

Структура курса

- Основные понятия и определения эконометрики. Эконометрическое моделирование.
- Парная линейная регрессионная модель.
- Множественная линейная регрессионная модель.
- Статистические свойства МНК-оценок МЛРМ.
- Проверка гипотез относительно возможных значений коэффициентов МЛРМ.
- Мультиколлинеарность.
- Ошибки спецификации.
- Процедуры отбора регрессоров.
- Обобщенный метод наименьших квадратов
- Гетероскедастичность.
- Автокорреляция
- Прогнозирование при помощи МЛРМ.
- Временные ряды
- Системы одновременных уравнений

Необходимые требования и НАВЫКИ

- Операции с векторами и матрицами.
- Дифференциальное и интегральное исчисление.
- Случайные величины. Функция распределения, закон распределения случайной величины Математическое ожидание, дисперсия, моменты распределения, асимметрия, эксцесс.
- Нормальное распределение.
- Предельные теоремы и закон больших чисел.
- Статистическое оценивание неизвестных параметров. Точные и интервальные оценки. Состоятельность, эффективность, несмещенность оценок.
- Проверка статистических гипотез.
- Дисперсионный анализ.

Литература

- Магнус Я. Р., Катышев П. К., Пересецкий А. А. Эконометрика. Начальный курс. (любое издание).
- Доугерти К. Введение в эконометрику. Москва, 2001.
- Эконометрика. Под ред. И. И. Елисеевой. Москва. Финансы и статистика, 2001.
- Практикум по эконометрике. Под ред. И. И. Елисеевой. Москва, финансы и статистика, 2001
- Айвазян С. А., Мхитарян В. С. Прикладная статистика и основы эконометрики. Москва, Ю ЮНИТИ (любое издание).
- Кремер Н. Ш., Путко Б. А. Эконометрика. М. ЮНИТИ, 2002

Тема 1. Эконометрическое моделирование

- Возникновение эконометрики как науки
- Определение эконометрики
- Прикладные цели эконометрики
- Этапы эконометрического
моделирования

История эконометрики как науки

- 1910, Австро-Венгрия – бухгалтер П. Цьемпа ввел термин «эконометрика»

Цьемпа считал, что если к данным бухгалтерского учета применить методы алгебры и геометрии, то будет получено новое, более глубокое представление о результатах хозяйственной деятельности.

Современное определение эконометрики

Эконометрика – научная дисциплина, объединяющая совокупность теоретических результатов, приемов, методов и моделей, предназначенных для того, чтобы на базе

- экономической теории;
- экономической статистики;
- математико-статистического инструментария

придавать конкретное количественное выражение общим (качественным) закономерностям, обусловленным экономической теорией. (С. А. Айвазян, В. С. Мхитарян. Прикладная статистика и основы эконометрики.)

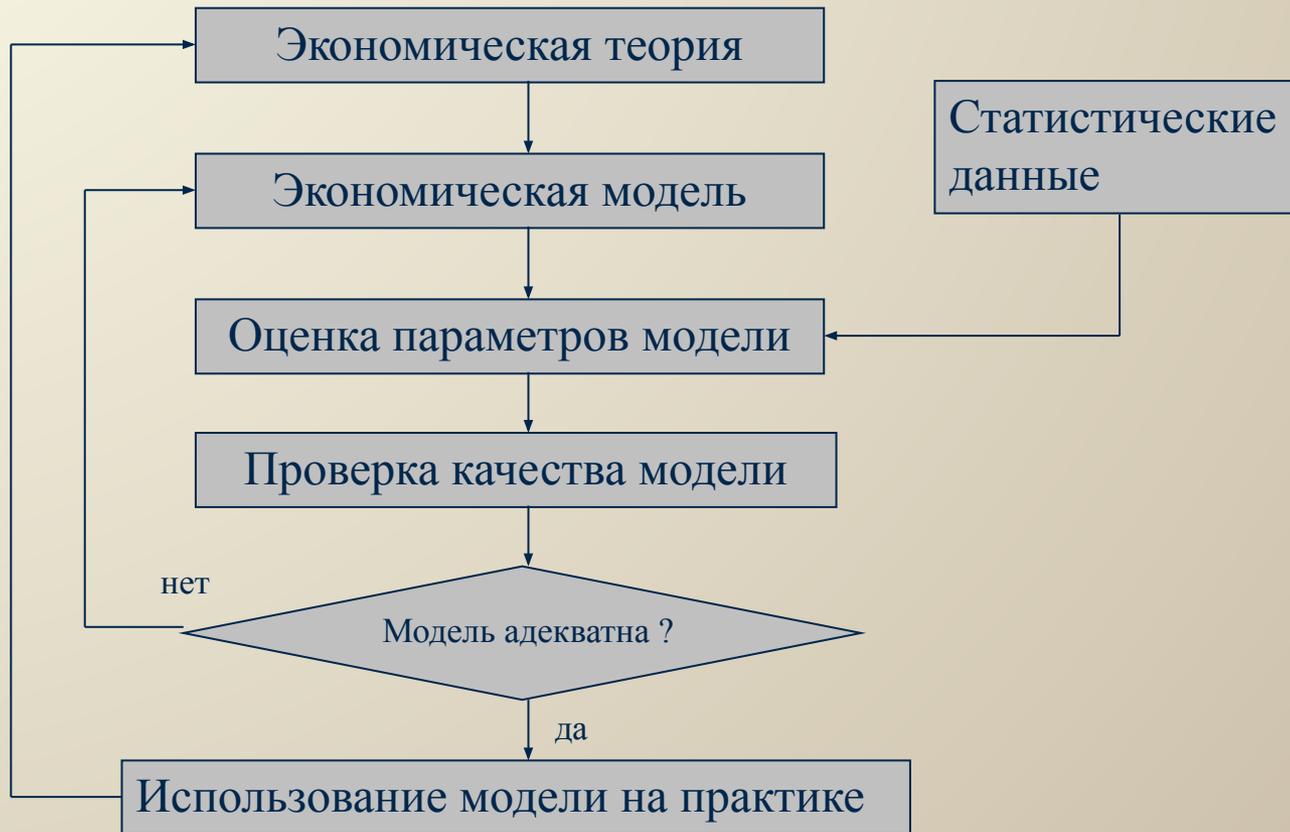
Прикладные цели эконометрики

- вывод экономических законов;
- формулировка экономических моделей, основываясь на экономической теории и эмпирических данных;
- оценка неизвестных величин (параметров) в этих моделях;
- прогнозирование и оценка точности прогноза;
- выработка рекомендаций по экономической политике.

