Временные ряды

Временной ряд

Ряд значений одной и той же переменной, полученных в результате измерений, произведенных в последовательные моменты (периоды) времени.

Временной ряд

Две точки зрения на ВР:

- Практическая: последовательность чисел;
- Теоретическая: последовательность случайных величин.

Таким образом ВР обычно интерпретируется как одна из возможных реализаций последовательности случайных величин, представляющей собой случайный (стохастический) процесс с дискретным временем.

Понятие случайного процесса

- Случайная функция, случайный процесс
- Траектория случайного процесса
- Случайный процесс с непрерывным временем
- Случайный процесс с дискретным временем
- Конечномерные распределения сл.процесса
- Числовые характеристики случайного процесса

Временной ряд = результат последовательных наблюдений за случайным процессом, реализация случайного процесса, наблюдённая траектория случайного процесса.

Различие между временным рядом и случайной выборкой

- Случайная выборка – реализации набора *независимых, одинаково* распределенных случайных величин;
- ВР - реализации набора *зависимых* случайных величин с различными (*меняющимися* во времени) распределениями.

=> Случайная выборка – простейший частный случай ВР

Моделирование временных рядов

- **Модель** – это упрощённое описание реальной системы (процесса), отражающее ее важнейшие свойства и предназначенное для ее изучения.
- Модель - это абстрактное представление реальности в какой-либо форме (например, в математической, физической, графической и т.п.), предназначенное для представления определённых аспектов этой реальности и позволяющее получить ответы на изучаемые вопросы.
- Модель: лаконизм $\Longleftrightarrow$ адекватность
- Структурное и неструктурное моделирование (механизм формирования результирующего показателя, предсказание численного значения показателя).

Зачем нужно моделировать временные ряды?

1. Чтобы строить прогнозы

   Каким будет уровень инфляции через год?

2. Чтобы исследовать динамические причинно-следственные связи

   Если фирма сегодня увеличит расходы на рекламу на миллион рублей, то как это скажется на объеме продаж через месяц? А через полгода?

3. Потому что иногда у нас просто нет альтернативных данных

   Исследовать инфляцию России можно только при помощи анализа временного ряда

Обозначения (1)

Y_t — значение переменной Y в период t

Выборка: $Y_1, \ldots Y_T$ — T наблюдений случайной переменной Y_t

Первый лаг переменной Y_t — значение этой переменной в предыдущий период времени, то есть Y_{t-1}

k-ый лаг переменной Y_t — это Y_{t-k}

Обозначения (2)

Первая разность переменной

$$\Delta Y_t = Y_t - Y_{t-1}$$

Темп прироста переменной $\frac{\Delta Y_t}{Y_{t-1}}$

Логарифмическая аппроксимация темпа прироста (верна для малых изменений):

$$\frac{\Delta Y_t}{Y_{t-1}} \approx \Delta ln Y_t = ln Y_t - ln Y_{t-1}$$

Лаговый оператор

По определению: $L\,Y_t = Y_{t-1}$

Примеры записей с использованием лагового оператора:

$$L^2 Y_t = L\big(L(Y_t)\big) = L(Y_{t-1}) = Y_{t-2}$$

$$Y_t = 0{,}4 * Y_{t-1} + u_t \Rightarrow Y_t = 0{,}4 * L\,Y_t + u_t$$
$$\Rightarrow (1 - 0{,}4L)Y_t = u_t$$

Стационарность временного ряда

Интуитивно, «стационарность» означает, что поведение ВР в будущем «похоже» на его поведение в прошлом.

Точнее, основные вероятностные характеристики ВР неизменны во времени

Именно «неизменность поведения» во времени позволяет строить прогнозы стационарных временных рядов на основе их предыстории

Формальные определения

Два типа стационарности:

1. Сильная стационарность (strictly stationarity)

   Стационарность в узком смысле
   «Постоянство законов распределения»

2. Слабая стационарность (weakly stationarity)

   Стационарность в широком смысле

   Ковариационная (2-го порядка)
   «Постоянство первых и вторых моментов распределения (мат.ожиданий, дисперсий, ковариаций)»

Строгая стационарность

Временной ряд называется стационарным в узком смысле (строго стационарным),

если совместное распределение m наблюдений $Y_{t_1}, Y_{t_2} \dots Y_{t_m}$ не зависит от сдвига по времени, то есть совпадает с распределением $Y_{t_1+k}, Y_{t_2+k} \dots Y_{t_m+k}$ для любых m, $t_1, \dots t_m$, k.

