



Экономико-математическое моделирование

Магомедов Магомед-Рамзан Бувайсарович,
кандидат экономических наук, доцент

Основы статистического моделирования



1. Обозначения
2. Парная регрессия
3. Метод наименьших квадратов,
4. Коэффициент детерминации

Основные типы кривых, используемые при количественной оценке связей между двумя переменными

$$a) \bar{y}_x = a + b \cdot x;$$

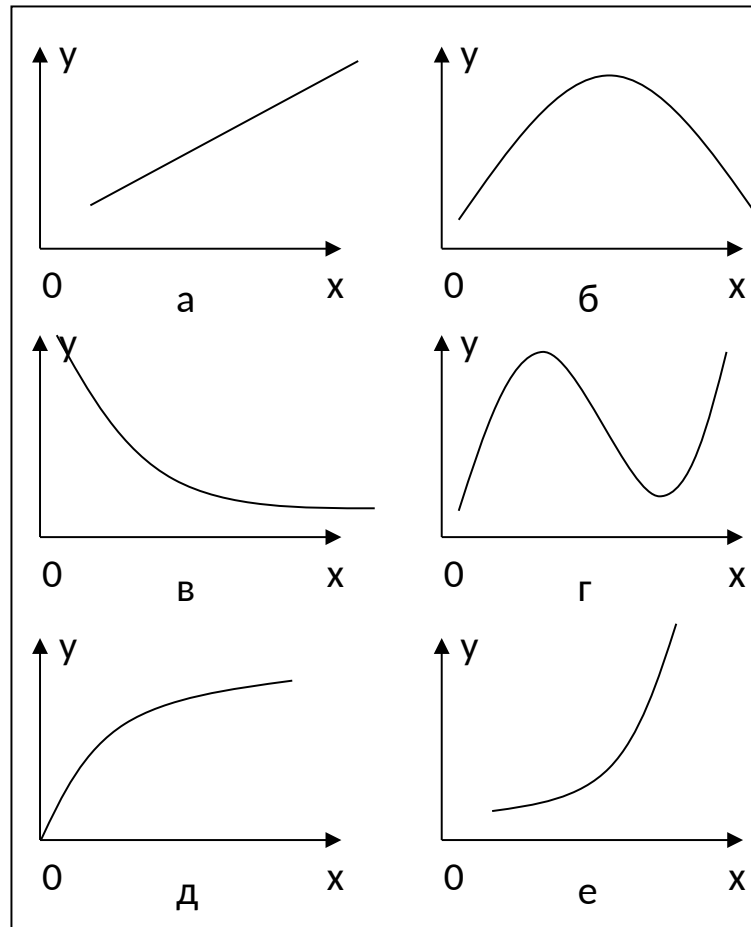
$$б) \bar{y}_x = a + b \cdot x + c \cdot x^2$$

$$в) \bar{y}_x = a + b / x;$$

$$г) \bar{y}_x = a + b \cdot x + c \cdot x^2 + d \cdot x^3$$

$$д) \bar{y}_x = a \cdot x^b;$$

$$е) \bar{y}_x = a \cdot b^x$$

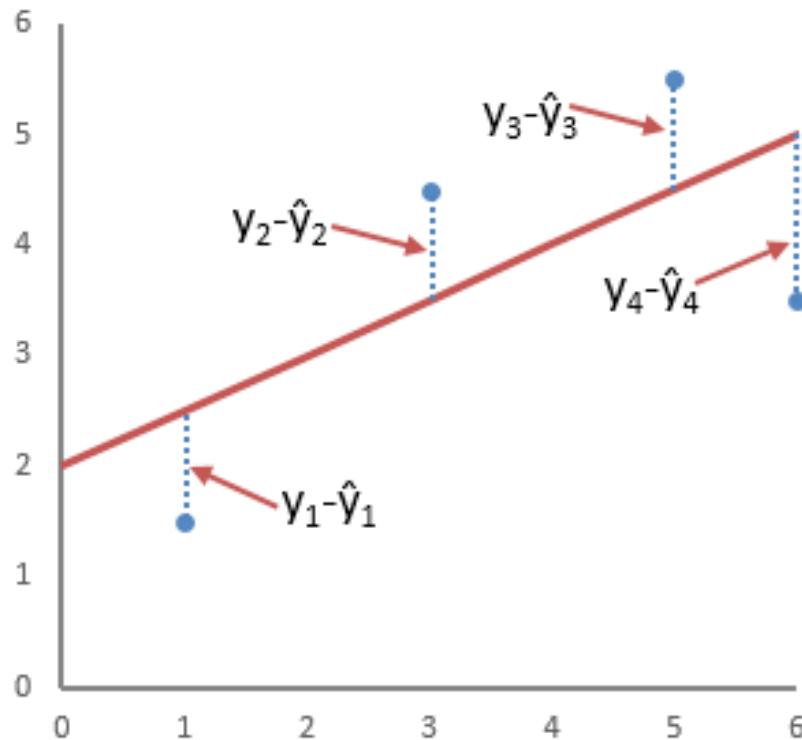


Метод наименьших квадратов (1)