Этапы эконометрического моделирования

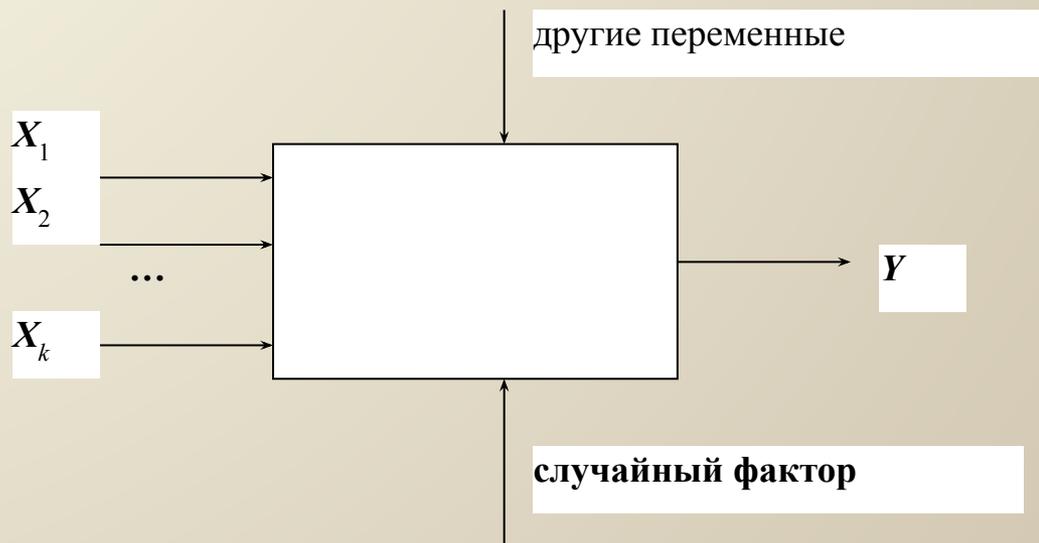
- Осознание того факта, что в экономике многие переменные связаны между собой
- Группировка отдельных соотношений в модель
- Сбор данных
- Идентификация
- Верификация

Этапы эконометрического моделирования



1. Переменные модели

- Переменную, процесс формирования значений которой нас по каким-то причинам интересует, будем обозначать Y и называть зависимой или объясняемой.
- Переменные, которые, как мы предполагаем, оказывают влияние на переменную Y , будем обозначать X_j и называть независимыми или объясняющими.



Другая классификация переменных

- Переменные, значения которых объясняются в рамках нашей модели, называются *эндогенными*.
- Переменные, значения которых нашей моделью не объясняются, являются для нее внешними, ничего о том, как формируются эти значения, мы не знаем, называются *экзогенными*

2. Спецификация модели

- определение цели моделирования;
- определения списка экзогенных и эндогенных переменных;
- определение форм зависимостей между переменными;
- формулировка априорных ограничений на случайную составляющую, что важно для свойств оценок и выбора метода оценивания;
- формулировка априорных ограничений на коэффициенты

Виды эконометрических моделей

- Модели временных рядов.
- Регрессионные модели с одним уравнением.
- Системы одновременных уравнений.

Модели временных рядов.

- Такие модели объясняют поведение переменной, меняющейся с течением времени, исходя только из ее предыдущих значений. К этому классу относятся модели тренда, сезонности, тренда и сезонности (аддитивная и мультипликативная формы) и др.

Регрессионные модели с одним уравнением.

- В таких моделях зависимая (объясняемая) переменная представляется в виде функции от независимых (объясняющих) переменных и параметров. В зависимости от вида функции модели бывают линейными и нелинейными.

Системы одновременных уравнений.

- Ситуация экономическая, поведение экономического объекта описывается системой уравнений. Системы состоят из уравнений и тождеств, которые могут содержать в себе объясняемые переменные из других уравнений (поэтому вводят понятия экзогенных и эндогенных переменных).

3. Сбор данных.

- cross-sectional data – пространственные данные – набор сведений по разным экономическим объектам в один и тот же момент времени;
- time-series data – временные ряды – наблюдение одного экономического параметра в разные периоды или моменты времени. Эти данные естественным образом упорядочены во времени.
- panel data – панельные данные – набор сведений по разным экономическим объектам за несколько периодов времени (данные переписи населения).

4. Идентификация.

- Идентификация модели – статистический анализ модели и, прежде всего – статистическое оценивание параметров. Выбор метода оценивания сюда тоже входит. Зависит от особенностей модели.

5. Верификация.

- Верификация модели – сопоставление реальных и модельных данных, проверка оцененной модели с тем, чтобы прийти к выводу о достаточной реалистичности получаемой с ее помощью картины объекта, либо признать необходимость оценки другой спецификации модели.

Вопросы для самопроверки

- Кто первый ввел в употребление термин «Эконометрика».
- В каком году был основан журнал «Econometrics».
- Каких вы знаете лауреатов нобелевской премии по экономике за достижения в эконометрических методах.
- На каких «трех китах» базируется современная экономическая теория.
- Приведите определение эконометрики, отражающее современный взгляд на эту науку.
- Каковы прикладные цели эконометрики.
- Перечислите основные этапы эконометрического моделирования.
- Что входит в спецификацию модели.
- Что происходит на этапе идентификации модели.
- Какие основные типы экономических данных вы знаете.
- Основные типы эконометрических моделей.
- Как происходит верификация модели

Тема 2. Парная линейная регрессионная модель

ПЛРМ

Две переменные X и Y

могут быть связаны

- функциональной зависимостью (т.е. существует функция f что $Y = f(X)$, значения переменной Y полностью определяются значениями переменной X)
- статистической зависимостью
- независимы.