Слабая стационарность

Временной ряд называется стационарным в широком смысле (слабо стационарным),

если для всех t:

$$E(Y_t) = \mu$$
$$V(Y_t) = \gamma_0$$
$$Cov(Y_t, Y_{t-k}) = \gamma_k$$

Важно: μ, γ_0 и γ_{k} конечны и не зависят от t

Слабая стационарность

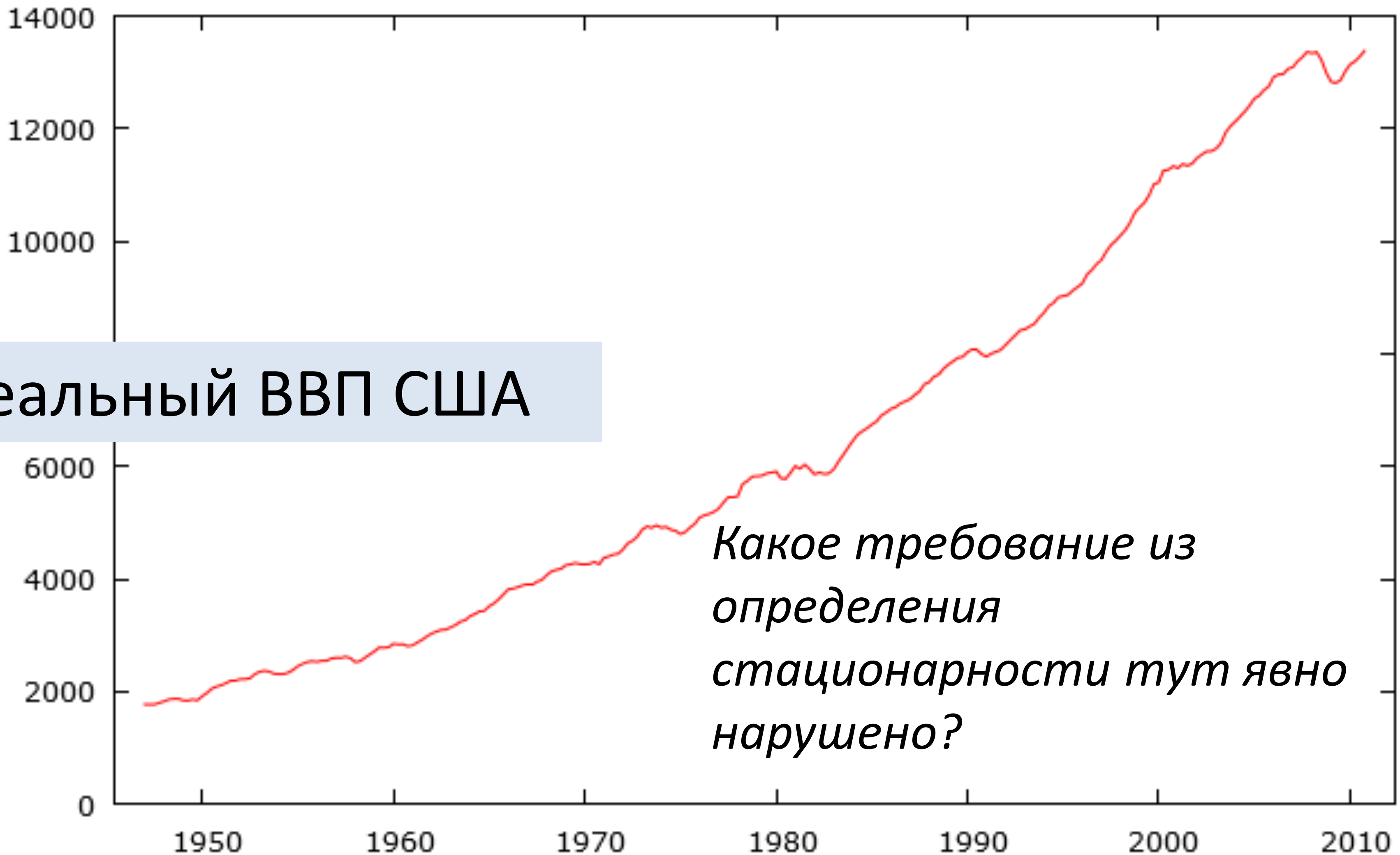
Временной ряд называется стационарным в широком смысле (слабо стационарным),

