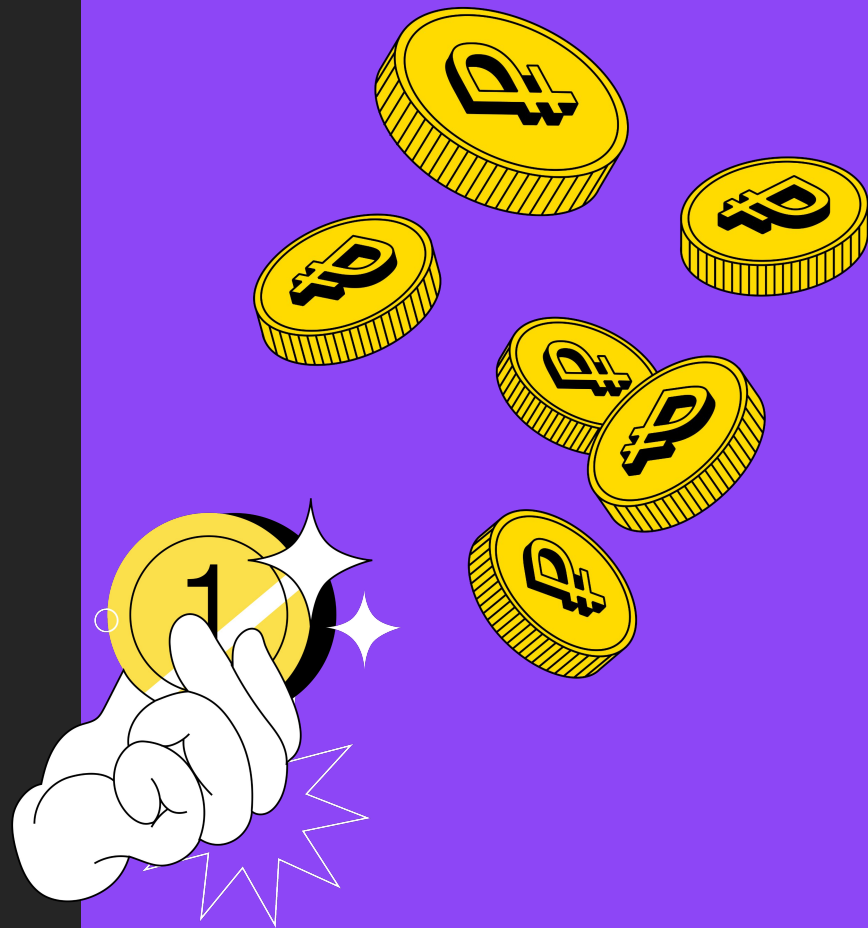


Введение в эконометрию

Урок 5





Знакомство и содержание урока



Дмитрий Бородин, CFA, FRM, Ph.D.

Инвестиционный директор венчурного фонда

- ✦ **12 лет опыта** в корпоративных финансах, инвестиционной оценке и финансовом моделировании.
- ✦ **14 лет опыта** преподавания математических и финансово-экономических дисциплин.
- ✦ Кандидат экономических наук, доцент **МГТУ им. Н.Э. Баумана**, приглашенный преподаватель РЭУ им. Плеханова.
- ✦ Сертификаты финансового аналитика **CFA** и финансового риск-менеджера **FRM**.



План курса

1

Временная стоимость денег

2

Проведение процентных расчетов

3

Оценка эффективности инвестиций

4







Применение статистики в экономике и финансах

5

Введение в эконометрику



Что будет на уроке сегодня

-  Познакомимся с понятиями линейной корреляции.
-  Изучим применение корреляционного анализа в экономике и финансах.
-  Разберёмся с теорией регрессионного анализа.
-  Изучим применение регрессий в экономике и финансах для прогнозирования.
-  Научимся рассчитывать корреляции и определять регрессионные коэффициенты в Excel.
-  Познакомимся с основами анализа временных рядов.



Эконометрика и её роль в экономике и финансах



Что такое эконометрика?

Эконометрика — наука, изучающая количественные и качественные экономические взаимосвязи с помощью статистических и других математических методов и моделей.

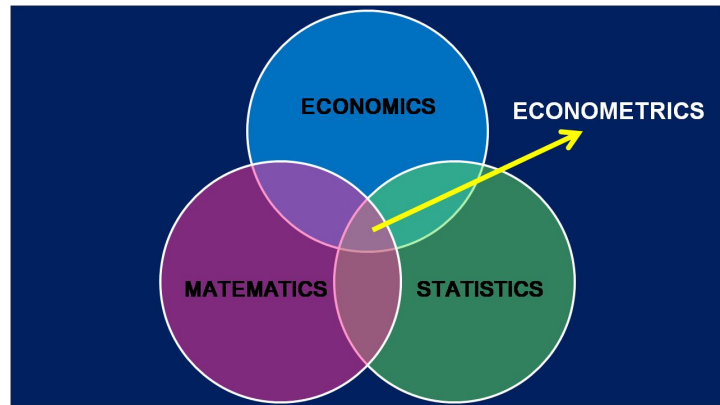
Проще говоря, эконометрика применяет инструментарий современной математики (в том числе математической статистики) к анализу экономических явлений.