Статистическая зависимость

- Если при изменении X меняется закон распределения случайной величины Y , то говорят, что величины (X, Y) связаны статистической зависимостью.
- Статистическая зависимость называется корреляционной, если при изменении X меняется математическое ожидание случайной величины Y .

Корреляционная зависимость

Если каждому значению величины X соответствует свое значение $M(Y | X)$

то говорят, что существует
регрессионная функция

$$M(Y | X) = f(X)$$

Случайная составляющая

Отклонение переменной Y от математического ожидания для соответствующего значения переменной X называется ошибкой и обозначается ε

$$\varepsilon(X) = Y(X) - f(X)$$

Регрессионное уравнение

Уравнение

$$Y = f(X) + \varepsilon$$

называется уравнением регрессии
переменной Y на переменную X

ЭКОНОМИЧЕСКИЙ СМЫСЛ ε

- невключение объясняющих переменных в уравнение. На самом деле на переменную Y влияет не только переменная X , но и ряд других переменных, которые не учтены в нашей модели по следующим причинам:
 - мы знаем, что другая переменная влияет, но не можем ее учесть, потому как не знаем, как измерить (психологический фактор, например);
 - существуют факторы, которые мы знаем, как измерить, но влияние их на Y так слабо, что их не стоит учитывать;
 - существенные переменные, но из-за отсутствия опыта или знаний мы их таковыми не считаем.

Экономический смысл ε (продолжение)

- Неправильная функциональная спецификация. Функциональное соотношение между Y и X может быть определено неправильно. Например, мы предположили линейную зависимость, а она может быть более сложной.
- Ошибки наблюдений (занижение реального уровня доходов). В этом случае наблюдаемые значения не будут соответствовать точному соотношению, и существующее расхождение будет вносить свой вклад в остаточный член.

Способы определения регрессионной функции $f(X)$

- параметрический – предполагаем, что вид регрессионной функции известен, неизвестны параметры функции
- непараметрический – предполагаем, что вид регрессионной функции неизвестен и мы составляем алгоритм расчета значений функции в каждой точке

Выбор вида $f(X)$

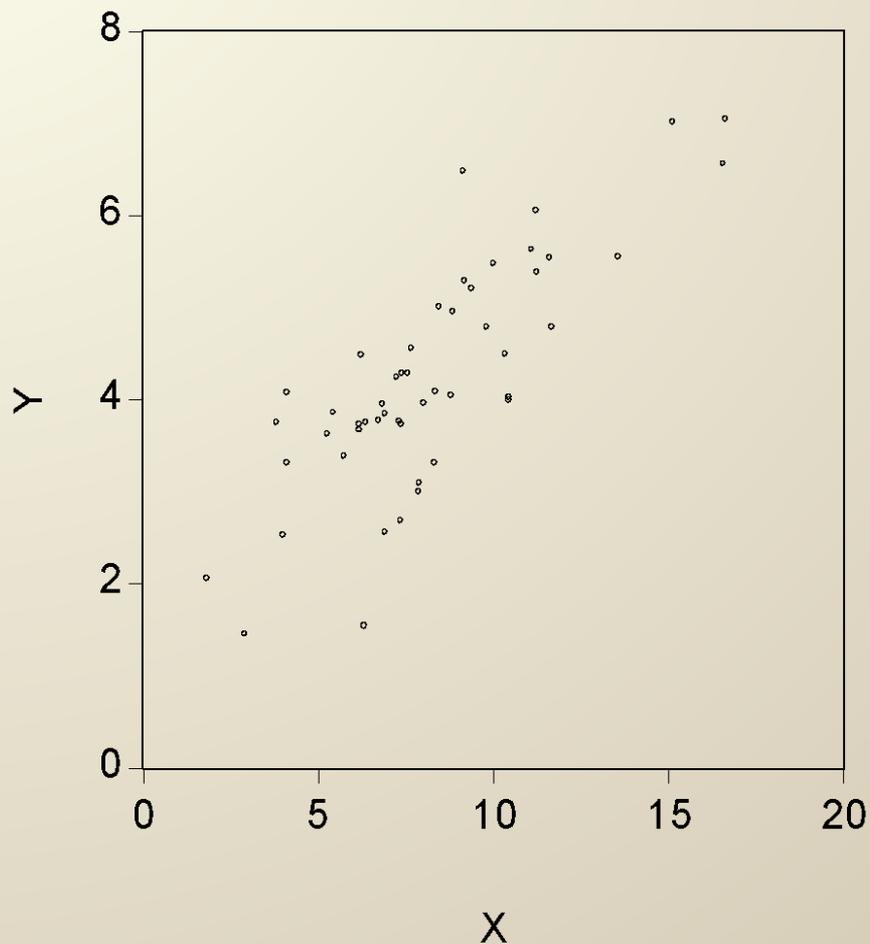
- экономическая теория
- опыт, интуиция исследователя
- эмпирический анализ данных

Эмпирический анализ данных

В парном случае материал наблюдений представляет собой набор пар чисел:

$$(X_i, Y_i) \quad i = 1, \dots, N$$

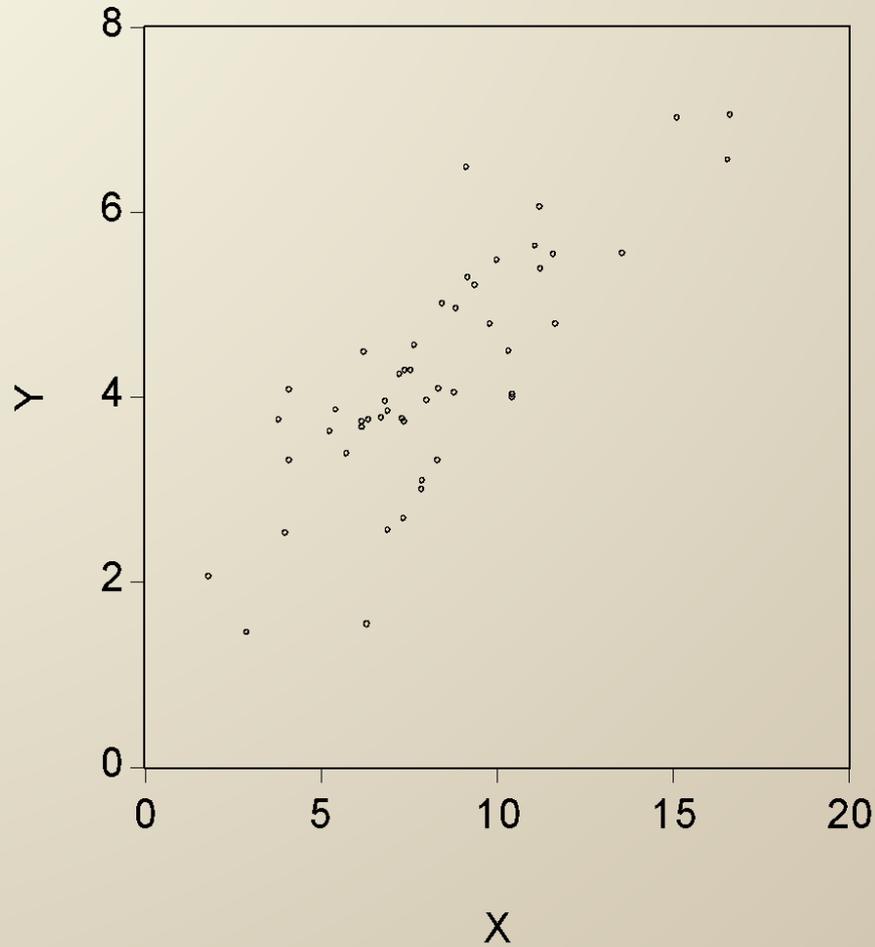
На плоскости каждому такому наблюдению соответствует точка:



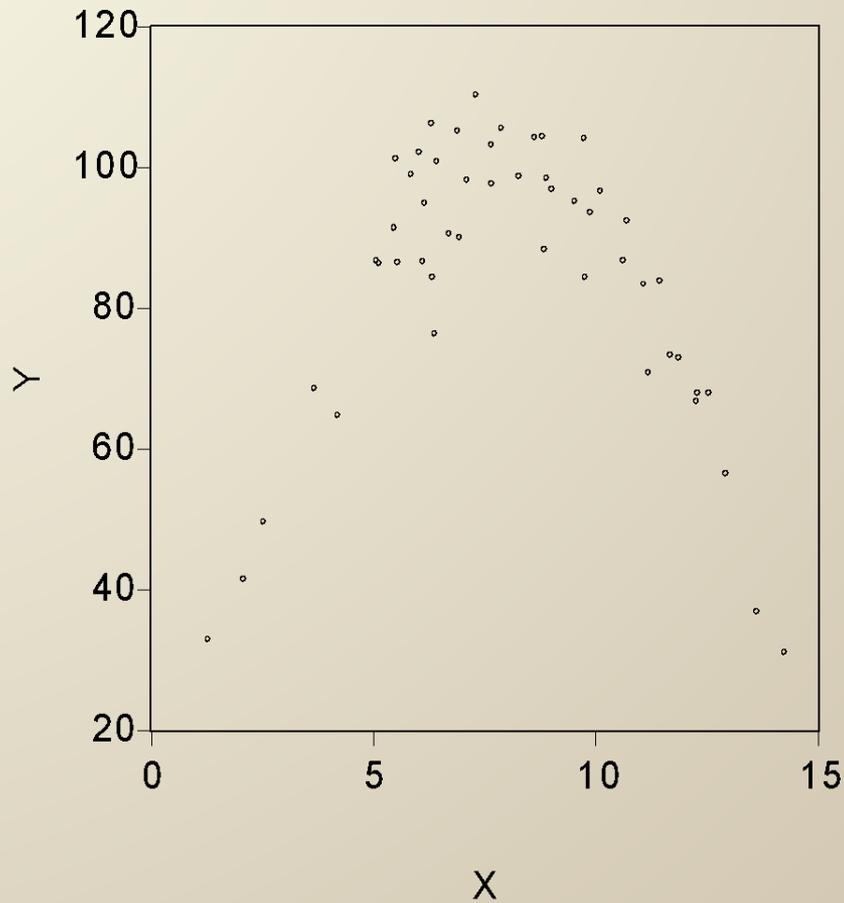
Полученный график называют облако наблюдений, поле корреляции или диаграмма рассеяния. По виду облака наблюдений можно определить вид регрессионной функции.