Пусть имеется набор из n наблюдений (y, x) , $i=1, 2, \dots, n$.

Задача: подобрать функцию, которая наилучшим образом описывает зависимость y от x .





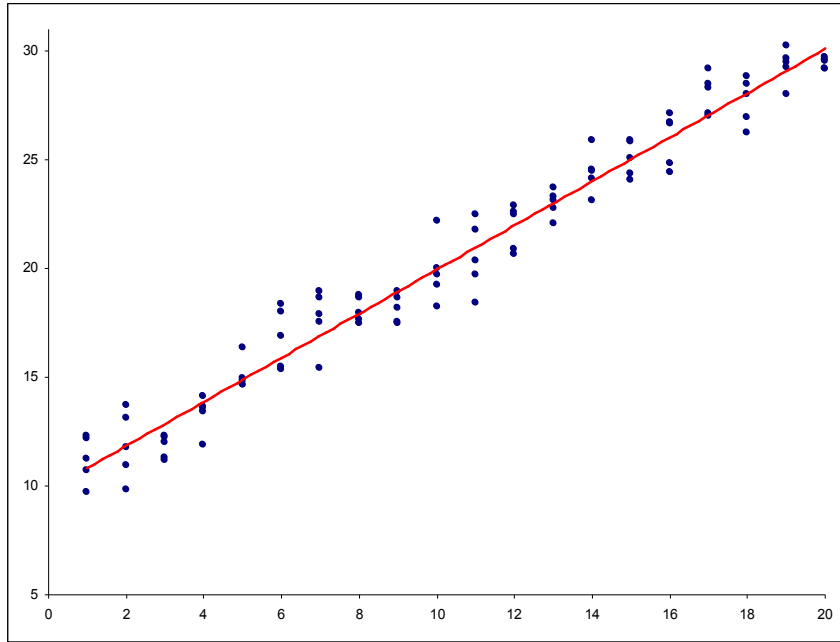
Парная регрессия

Пусть имеется информация о двух переменных: x и y . Выборка из n наблюдений переменной $\mathbf{x}$ ($x_1, x_2, \dots, x_n$) и n наблюдений переменной $\mathbf{y}$ ($y_1, y_2, \dots, y_n$).

Мы предполагаем, что переменные $\mathbf{x}$ и $\mathbf{y}$ взаимосвязаны и хотим проверить это предположение.

Парный регрессионный анализ как раз и заключается в исследовании зависимости между парой переменных.

Модель парной регрессии



В статистических
данных редко
встречаются точные
линейные соотношения:
 $y_i = a + bx_i$

Обычно они бывают
приблизженными:
 $y_i \approx a + bx_i$

<= Как на этом графике

Модель парной регрессии



Приблизительные взаимосвязи вида

$$y_i = a + bx_i$$

специалисты обычно описывают следующим образом:

$$y_i \approx a + bx_i + \varepsilon_i$$

y_i — значения зависимой переменной

x_i — значения независимой переменной
(регрессора)

ε_i — случайные ошибки

Модель парной регрессии



$$y_i = a + bx_i + \varepsilon_i$$

Откуда берутся случайные ошибки?

1. Существуют другие, неучтенные в нашей упрощенной модели факторы. Эти факторы также оказывают влияние на зависимую переменную ***y***
2. Присутствуют ошибки измерений зависимой переменной

Модель парной регрессии



В чем разница между a и b ?