Особенности эконометрики

Есть ряд причин, по которым эконометрику выделяют как отдельную науку, несмотря на то, что практически все инструменты и методы взяты из других областей математики:

- сложность интерпретации результатов эконометрических расчётов без чёткого понимания предметной (экономической) области;
- малое количество данных и ограниченный размер выборок;
- нерегулярный характер данных и наличие пропусков/«выбросов»;
- косвенный характер измерений экономических величин;
- невозможность проведения дополнительных экспериментов/наблюдений.





Основные разделы эконометрики

Эконометрику можно разбить на несколько областей, отличающихся **функционалом** и **целями**:

- **корреляционный анализ** предназначен для качественного поиска зависимостей между экономическими случайными величинами (например, между доходностью инвестиционного фонда и доходностью рыночного индекса);
- **регрессионный анализ** предназначен для количественного определения аналитической зависимости между двумя экономическими величинами (например, поиск уравнения зависимости между ценой золота и ценой акции золотодобывающей компании);
- **анализ временных рядов** используется для прогнозирования будущих значений экономической величины на основе её исторических значений (например, прогнозирование будущей волатильности индекса на основе его исторической динамики);
- **панельный анализ** предназначен для одновременного изучения статистических данных и во времени, и между экономическими величинами (например, анализ данных по бедности или преступности)



Корреляционный анализ



Что такое статистическая зависимость?

Статистическая зависимость означает, что распределение вероятностей одной случайной величины зависит от значения другой случайной величины.

Если случайные величины статистически зависимы, то изменения значений одной приводят к систематическому изменению значений другой величины.

Корреляционный анализ — метод обработки статистических данных, с помощью которого измеряется теснота связи между случайными величинами.



Линейная корреляция

При этом функциональная форма статистической зависимости может быть любой. Обычно её классифицируют на **линейную** и **нелинейную** зависимости.

Подавляющее большинство экономических величин связаны именно **линейной** зависимостью, поэтому ей посвящена существенная часть прикладной эконометрики.

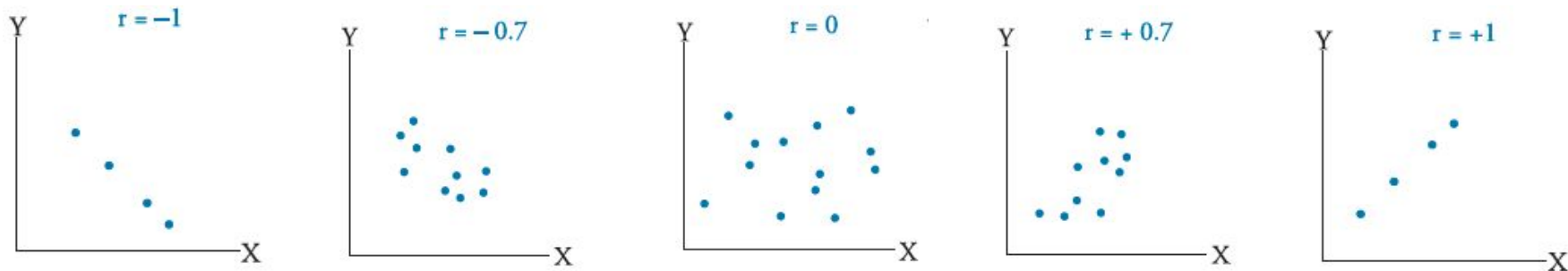
Линейная корреляция — мера линейной связи между двумя случайными величинами. Измеряется коэффициентом линейной корреляции Пирсона, обладающего следующим свойством:

$$-1 \leq r \leq +1$$

- **$r = +1$** — совершенная положительная корреляция,
- **$0 < r < +1$** — несовершенная положительная корреляция,
- **$r = -1$** — совершенная отрицательная корреляция,
- **$-1 < r < 0$** — несовершенная отрицательная корреляция,
- **$r = 0$** — линейная зависимость отсутствует.