Линейная

$$Y = \alpha + \beta X + \varepsilon.$$

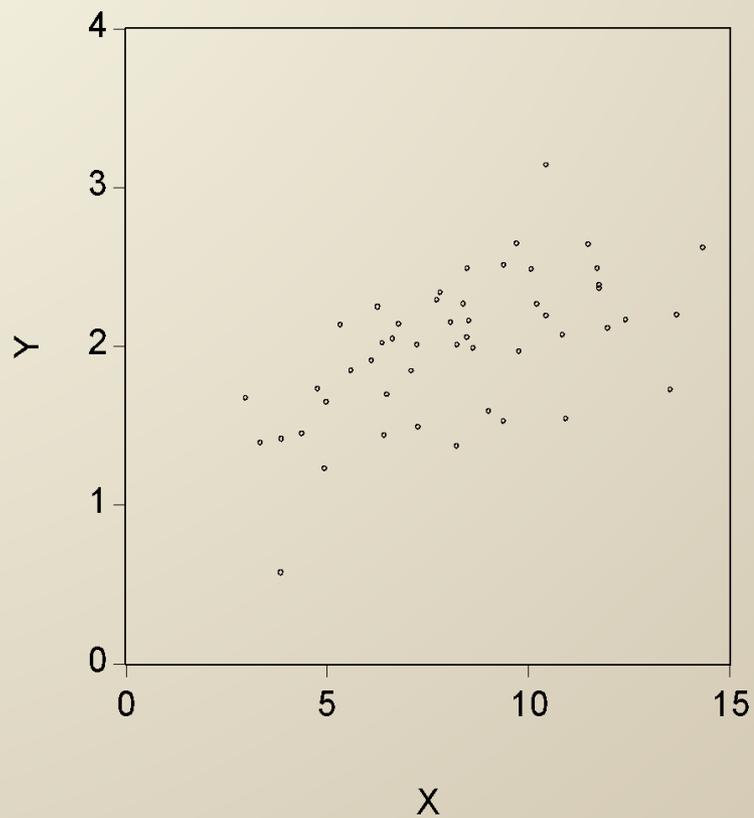


Квадратичная $Y = \alpha + \beta X + \gamma X^2 + \varepsilon$



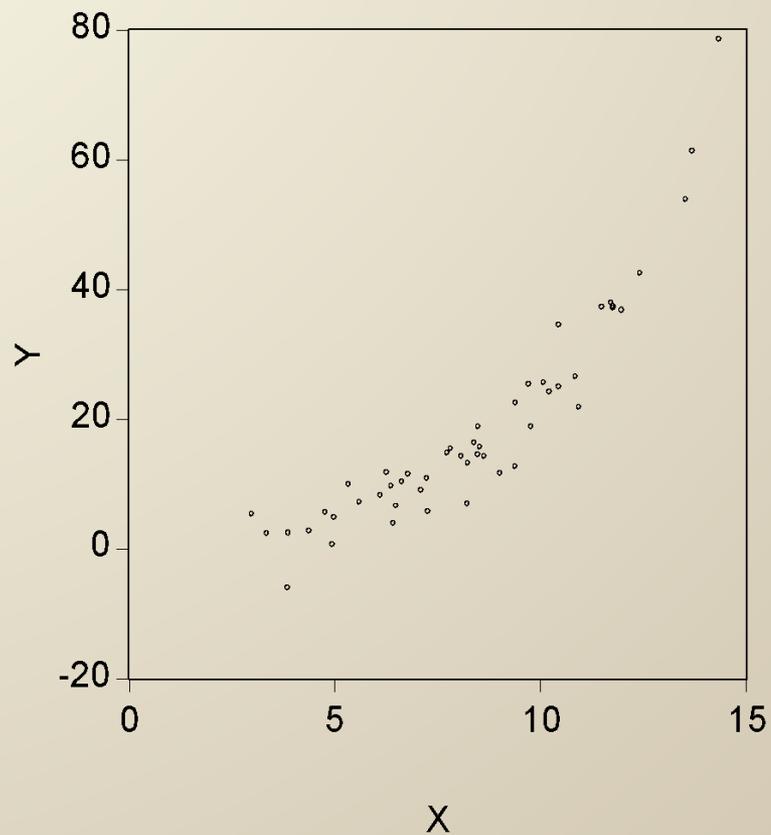
Показательная

$$Y = \alpha X^{\beta} \varepsilon$$



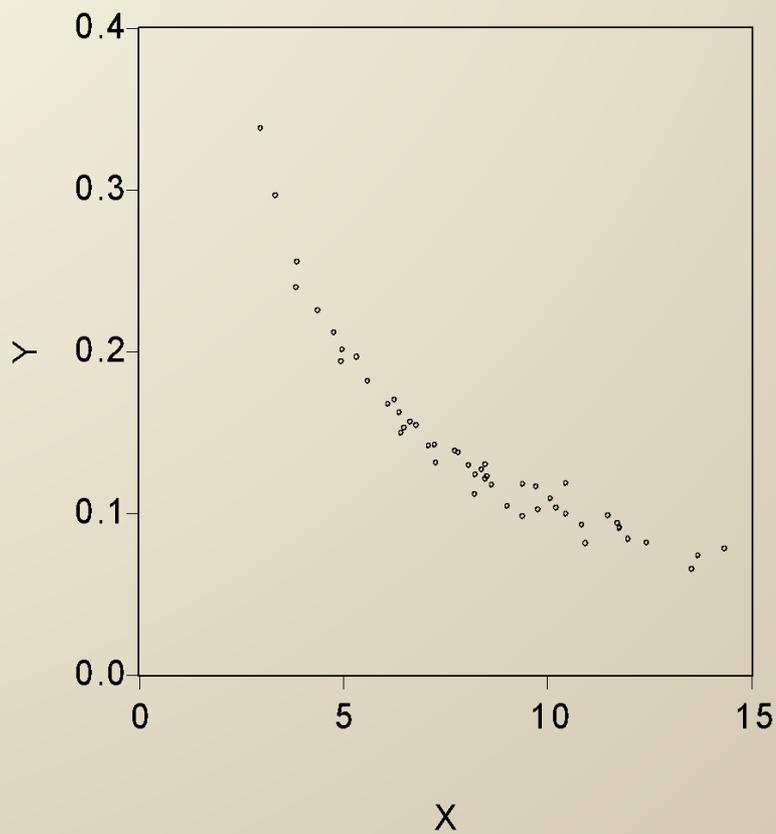
Степенная

