на 1%.

Логарифмическую форму следует использовать там, где есть основание предполагать постоянство эластичности

# Логарифмическая форма

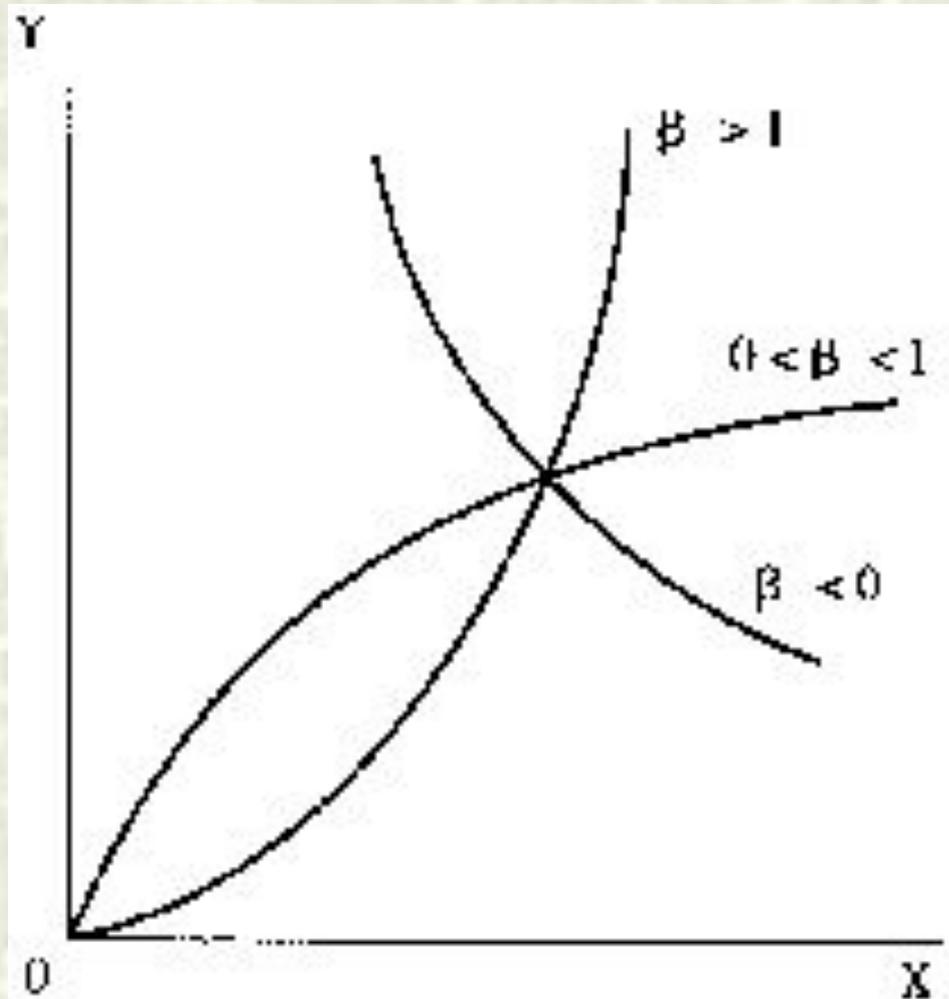
$$\ln Y_i = \alpha + \beta \ln X_i + \varepsilon_i$$

Вычисление наклона (скорости роста)

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

Наклон постоянно меняется с изменением номера наблюдения

# Графики логарифмической формы зависимости



# Полулогарифмические формы

---

1. Линейно-логарифмическая форма  
(логарифм при объясняющей переменной)
2. Логарифмически-линейная форма  
(логарифм при зависимой переменной)

# Линейно-логарифмическая форма

$$Y = \beta_0 + \beta_1 \ln X_1 + \beta_2 X_2 + \varepsilon$$

Интерпретация коэффициента регрессии  $\beta$ :

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

Коэффициент при объясняющей переменной показывает на сколько единиц возрастает  $Y$  при возрастании  $X$  на 1%