$$y_i = a + bx_i + \varepsilon_i$$

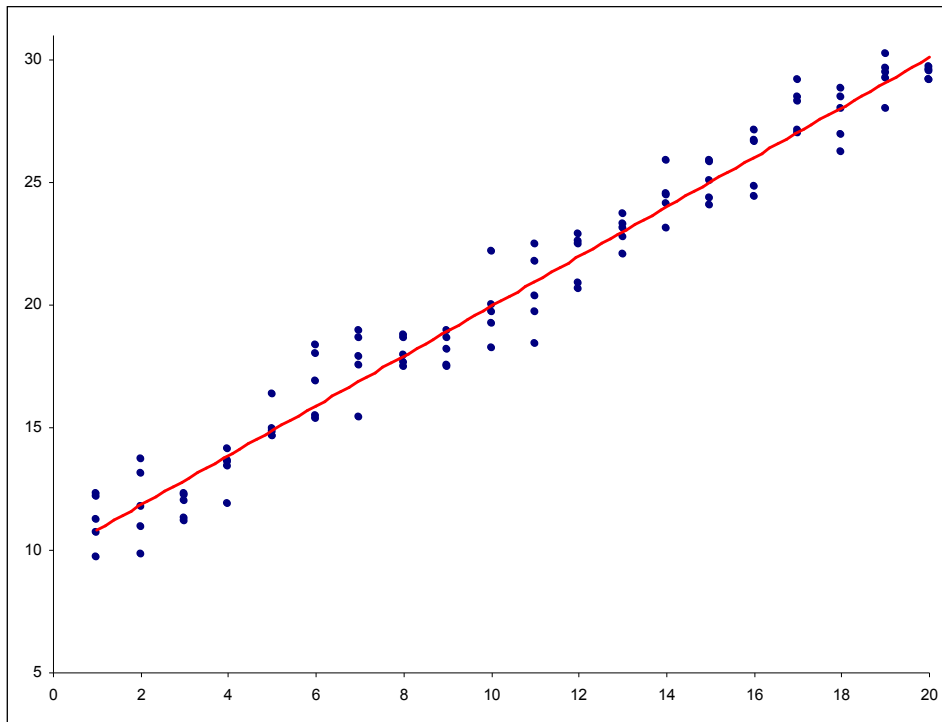
a и b — истинные значения параметров модели, которые на практике **никогда не известны** исследователю

$$\check{y}_i = a + bx_i$$

a и b — их оценки, полученные при помощи МНК, на основе случайной выборки

Следовательно, a и b — случайные величины

Модель парной регрессии



В статистических данных редко встречаются точные линейные соотношения:

$$y_i = a + bx_i$$

Обычно они бывают приближенными:

$$y_i \approx a + bx_i$$

<= Как на этом графике

Модель парной регрессии

Приблизительные взаимосвязи вида

$$y_i \approx a + bx_i ,$$

эконометристы обычно описывают следующим образом:

$$y_i = a + bx_i + \varepsilon_i .$$

y_i — значения зависимой переменной

x_i — значения независимой переменной (регрессора)

ε_i — случайные ошибки

Модель парной регрессии

$$y_i = a + bx_i + \varepsilon_i$$

Откуда берутся случайные ошибки ε_i ?

1. Существуют другие, неучтенные в нашей упрощенной модели факторы. Эти факторы также оказывают влияние на зависимую переменную y
2. Присутствуют ошибки измерений зависимой переменной

Модель парной регрессии

В чем разница между a и b ?

$$y_i = a + bx_i + \varepsilon_i$$

a и b — истинные значения параметров модели, которые на практике **никогда не известны** исследователю

$$\hat{y}_i = a + bx_i$$

a и b — их оценки, полученные при помощи МНК, на основе случайной выборки

Следовательно, a и b — случайные величины

Модель парной регрессии

Разумеется, хочется, чтобы полученные оценки a и b были близки к истинным значениям.

При каких условиях можно на это надеяться?

Эти условия называют **предпосылками классической линейной модели парной регрессии (КЛМПР)**

Предпосылки КЛМНР

(1) Модель линейна по параметрам и правильно специфицирована

$$y_i = a + bx_i + \varepsilon_i$$

Предпосылки КЛМПР

(2) $x_1, \dots, x_n$ — детерминированные (неслучайные)
величины (не все одинаковые)

Предпосылки КЛМПР

(3) Математическое ожидание случайных ошибок

равно нулю: $E(\varepsilon_i) = 0$

Предпосылки КЛМПР

(4) случайные ошибки имеют постоянную

дисперсию: $V(\varepsilon_i) = \sigma^2 = \text{const}$

Предпосылки КЛМПР

- (5) Случайные ошибки, соответствующие разным наблюдениям не зависят друг от друга (не коррелированы)

Предпосылки КЛМПР

(6)* Случайные ошибки ε_i имеют нормальное распределение



Теорема Гаусса — Маркова

Если выполнены условия (1)–(5), **то** оценки a и b , полученные по методу наименьших квадратов (МНК), являются:

(а) несмещенными

(б) эффективными, то есть имеют наименьшую дисперсию в классе всех линейных несмещенных оценок.