$$Y = \alpha e^{\beta X} \varepsilon$$

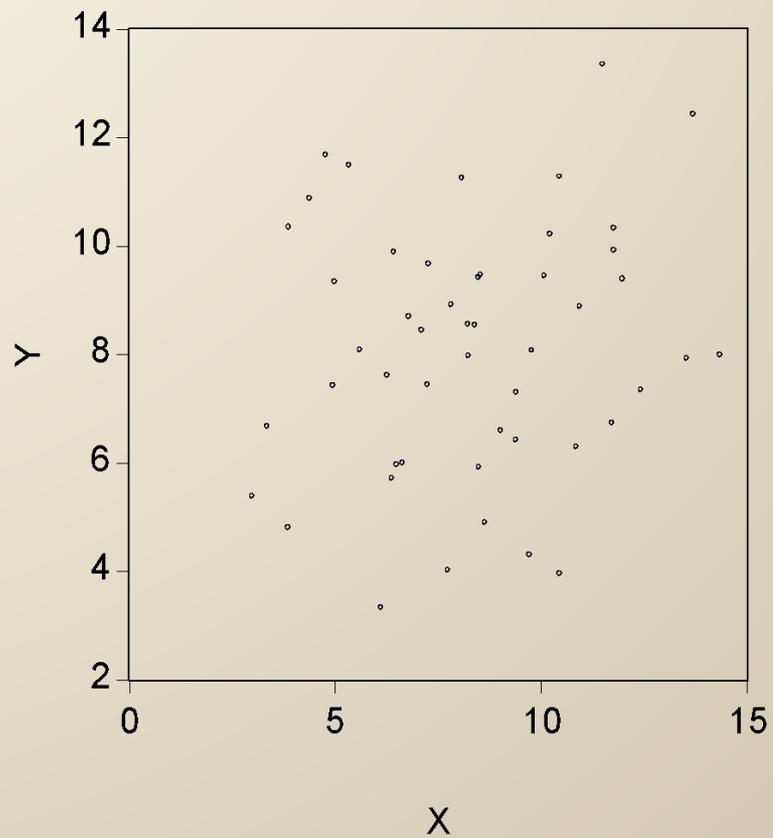


Гиперболическая

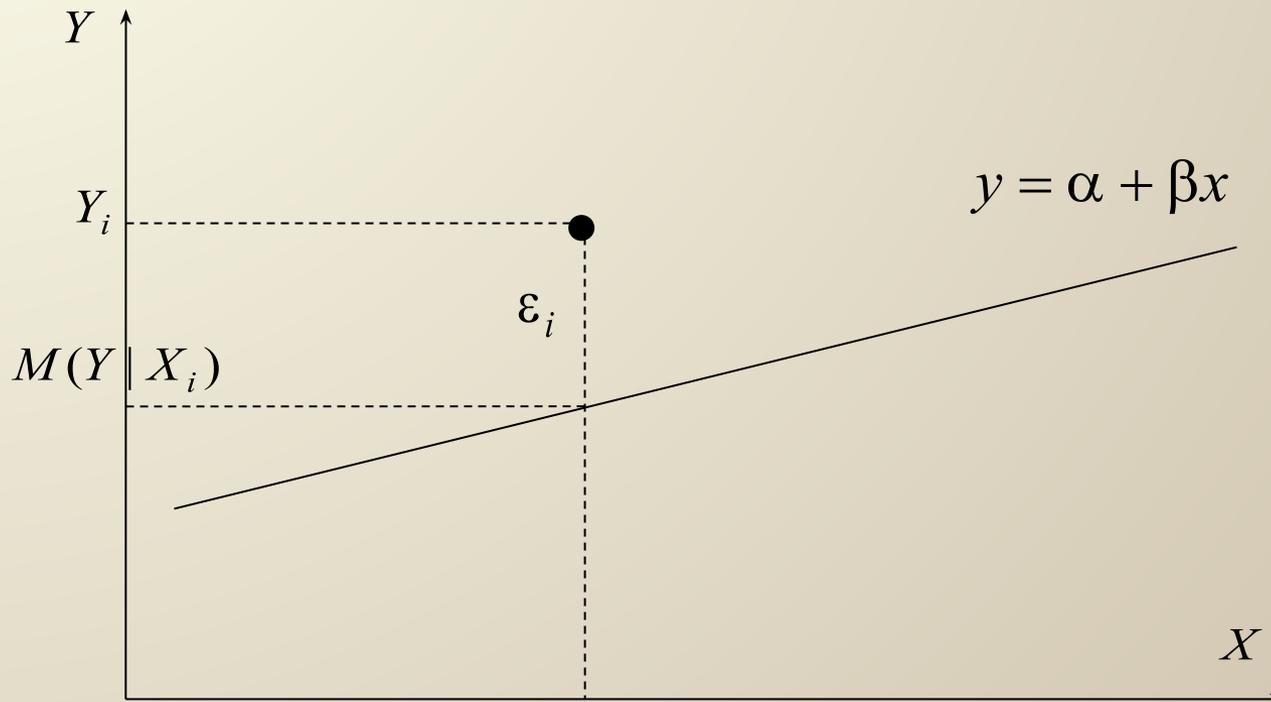
$$Y = \alpha + \frac{\beta}{X} + \varepsilon$$



X и Y независимы

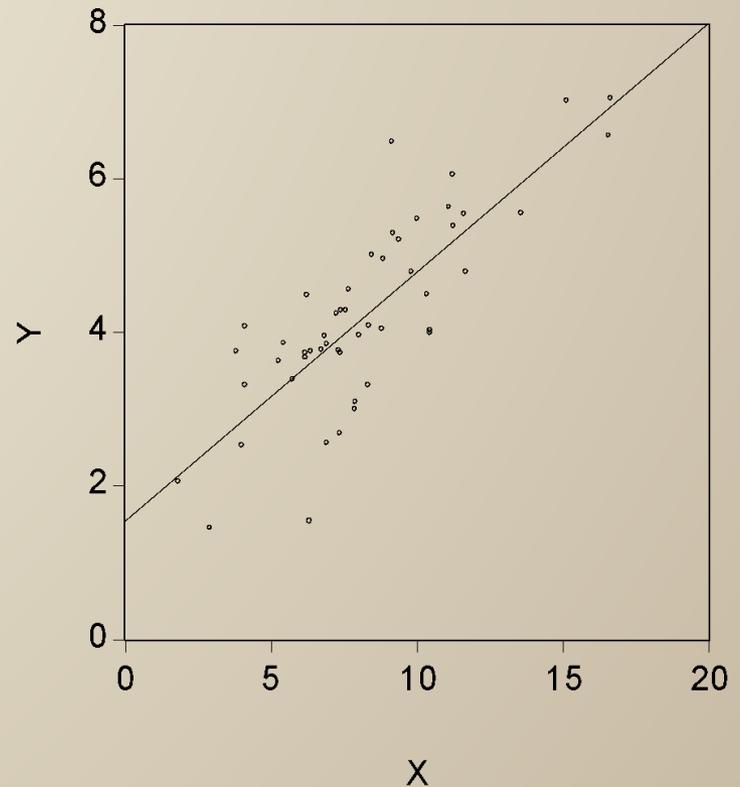


Парная линейная регрессионная модель $Y = \alpha + \beta X + \varepsilon$.