При интерпретации коэффициент следует делить на 100

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

# Линейно-логарифмическая форма

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

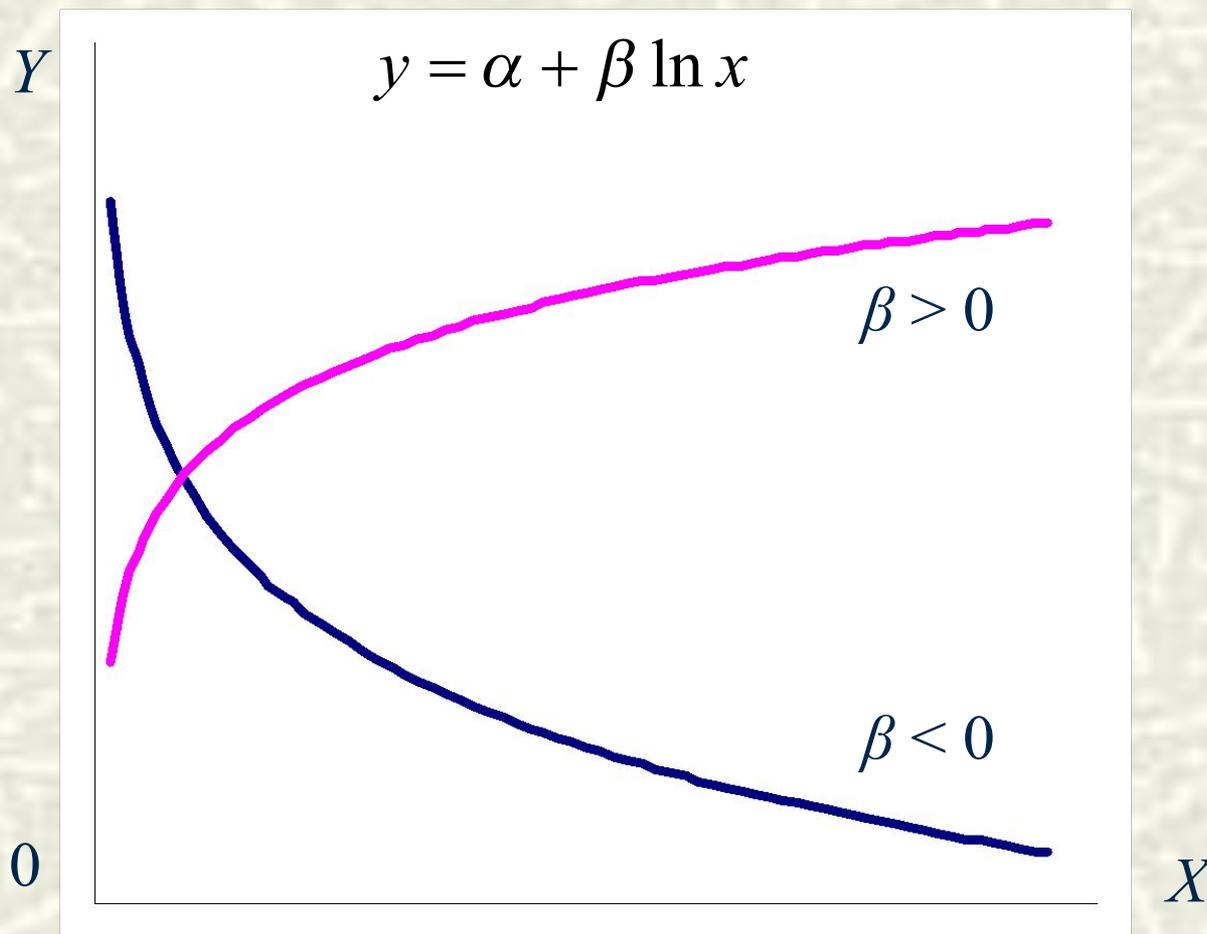
Эластичность убывает с ростом  $Y$ :

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

Это указывает на класс зависимостей, где следует применять линейно-логарифмическую форму регрессии

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»