Теорема Гаусса — Маркова

- Часто ли на практике выполняются все шесть предпосылок классической линейной модели?
- Нет.

Поэтому в рамках курса мы обсудим, что делать в случае их нарушения.

Оценка дисперсии случайной ошибки

Если выполнены условия (1)–(6), то несмещенная оценка дисперсии случайных ошибок σ^2 имеет вид

$$\hat{\sigma}^2 = s^2 = \frac{1}{n-2} \sum_{i=1}^n e_i^2$$

s^2 используется для расчета стандартных ошибок коэффициентов

Стандартные ошибки коэффициентов

Используются для измерения точности полученных оценок: чем ниже стандартные ошибки, тем точнее оценки коэффициентов

В модели парной регрессии стандартные ошибки (Standard errors) коэффициентов оцениваются по следующим формулам (откуда эти формулы обсудим немного позже):

$$SE(\hat{\beta}_1) = \sqrt{\frac{s^2}{n} \left(1 + \frac{\bar{x}^2}{Var(x)} \right)}, \quad SE(\hat{\beta}_2) = \sqrt{\frac{s^2}{\sum (x_i - \bar{x})^2}}$$

Тестирование значимости коэффициента

a и b — оценки, полученные при помощи МНК, на основе случайной выборки. Следовательно, они сами являются случайными величинами.

Поэтому даже, если истинное значение коэффициента b равно нулю, его оценка b может отклоняться от нуля.

Нужно уметь определять, достаточно ли сильно b отличается от нуля для того, чтобы можно было с уверенностью утверждать, что и истинное значение коэффициента также не равно нулю?

Тестирование значимости коэффициента

Нужно уметь определять, достаточно ли сильно b отличается от нуля для того, чтобы можно было с уверенностью утверждать, что и истинное значение коэффициента также не равно нулю?

На практике для решения этой задачи используется тест на значимость коэффициента

Тестирование значимости коэффициента

Рассматриваемая модель $y_i = a + bx_i + \varepsilon_i$

Тестируемая гипотеза

$H_0: b = 0$ «Переменная **x** не оказывает значимого влияния на переменную **y**»

Альтернативная гипотеза

$H_1: b \neq 0$ «Переменная **x** оказывает значимое влияние на переменную **y**»

Тестирование значимости коэффициента Способ 1

Алгоритм проведения теста

Шаг 1

Вычисляем расчетное значение t-статистики

$$t_{расч} = \frac{b}{SE(b)}$$

Шаг 2

Выбираем **уровень значимости**

Уровень значимости — **вероятность ошибки** первого рода, то есть вероятность отклонить гипотезу H_0 , если на самом деле гипотеза H_0 верна.

В эконометрике обычно используется уровень значимости $\alpha = 0,01 = 1\%$ или $\alpha = 0,05 = 5\%$.

Тестирование значимости коэффициента Способ 1

Шаг 3

Из таблиц t-распределения Стьюдента находим критическое значение t-статистики $t_{кр}$.

Оно зависит от уровня значимости (двусторонний тест) α и от так называемого числа степеней свободы, которое в случае нашего теста равно $(n - 2)$.

Тестирование значимости коэффициента Способ 1

Шаг 4

Сравниваем расчетное и критическое значение t-статистик

Если $\left| t_{расч} \right| < t_{кр}$,

то гипотеза H_0 не отклоняется (принимается),

то есть мы делаем вывод о том, что переменная x не оказывает значимого влияния на переменную y .

В этом случае коэффициент при переменной x называют незначимым.

В противном случае гипотеза H_0 не принимается (отклоняется).

Literature

- **Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика. Начальный курс: Учеб. — 6-е изд., перераб. и доп. — М.: Дело, 2004.**
Глава 2, параграфы 2.1–2.6
- **Доугерти К. Введение в эконометрику: Учебник. 3-е изд. / Пер. с англ. — М.: ИНФРА-М, 2009. Глава 1**

Узнайте больше о ФТИМО!

Спасибо за
ВНИМАНИЕ

Контакты:

Сайт: <https://clck.ru/dyFGK>

E-mail: ftmo@chspu.ru

Telegram-канал: https://t.me/ftimo_chspu

ВКонтакте: https://vk.com/ftmo_chspu