Визуализация корреляции

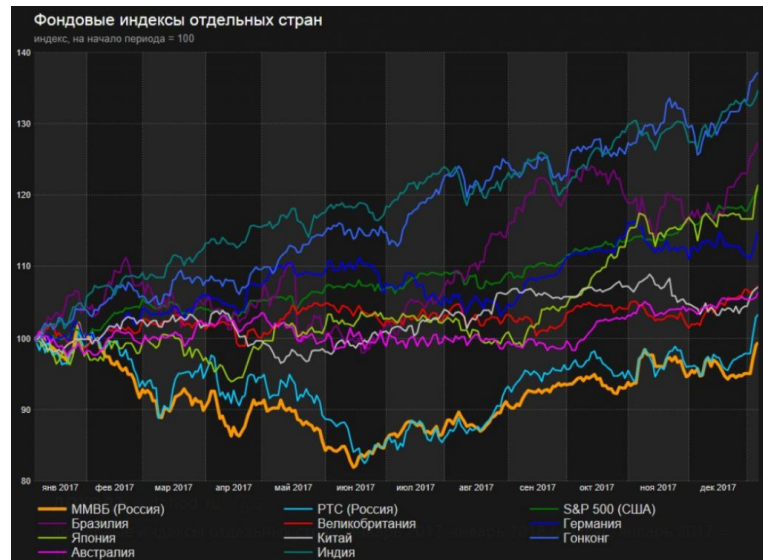


Интерпретация коэффициента линейной корреляции: чем больше значение коэффициента корреляции по модулю (чем он сильнее отличается от 0), тем сильнее линейная взаимосвязь между исследуемыми величинами.



Ограничения корреляционного анализа

- требуется достаточное количество данных;
- измеряется только линейная взаимосвязь, в то время как статистическая взаимосвязь может быть нелинейной, и коэффициент корреляции не сможет её найти;
- не позволяет выделять причинно-следственные связи;
- большое влияние оказывают «выбросы» в данных;
- существует так называемая ложная корреляция.





Регрессионный анализ



Что такое регрессия?

Регрессия — поиск уравнения зависимости одной случайной величины (**зависимой**) от другой (**независимой**), коэффициенты которого подбираются по некоторому критерию, позволяющему лучшим образом объяснить изменение зависимой величины.

В основе регрессии лежат определённые статистические допущения, которые выходят за рамки нашего курса, но их желательно знать при выполнении серьёзных эконометрических расчётов.



Линейная регрессия

Если предполагаемое уравнение зависимости величин имеет линейный вид, то такая регрессия будет называться **линейной**. Формула линейной регрессии:

$$Y_i = b_0 + b_1 X_i + \varepsilon_i, i=1, \dots, n$$

где X — независимая величина (переменная), например, цена золота;

Y — зависимая величина, например, цена акции золотодобывающей компании.

Цель расчёта регрессии — определение коэффициентов b_0 и b_1 . Точного значения этих коэффициентов мы никогда не узнаем, поскольку всегда имеем дело не с генеральными совокупностями, а с ограниченными выборками. Поэтому вместо точных значений коэффициентов мы рассчитываем их оценки, которые используем для записи уравнения:

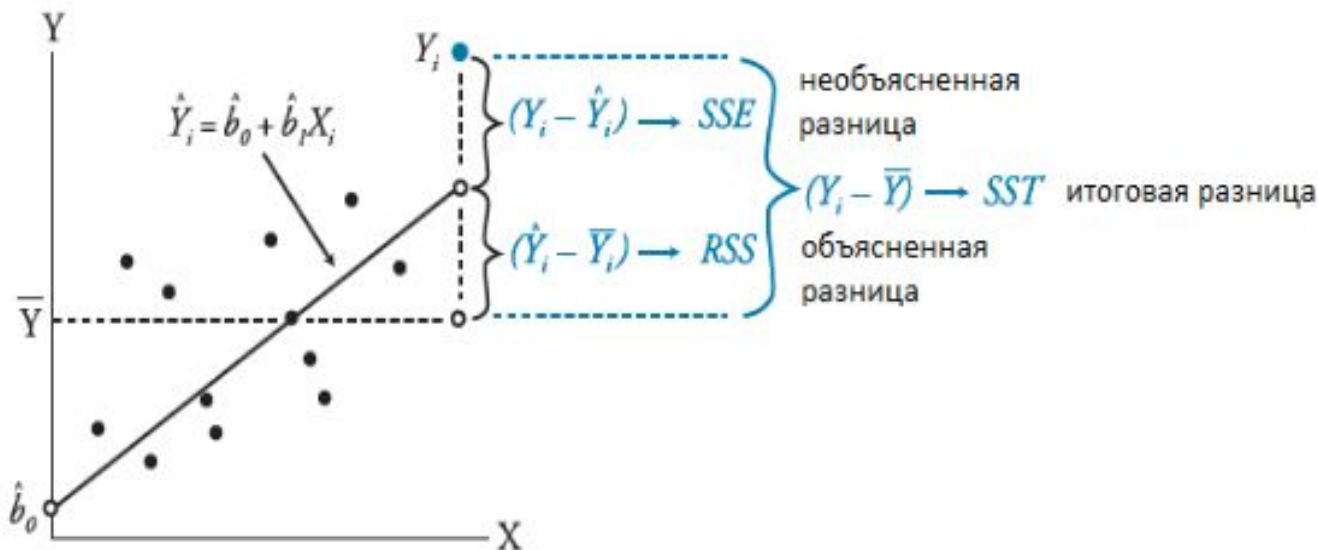
$$\hat{Y}_i = \hat{b}_0 + \hat{b}_1 X_i, i=1, \dots, n$$

Теперь можно прогнозировать значения величины Y в будущем для разных прогнозных значений величины X .