# Графики линейно-логарифмической формы зависимости



# Логарифмически-линейная форма

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

Интерпретация коэффициента регрессии  $\beta$ :

$$\frac{dY}{Y} = \beta dX \quad \Rightarrow \quad \beta = \frac{dY}{Y dX} \quad \xrightarrow{dX=1} \quad \beta \cdot 100\% = \frac{dY}{Y} \cdot 100\%$$

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

При интерпретации коэффициент следует умножать на 100

# Логарифмически-линейная форма

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

Эластичность растет с ростом  $Y$ :

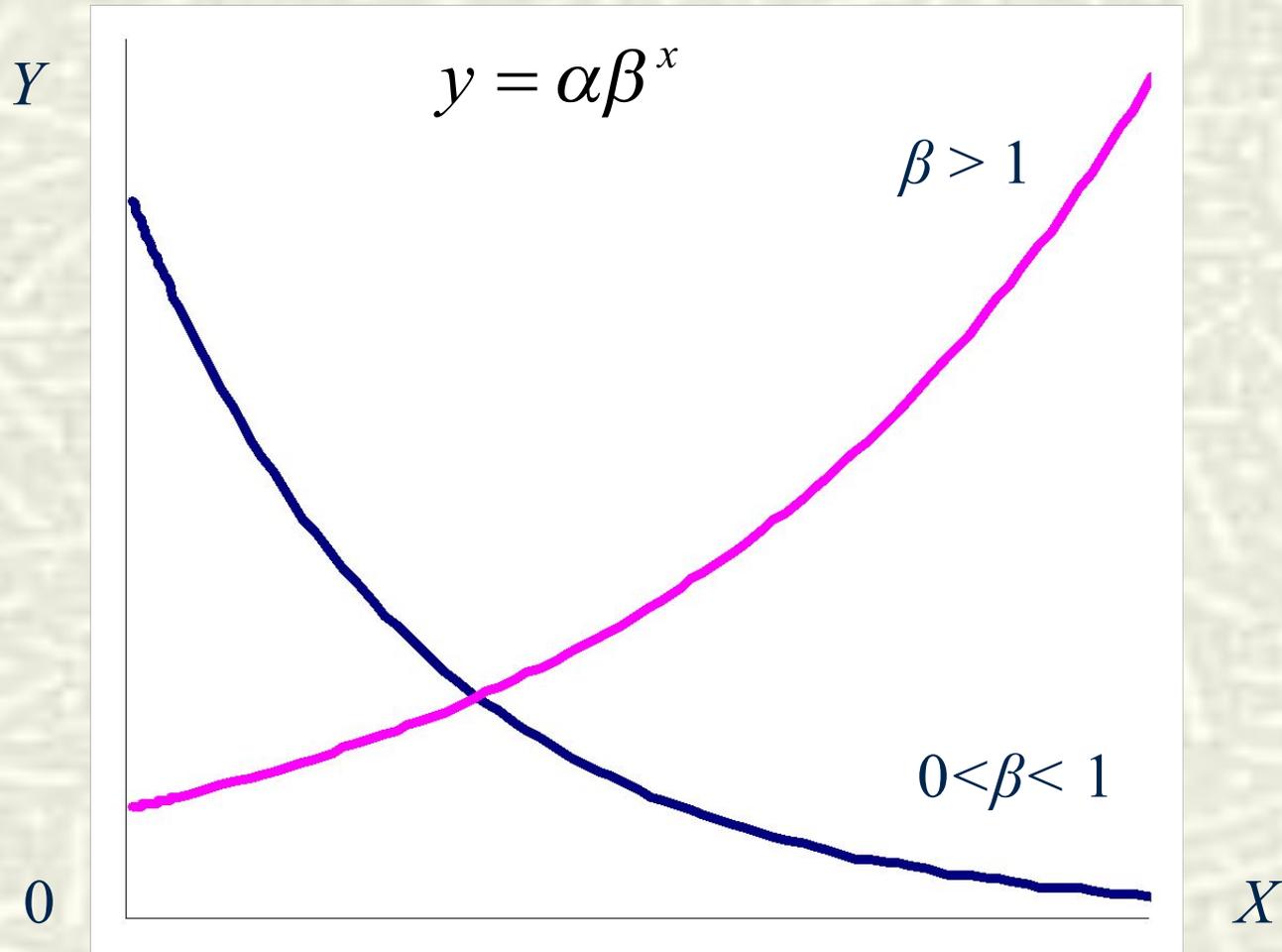
$$\beta = \frac{dY}{YdX} \Rightarrow L = \frac{dY}{dX} \cdot \frac{X}{Y} = \beta \frac{YX}{Y} = \beta X$$