если для всех t:

$$E(Y_t) = \mu$$
$$V(Y_t) = \gamma_0$$
$$Cov(Y_t, Y_{t-k}) = \gamma_k$$

Следует ли из строгой стационарности слабая?

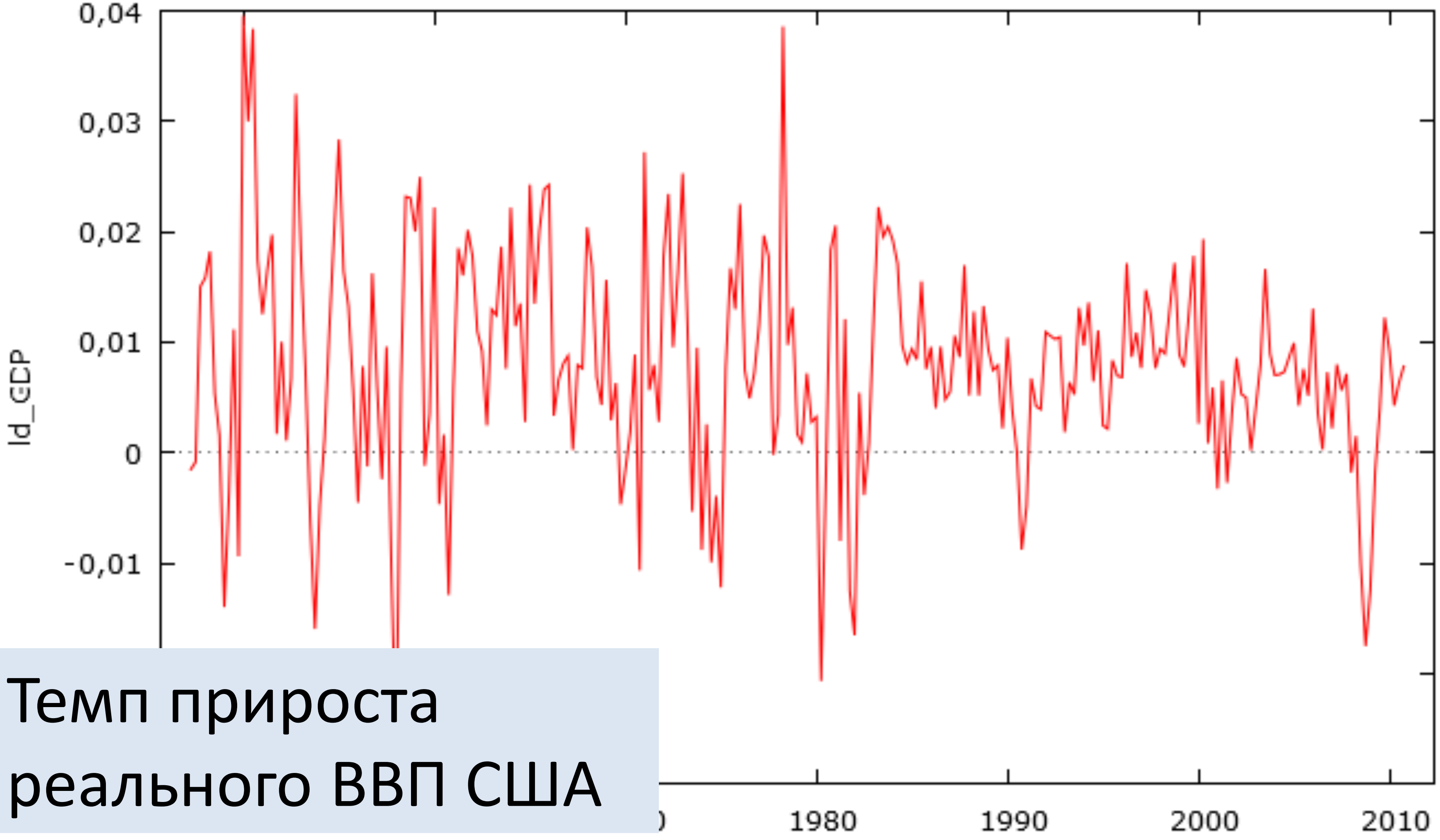
Пример нестационарного ряда



Реальный ВВП США

Какое требование из определения стационарности тут явно нарушено?

Пример стационарного ряда



Темп прироста реального ВВП США

Стационарные ряды: автокорреляционная функция

Коэффициент автоковариации k-го порядка:

$$Cov(Y_t, Y_{t-k}) = \gamma_k$$

Коэффициент автокорреляции k-го порядка:

$$\rho_k = Corr(Y_t, Y_{t-k}) = \frac{Cov(Y_t, Y_{t-k})}{\sqrt{V(Y_t)V(Y_{t-k})}} =$$

$$= \frac{\gamma_k}{\sqrt{\gamma_0\gamma_0}} = \frac{\gamma_k}{\gamma_0}$$

Стационарные ряды: автокорреляционная функция

Коэффициент автокорреляции k-го порядка:

$$\rho_k = Corr(Y_t, Y_{t-k}) = \frac{\gamma_k}{\gamma_0}$$

ρ_k как функция от k называется автокорреляционной функцией (АКФ, ACF).

График АКФ — это наглядный инструмент для анализа свойств временного ряда

Одномерные и многомерные модели временных рядов (ВР)

Мы начнем с одномерных моделей ВР, в которых текущее значение ВР зависит только от его собственной предыстории (и не зависит от других ВР)

Такие модели часто используются при построении прогнозов.

Для анализа причинно-следственных связей используются многомерные модели ВР, которые будут обсуждаться позже.

Белый шум (White noise)

Процессом белого шума называется последовательность независимых и одинаково распределенных случайных величин с нулевым математическим ожиданием и конечной дисперсией σ^2:

$$\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{t-1}, \varepsilon_t, \dots$$

Более слабый вариант белого шума:

$$E\varepsilon_t = 0\ ,\ \ D\varepsilon_t = \sigma^2\ ,\ \ \ cov(\varepsilon_t, \varepsilon_s) = 0$$

Белый шум (White noise)

Процессом белого шума называется последовательность независимых и одинаково распределенных случайных величин с нулевым математическим ожиданием и конечной дисперсией σ^2:

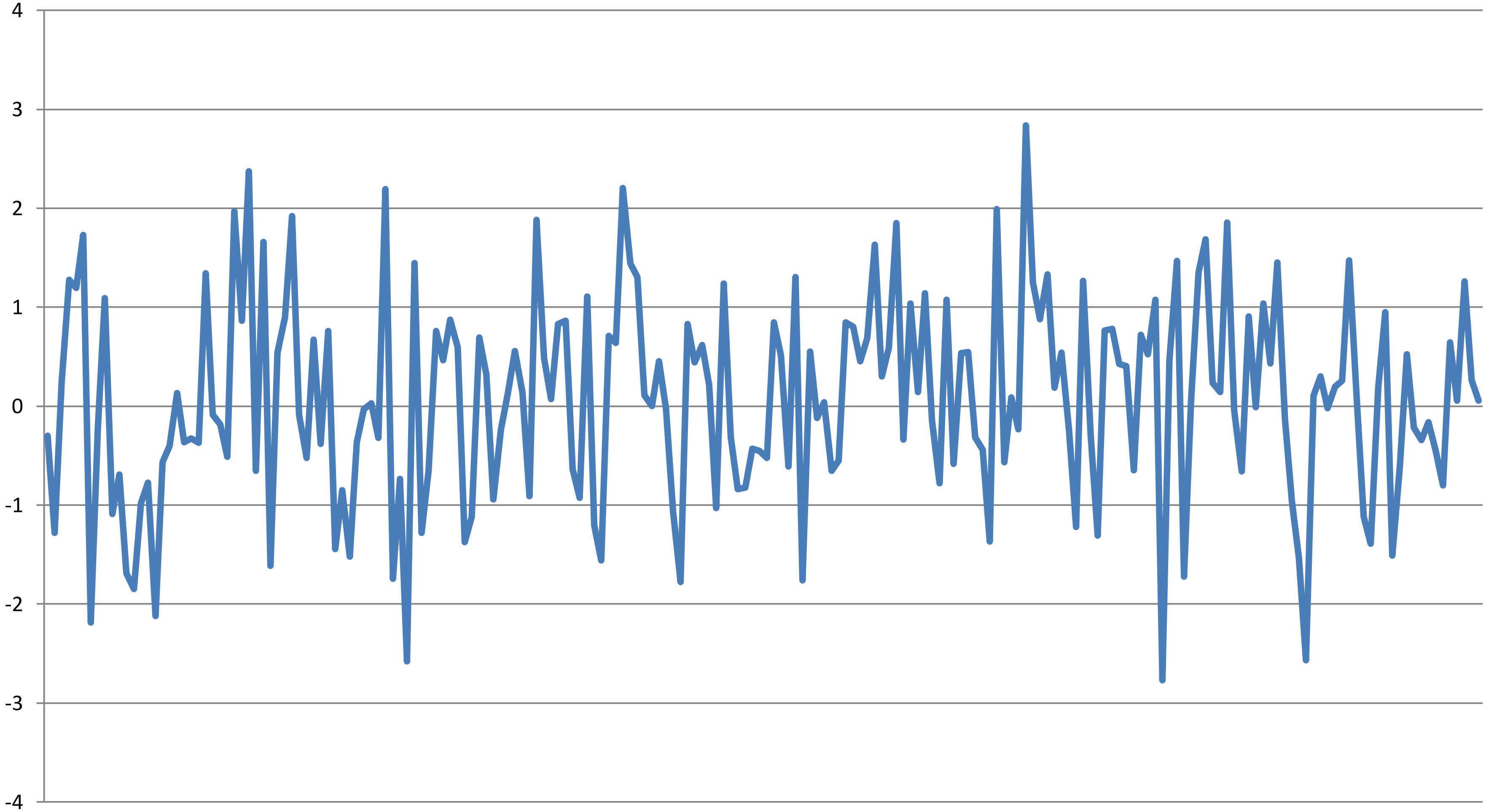
$$\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{t-1}, \varepsilon_t, \dots$$

Более слабый вариант белого шума:

$$E\varepsilon_t = 0\,,\ \ D\varepsilon_t = \sigma^2\,,\ \ cov(\varepsilon_t, \varepsilon_s) = 0$$

Вопрос: Белый шум стационарен ?

Белый шум



Модель авторегрессии AR(1)

$$y_t = \delta + \theta y_{t-1} + \varepsilon_t,$$

где δ и θ – коэффициенты, $|\theta| < 1$,

ε_t- белый шум

y_t – процесс авторегрессии первого порядка

Представление процесса AR(1) при $|\theta| < 1$

в виде линейной комбинации текущего и прошлых значений белого шума.

Условие стационарности процесса AR(1)

AR(1) — процесс авторегрессии 1-го порядка

$$y_t = \delta + \theta y_{t-1} + \varepsilon_t, |\theta| < 1$$

Докажем, что:

$$\mu = \frac{\delta}{1-\theta},$$

$$\gamma_0 = \frac{\sigma^2}{1-\theta^2},$$

$$\gamma_k = \theta^k \frac{\sigma^2}{1-\theta^2}$$

AR(1) — вычисление математического ожидания

$$y_t = \delta + \theta y_{t-1} + \varepsilon_t, |\theta| < 1$$

$$Ey_t = E(\delta + \theta y_{t-1} + \varepsilon_t)$$

$$Ey_t = \delta + \theta E(y_{t-1}) + E(\varepsilon_t)$$

$$\mu = \delta + \theta\mu + 0$$

$$\mu = \frac{\delta}{1 - \theta}$$

AR(1) — вычисление дисперсии

$$y_t = \delta + \theta y_{t-1} + \varepsilon_t, |\theta| < 1$$

$$V(y_t) = V(\delta + \theta y_{t-1} + \varepsilon_t)$$

$$V(y_t) = \theta^2 V(y_{t-1}) + V(\varepsilon_t)$$

$$\gamma_0 = \theta^2 \gamma_0 + \sigma^2$$

$$\gamma_0 = \frac{\sigma^2}{1 - \theta^2}$$

AR(1) — вычисление автоковариаций

$$\gamma_1 = Cov(y_t, y_{t-1}) = Cov(\delta + \theta y_{t-1} + \varepsilon_t, y_{t-1}) =$$

$$= Cov(\theta y_{t-1} + \varepsilon_t, y_{t-1}) =$$

$$= \theta * Cov(y_{t-1}, y_{t-1}) + Cov(\varepsilon_t, y_{t-1}) =$$

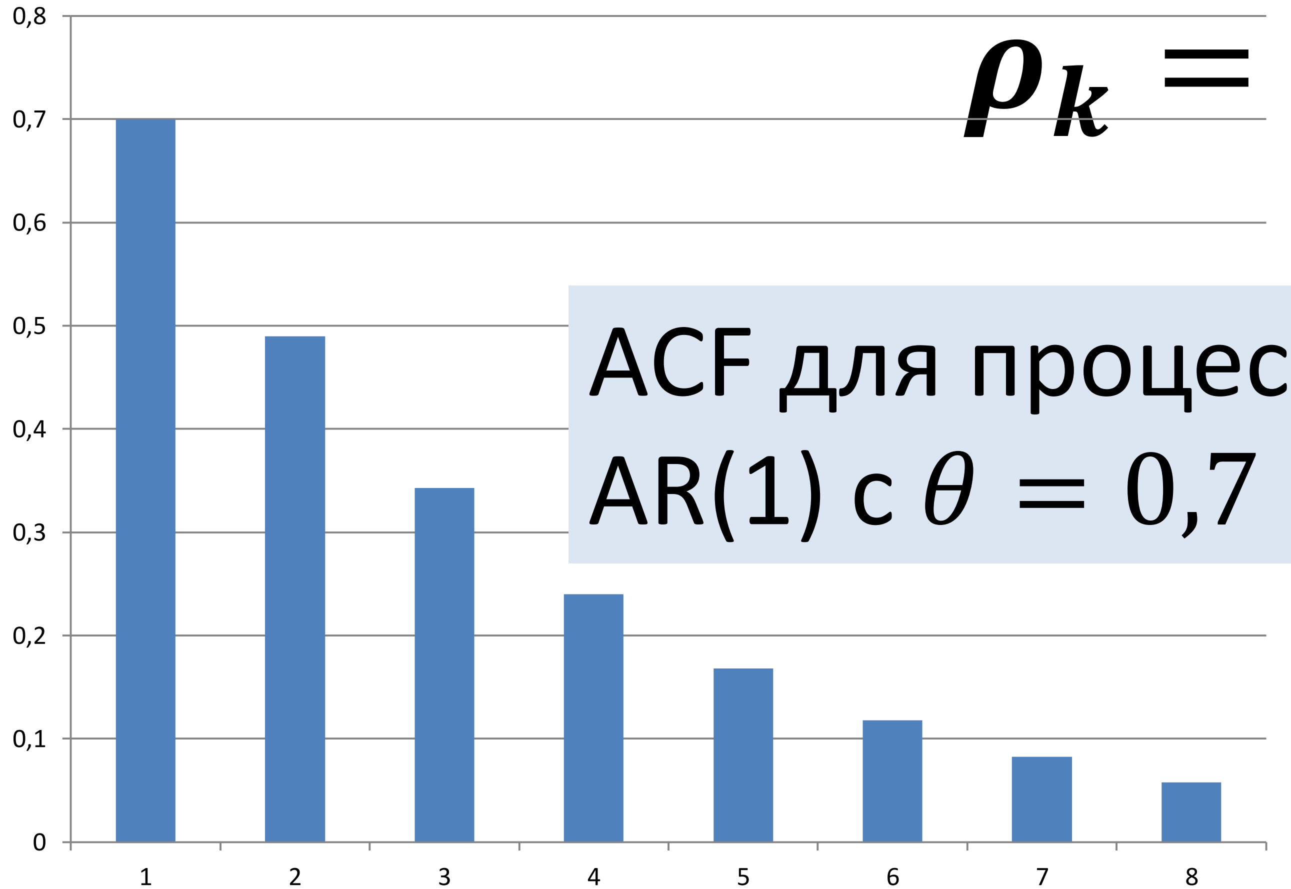
$$= \theta * \gamma_0 + 0 = \theta \frac{\sigma^2}{1 - \theta^2}$$