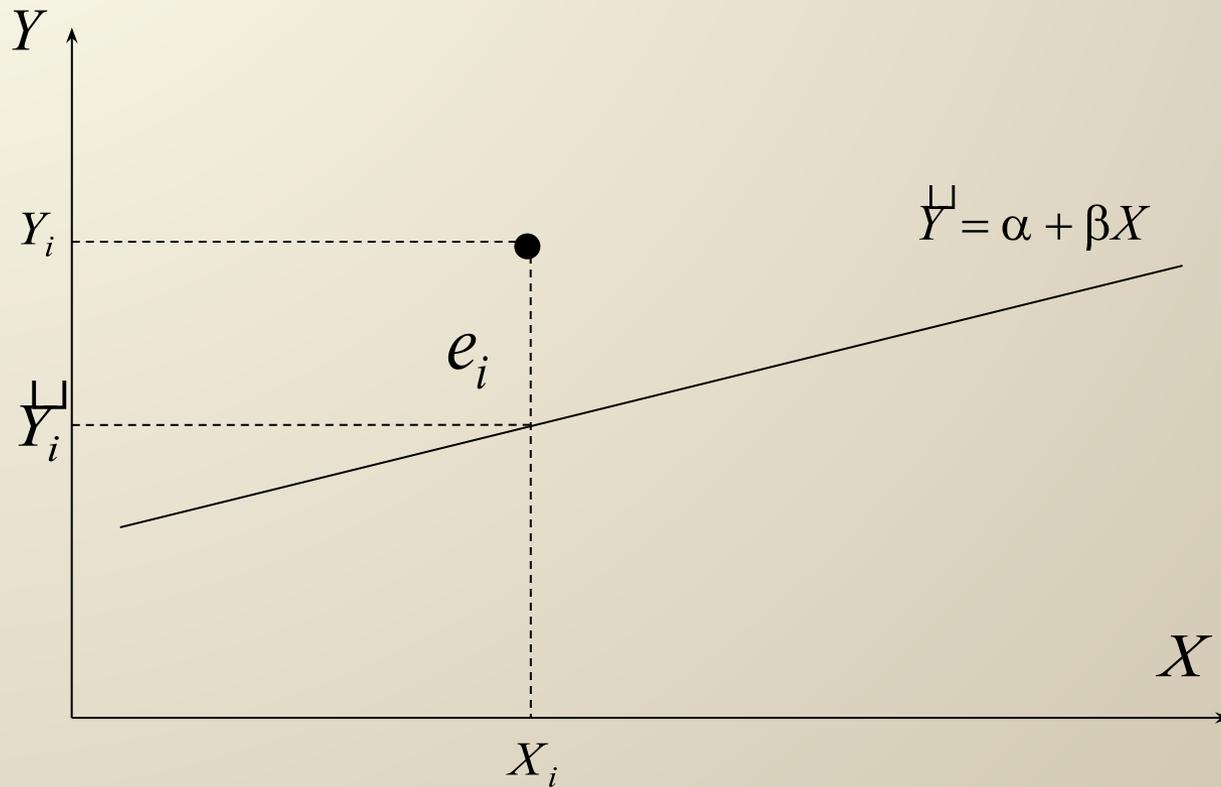


Выбор коэффициентов регрессионной прямой

Из всех возможных прямых мы хотим выбрать ту, чтобы она «наилучшим образом» подходила к нашим данным, т. е. отражала бы линейную зависимость Y от X . Иными словами, чтобы каждое Y_i лежало бы как можно ближе к прямой. Можно сказать, мы хотим, чтобы желаемая прямая была бы в центре скопления наших данных.



Рассмотрение остатков на графике



Интегральная мера близости

$$\sum_{i=1}^N e_i^2 = \sum_{i=1}^N (Y_i - a - bX_i)^2 = S(a, b) \rightarrow \min_{(a, b)}$$

$$\sum_{i=1}^N |e_i| = \sum_{i=1}^N |Y_i - a - bX_i| = S(a, b) \rightarrow \min_{(a, b)}$$

Метод наименьших квадратов

$$\sum_{i=1}^N e_i^2 \rightarrow \min_{(a,b)}$$

Среди всех возможных
прямых выбираем ту,
для которой сумма
квадратов остатков
МИНИМАЛЬНА

Минимизация

$$\frac{\partial S}{\partial a} = -2 \sum_{i=1}^N (Y_i - a - bX_i) = 0$$

$$\frac{\partial S}{\partial b} = -2 \sum_{i=1}^N (Y_i - a - bX_i)X_i = 0$$

$$\begin{cases} \sum_{i=1}^N (Y_i - a - bX_i) = 0 \\ \sum_{i=1}^N (Y_i - a - bX_i)X_i = 0 \end{cases}$$

ИЛИ

$$\begin{cases} \sum_{i=1}^N e_i = \mathbf{0} \\ \sum_{i=1}^N X_i e_i = \mathbf{0} \end{cases}$$

Система нормальных уравнений

$$\begin{cases} Na + b \sum_{i=1}^N X_i = \sum_{i=1}^N Y_i \\ a \sum_{i=1}^N X_i + b \sum_{i=1}^N X_i^2 = \sum_{i=1}^N X_i Y_i \end{cases}$$

МНК-коэффициенты ПЛРМ

$$\beta = \frac{\frac{\sum_{i=1}^N X_i Y_i}{N} - \frac{\sum_{i=1}^N X_i}{N} \frac{\sum_{i=1}^N Y_i}{N}}{\frac{\sum_{i=1}^N X_i^2}{N} - \left(\frac{\sum_{i=1}^N X_i}{N} \right)^2} \quad - \text{ коэффициент наклона}$$