Это указывает на класс зависимостей, где следует применять линейно-логарифмическую форму регрессии

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

Примеры: кривые Энгеля для товаров роскоши, моделирование оплаты труда (процентная надбавка за стаж и опыт)

# Графики логарифмически-линейной формы зависимости



# Логарифмически-линейная форма от времени

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

Вид уравнения:  $Y_i = e^\alpha e^{\beta t_i} e^{\varepsilon_i}$

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

Коэффициент при переменной времени выражает темп прироста. Он показывает на сколько процентов (если умножить его на 100) возрастает  $Y$  ежегодно

Эту функциональную форму удобно использовать для моделирования процессов экономического роста

# Обратные зависимости

$$Y_i = \alpha + \beta \frac{1}{X_i} + \varepsilon_i$$

Вычисление эластичности

$$L = \frac{dY / Y}{dX / X} = \beta \cdot \left( -\frac{1}{XY} \right)$$

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

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

# Сводка 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	$Y_i = \alpha + \beta X_i + u$	$\beta$	$\beta \left( \frac{X_i}{Y_i} \right)$
Двойная логарифмическая	$\ln Y_i = \alpha + \beta \ln X_i + u$	$\beta \left( \frac{Y_i}{X_i} \right)$	$\beta$
Линейно-логарифмическая (ln X)	$Y_i = \alpha + \beta \ln X_i + u$	$\beta \left( \frac{1}{X_i} \right)$	$\beta \left( \frac{1}{Y_i} \right)$
Логарифмически-линейная (ln Y)	$\ln Y_i = \alpha + \beta X_i + u$	$\beta Y_i$	$\beta X_i$
Обратная	$Y_i = \alpha + \beta \frac{1}{X_i} + u$	$-\beta \left( \frac{1}{X_i^2} \right)$	$-\beta \left( \frac{1}{X_i Y_i} \right)$

# Сводка линеаризующих преобразований для основных зависимостей в экономике

Вид уравнения регрессии	Постоянство отношения приращения	Функциональная сетка, выпрямляющая график	Система уравнений для определения параметров по МНК
1	2	3	4
$\hat{y} = kx + b$	$\frac{\Delta y}{\Delta x} = const$	-	$\begin{cases} bn + k \sum x_i = \sum y_i \\ b \sum x_i + k \sum x_i^2 = \sum x_i y_i \end{cases}$

# Сводка линеаризующих преобразований для основных зависимостей в экономике

1	2	3	4
$\hat{y} = ax^2 + bx + c$	$\frac{\Delta^2 y}{\Delta^2 x} = const$	.	$\begin{cases} cn + b \sum x_i + a \sum x_i^2 = \sum y_i \\ c \sum x_i + b \sum x_i^2 + a \sum x_i^3 = \sum y_i x_i \\ c \sum x_i^2 + b \sum x_i^3 + a \sum x_i^4 = \sum y_i x_i^2 \end{cases}$

# Сводка линеаризующих преобразований для основных зависимостей в экономике

1	2	3	4
$\hat{y} = \frac{k}{x} + b$	$\frac{\Delta y}{\Delta \frac{1}{x}} = \text{const}$	$u = \frac{1}{x},$ $v = y$	$\begin{cases} bn + k \sum u_i = \sum v_i \\ b \sum u_i + k \sum u_i^2 = \sum u_i v_i \end{cases}$