Аналогично: $\gamma_k = Cov(y_t, y_{t-k}) = \theta^k \frac{\sigma^2}{1-\theta^2}$

AR(1) — процесс авторегрессии 1-го порядка

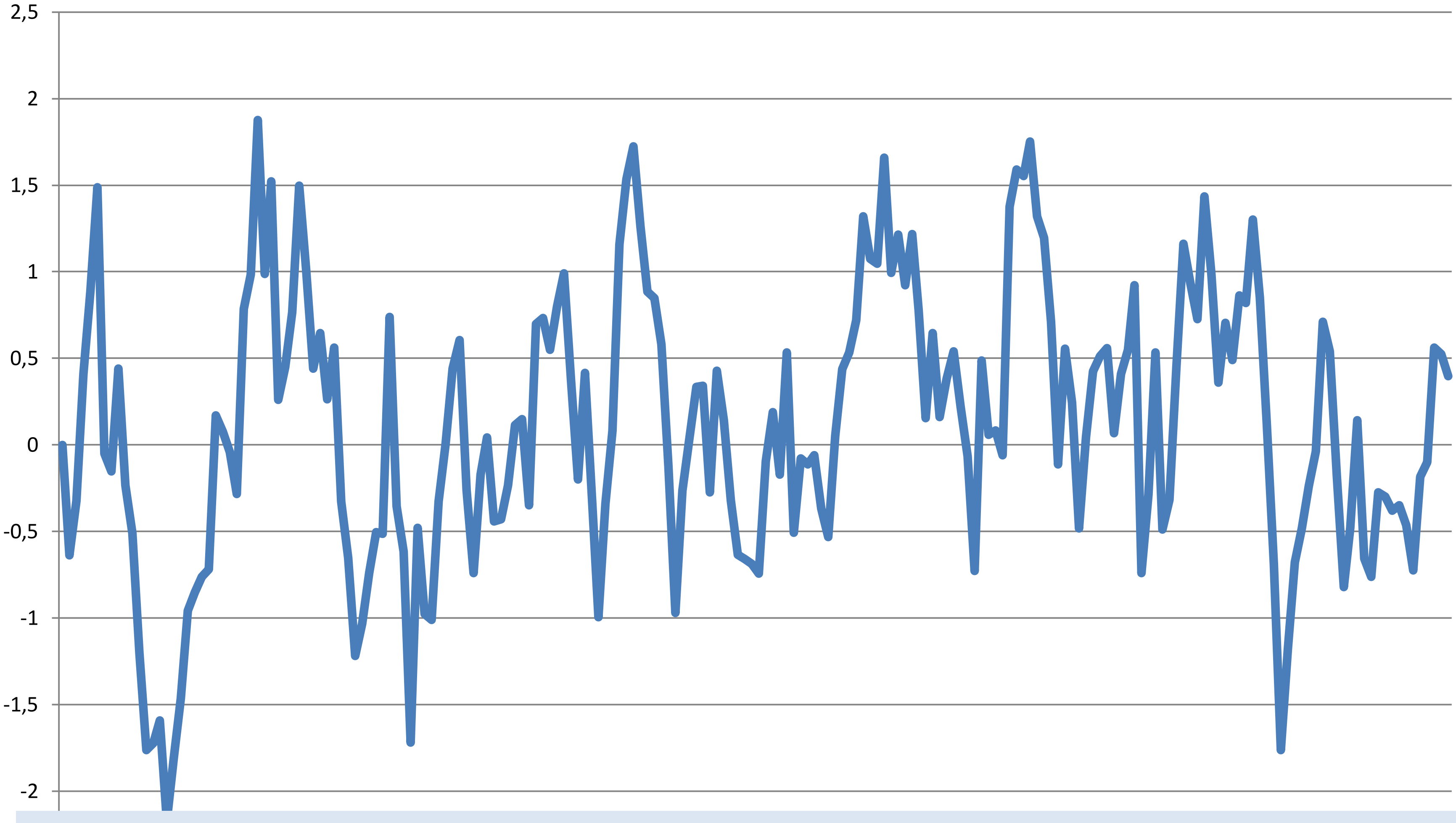
$$y_t = \delta + \theta y_{t-1} + \varepsilon_t, |\theta| < 1$$