Определение коэффициентов регрессии

Поиск коэффициентов линейной регрессии может осуществляться по-разному, однако в большинстве случаев используется **метод наименьших квадратов**: подбираются такие коэффициенты, чтобы сумма квадратов отклонений реального значения величины Y от её прогнозного значения согласно уравнению регрессии для всех значений величины X была минимальной.





Множественная регрессия

Множественная регрессия — поиск коэффициентов линейной зависимости одной случайной величины Y от нескольких случайных величин X_i , например, зависимость цены продукции предприятия от цен на различные ресурсы, которые используются в производстве.

Формула:

$$Y_i = b_0 + b_1 X_{1i} + b_2 X_{2i} + \dots + b_k X_{ki} + \varepsilon_i$$

Цель — поиск коэффициентов b_1, b_2, \dots, b_k , чтобы получить возможность прогнозировать будущие значения величины Y (аналогично классической линейной регрессии, разобранной выше).

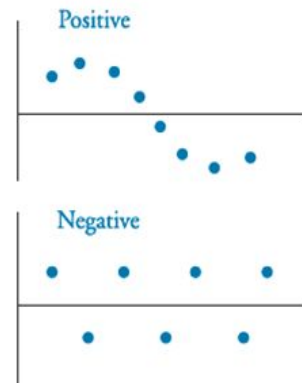
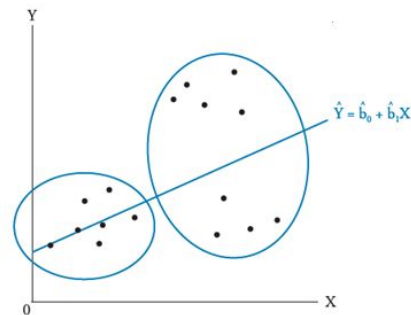


Оценка качества регрессии и проблемы применимости

После расчёта регрессии, прежде чем пользоваться её результатами, необходимо оценить её **качество**. Главный показатель качества регрессии — **коэффициент детерминации (R-квадрат)**. Он показывает, какую долю вариации величины Y при разных значениях X удалось объяснить с помощью регрессии.

Для оценки применимости регрессии к прогнозированию реальных данных следует обратить внимание на потенциальные проблемы:

- **условная гетероскедастичность** — зависимость ошибки регрессии от величины независимой величины (X);
- **автокорреляция** — ошибка регрессии зависит от предыдущих значений ошибки;
- **мультиколлинеарность** — две или более независимых величин (X) оказываются связаны (коррелированы) между собой.





Анализ финансовых временных рядов



Анализ временных рядов

Анализ временных рядов — совокупность математических методов анализа, предназначенных для прогнозирования временного ряда (значений случайной величины во времени).

При этом, в отличие от вышеизложенного, не проводится оценка зависимости от других случайных величин.

В рамках анализа временных рядов выделяют две группы методов:

- трендовые модели;
- авторегрессионные модели.



Реализация методов анализа временных рядов

Трендовые модели изучают зависимость некоторой случайной величины от времени (например, зависимость процентной ставки от времени):

$$y_t = b_0 + b_1(t) + \varepsilon_t$$

В этом случае может использоваться классический регрессионный аппарат, изученный ранее, где в качестве независимой переменной выступает время (номер точки данных).

Авторегрессионные модели изучают зависимость случайной величины от её предыдущих значений.







$$x_t = b_0 + b_1x_{t-1} + b_2x_{t-2} + \dots + b_px_{t-p} + \varepsilon_t$$

В этом случае формулы для стандартной регрессии также применимы, однако существуют более строгие технические требования к характеристикам данных.

Самое частое применение авторегрессионных моделей в финансах — прогнозирование цен и доходностей акций, процентных ставок и волатильности финансовых активов.



На этой лекции мы

-  Познакомились с понятиями линейной корреляции.
-  Изучили применение корреляционного анализа в экономике и финансах.
-  Разобрались с теорией регрессионного анализа.
-  Изучили применение регрессий в экономике и финансах для прогнозирования.
-  Научились рассчитывать корреляции и определять регрессионные коэффициенты в Excel.
-  Познакомились с основами анализа временных рядов.



Спасибо за внимание!