# Сводка линеаризующих преобразований для основных зависимостей в экономике

1	2	3	4
$\hat{y} = ke^{-x} + b$	$\frac{\Delta y}{\Delta e^{-x}} = const$	$u = e^{-x},$ $v = y$	$\begin{cases} bn + k \sum u_i = \sum v_i \\ b \sum u_i + k \sum u_i^2 = \sum u_i v_i \end{cases}$

# Сводка линеаризующих преобразований для основных зависимостей в экономике

1	2	3	4
$\hat{y} = k \ln x + b$	$\frac{\Delta y}{\Delta \ln x} = const$	$u = \ln x,$ $v = y$	$\begin{cases} bn + k \sum u_i = \sum v_i \\ b \sum u_i + k \sum u_i^2 = \sum u_i v_i \end{cases}$

# Сводка линеаризующих преобразований для основных зависимостей в экономике

1	2	3	4
$\hat{y} = \frac{1}{\frac{k}{x} + b}$	$\frac{\Delta \frac{1}{y}}{\Delta \frac{1}{x}} = \text{const}$	$u = \frac{1}{x},$ $v = \frac{1}{y}$	$\begin{cases} bn + k \sum u_i = \sum v_i \\ b \sum u_i + k \sum u_i^2 = \sum u_i v_i \end{cases}$

# Сводка линеаризующих преобразований для основных зависимостей в экономике

1	2	3	4
$\hat{y} = \frac{1}{ke^{-x} + b}$	$\frac{\Delta \frac{1}{y}}{\Delta e^{-x}} = const$	$u = e^{-x},$ $v = \frac{1}{y}$	$\begin{cases} bn + k \sum u_i = \sum v_i \\ b \sum u_i + k \sum u_i^2 = \sum u_i v_i \end{cases}$

# Сводка линеаризующих преобразований для основных зависимостей в экономике

1	2	3	4
$\hat{y} = bx^k$	$\frac{\Delta \ln y}{\Delta \ln x} = \text{const}$	$u = \ln x,$ $v = \ln y$	$\begin{cases} \ln b \cdot n + k \sum u_i = \sum v_i \\ \ln b \cdot \sum u_i + k \sum u_i^2 = \sum u_i v_i \end{cases}$

# Сводка линеаризующих преобразований для основных зависимостей в экономике

1	2	3	4
$\hat{y} = be^{kx}$	$\frac{\Delta \ln y}{\Delta x} = \text{const}$	$u = x,$ $v = \ln y$	$\begin{cases} \ln b \cdot n + k \sum u_i = \sum v_i \\ \ln b \cdot \sum u_i + k \sum u_i^2 = \sum u_i v_i \end{cases}$

# Сводка линеаризующих преобразований для основных зависимостей в экономике

1	2	3	4
$\hat{y} = be^{\frac{k}{x}}$	$\frac{\Delta \ln y}{\Delta \frac{1}{x}} = const$	$u = \frac{1}{x},$ $v = \ln y$	$\begin{cases} \ln b \cdot n + k \sum u_i = \sum v_i \\ \ln b \cdot \sum u_i + k \sum u_i^2 = \sum u_i v_i \end{cases}$

# Сводка линеаризующих преобразований для основных зависимостей в экономике

1	2	3	4
$\hat{y} = cx^a e^{bx}$	$v = au + bx + \ln c$	$u = \ln x,$ $v = \ln y$	$\begin{cases} \ln c \cdot n + b \sum x_i + a \sum \ln x_i = \sum v_i \\ \ln c \cdot \sum \ln x_i + b \sum x_i \ln x_i + a \sum \ln^2 x_i = \sum v_i \ln x_i \\ \ln c \cdot \sum x_i + b \sum x_i^2 + a \sum x_i \ln x_i = \sum v_i x_i \end{cases}$

# Преобразование случайного отклонения

МНК применяется к преобразованным (линеаризованным) уравнениям. Поэтому необходимо особое внимание уделять рассмотрению свойств случайных отклонений – выполнимости предпосылок теоремы Гаусса-Маркова.