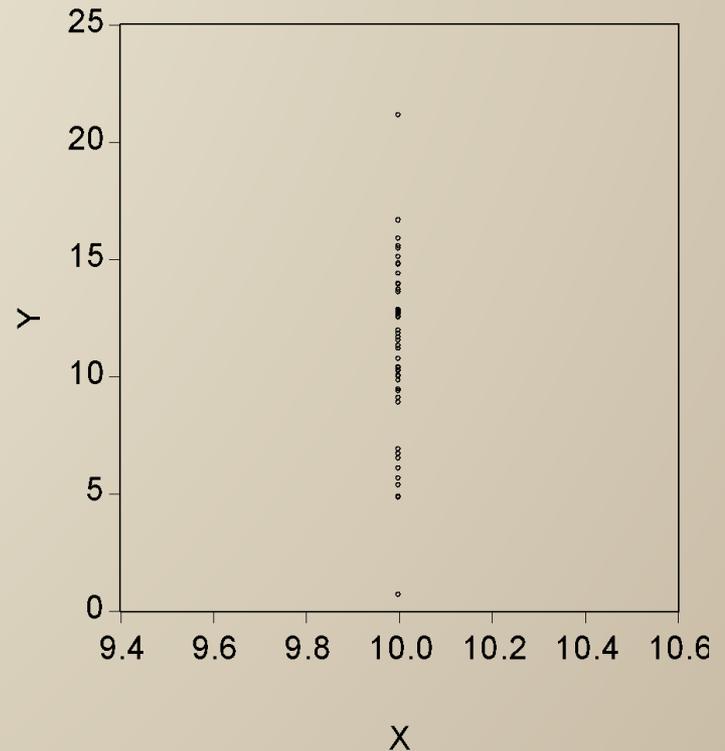
$$\alpha = \bar{Y} - \beta \bar{X} \quad - \text{ свободный коэффициент}$$

Другие формы записи коэффициента наклона

$$\hat{\beta} = \frac{\sum_{i=1}^N (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^N (X_i - \bar{X})^2} = \frac{\text{Cov}(X, Y)}{\sigma_X^2}$$

Замечания

- Линия регрессии проходит через точку (\bar{X}, \bar{Y})
- Мы предполагаем, что среди X_i есть разные, тогда $\sigma_X \neq 0$. В противном случае, оценок по методу наименьших квадратов не существует.



Теснота линейной корреляционной связи

В качестве меры близости данных наблюдений к линии регрессии служит выборочный коэффициент парной линейной корреляции (парный линейный коэффициент корреляции):

$$r_{xy} = \frac{\frac{\sum_{i=1}^N X_i Y_i}{N} - \bar{X}\bar{Y}}{\sqrt{\frac{\sum_{i=1}^N X_i^2}{N} - (\bar{X})^2} \sqrt{\frac{\sum_{i=1}^N Y_i^2}{N} - (\bar{Y})^2}}$$

Связь между коэффициентом корреляции и коэффициентом наклона

$$\beta = r_{XY} \frac{\sigma_Y}{\sigma_X} \quad r_{XY} = \beta \frac{\sigma_X}{\sigma_Y}$$

Знак коэффициента наклона линии регрессии и коэффициента корреляции совпадают

Свойства коэффициента корреляции

$$|r_{xy}| \leq 1$$

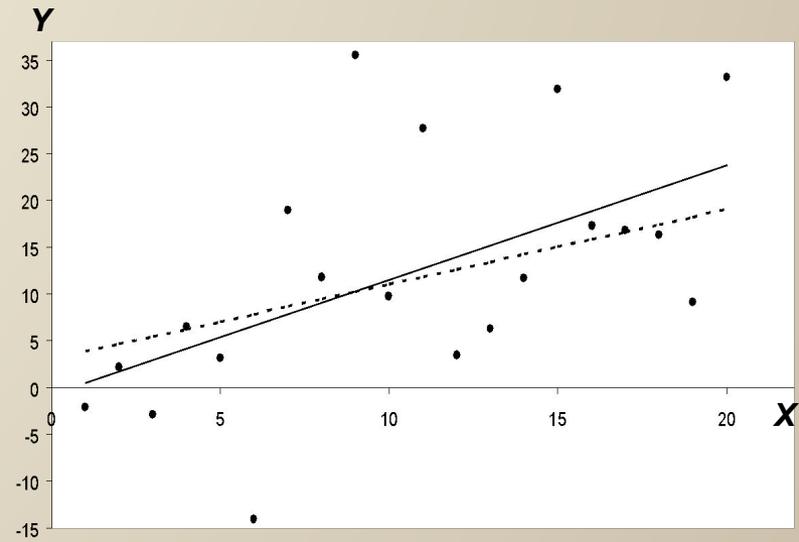
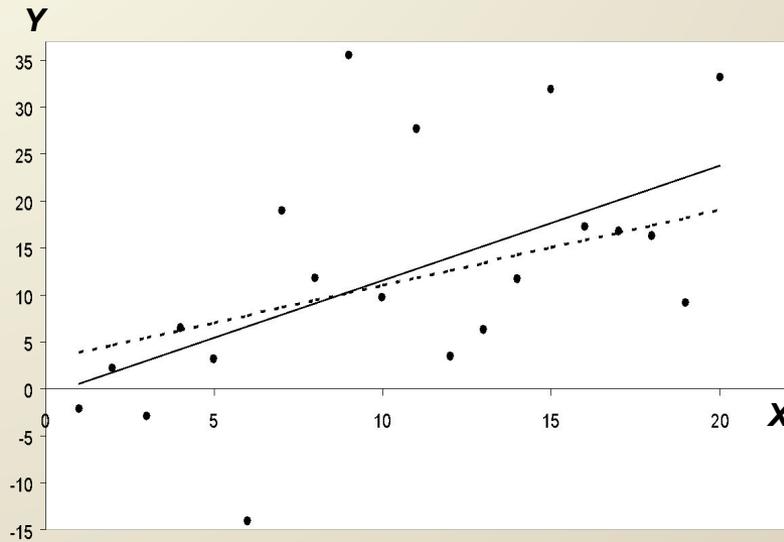
$r_{xy} = \pm 1$ - необходимое и достаточное условием того, что все наблюдаемые значения (X_i, Y_i) лежат на прямой регрессии

Свойства коэффициента корреляции (продолжение)

$r_{XY} = 0$ переменные не связаны линейной корреляционной связью. Линия регрессии проходит горизонтально.

$0 < |r_{xy}| < 1$ между переменными существует линейная корреляционная связь, которая тем лучше (ближе к линейной функциональной), чем ближе коэффициент корреляции по модулю к 1

Уравнение одно, коэффициенты корреляции разные



$$Y = 3.0 + 0.8X$$