Пример.

$$Y = \alpha X^\beta + \varepsilon \xrightarrow{\ln(\cdot)} \ln Y = \ln(\alpha X^\beta + \varepsilon)$$

Логарифмирование нелинейной модели с аддитивным случайным членом не приводит к линеаризации соотношения относительно параметров.

# Признаки качественной модели

---

1. *Простота модели* (из примерно одинаково отражающих реальность моделей, выбирается та, которая содержит меньше объясняющих переменных).
2. *Единственность* (для любых данных коэффициенты модели должны вычисляться однозначно).
3. *Максимальное соответствие* (модель тем лучше, чем больше скорректированный коэффициент детерминации).
4. *Согласованность с теорией* (уравнение регрессии должно соответствовать теоретическим предпосылкам).
5. *Прогнозные качества* (прогнозы, полученные на основе модели, должны подтверждаться реальностью).

# Сравнение различных моделей

---

1. Содержательный анализ
2. Формальный анализ:
  - *Метод Зарембки*
  - *Преобразование Бокса-Кокса*

# Метод Зарембки

---

Применим для выбора из двух форм (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

# Сравнение различных моделей парной регрессии методом Зарембки

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

$$Y_i^* = Y_i / \sqrt[n]{Y_1 Y_2 \dots Y_n} = Y_i / e^{\frac{1}{n} \sum_{i=1}^n \ln Y_i}$$

2. Рассчитываются линейная и логарифмическая регрессии, и сравниваются значения их сумм квадратов остатков ( $RSS$ )

$$Y_i^* = \alpha_1 + \beta_1 X_i + \varepsilon_i \quad (RSS_1) \quad \ln Y_i^* = \alpha_2 + \beta_2 \ln X_i + \varepsilon'_i \quad (RSS_2)$$

# Сравнение различных моделей парной регрессии методом Зарембки

3. Вычисляем  $\chi^2$ -статистику для оценки значимости различий

$$\chi^2 = \left( \frac{n}{2} \right) \cdot \left| \ln \frac{SSR_1}{SSR_2} \right|$$

4. Сравниваем с критическим значением  $\chi^2$ -распределения  $\chi_{\alpha; m}^2$ . Различия значимы на уровне значимости  $\alpha$ , если

$$\chi^2 > \chi_{\alpha; m}^2$$

# Метод Бокса-Кокса

Идея метода. Переменная  $\frac{Y^\lambda - 1}{\lambda}$  :

при  $\lambda=1$  превращается в линейную функцию  $\frac{Y - 1}{1}$

при  $\lambda \rightarrow 0$  переходит в логарифм  $\frac{Y^\lambda - 1}{\lambda} \rightarrow \ln Y$

Плавно изменяя  $\lambda$ , можно постепенно перейти от линейной регрессии к логарифмической, все время сравнивая качество

# Сравнение различных моделей парной регрессии методом Бокса-Кокса

1. Преобразуют зависимую переменную по методу Зарембки:

$$Y_i^* = Y_i / \sqrt[n]{Y_1 Y_2 \otimes Y_n} = Y_i / e^{\frac{1}{n} \sum_{i=1}^n \ln Y_i}$$

2. Рассчитывают новые переменные (преобразование Бокса-Кокса) при значениях  $\lambda$  от 1 до 0:

$$Y_i^{(B-C)} = \frac{Y_i^{*\lambda} - 1}{\lambda} \qquad X_i^{(B-C)} = \frac{X_i^\lambda - 1}{\lambda}$$

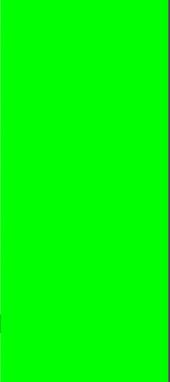
# Сравнение различных моделей парной регрессии методом Бокса-Кокса

---

3. Рассчитывают уравнения регрессии для новых переменных при значениях  $\lambda$  от 1 до 0:

$$Y_i^{(B-C)} = \alpha + \beta X_i^{(B-C)} + \varepsilon_i$$

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



---

Конец лекции