$$\boldsymbol{\rho_k = \frac{\gamma_k}{\gamma_0} = \theta^k}$$



ACF для процесса AR(1) с $\theta = 0{,}7$

AR(1) — процесс авторегрессии 1-го порядка



Пример временного ряда порожденного AR(1) процессом с $\theta = 0{,}7$

AR(p) — процесс авторегрессии порядка p

$$y_t = \delta + \theta_1 y_{t-1} + \cdots + \theta_p y_{t-p} + \varepsilon_t$$

AR(p) — процесс авторегрессии порядка p

$$y_t = \delta + \theta_1 y_{t-1} + \cdots + \theta_p y_{t-p} + \varepsilon_t$$

Запись модели с использованием лагового оператора:

$$y_t = \delta + \theta_1 L y_t + \cdots + \theta_p L^p y_t + \varepsilon_t$$

или

$$(1 - \theta_1 L - \cdots - \theta_p L^p) y_t = \delta + \varepsilon_t$$

AR(p) — процесс авторегрессии порядка p

$$y_t = \delta + \theta_1 y_{t-1} + \cdots + \theta_p y_{t-p} + \varepsilon_t$$

$$(1 - \theta_1 L - \cdots - \theta_p L^p) y_t = \delta + \varepsilon_t$$

Характеристическое уравнение:

$$1 - \theta_1 z - \cdots - \theta_p z^p = 0$$

AR(p) — процесс авторегрессии порядка p

$$y_t = \delta + \theta_1 y_{t-1} + \cdots + \theta_p y_{t-p} + \varepsilon_t$$

$$(1 - \theta_1 L - \cdots - \theta_p L^p) y_t = \varepsilon_t$$

Характеристическое уравнение:

$$1 - \theta_1 z - \cdots - \theta_p z^p = 0$$

Известно, что у многочлена степени p существуют p штук комплексных корней

Необходимое и достаточное условие стационарности процесса AR(p)

Теорема:

Процесс AR(p) является стационарным тогда и только тогда, когда все корни характеристического уравнения

$$1 - \theta_1 z - \cdots - \theta_p z^p = 0$$

по модулю больше единицы $|z_j| > 1$

Необходимое и достаточное условие стационарности процесса AR(p)

Любой многочлен степени p имеет ровно p штук комплексных корней, поэтому
$1-\theta_1 z-\cdots-\theta_p z^p=-\theta_p(z-z_1)(z-z_2)\ldots(z-z_p)$

С использованием *инвертированных* корней $\lambda_j=1/z_j$ (ясно, что все корни z_j ненулевые!), условие стационарности формулируется так:
$|\lambda_j|<1$ для всех j=1,…p.

Характеристический многочлен имеет вид
$(1-\lambda_1 z)(1-\lambda_2 z)\ldots(1-\lambda_p z)$,

Необходимое и достаточное условие стационарности процесса AR(p)

Характеристический многочлен теперь имеет вид

$$(1-\lambda_1 z)(1-\lambda_2 z)\ldots\left(1-\lambda_p z\right),$$

а характеристическое уравнение

$$(1-\lambda_1 L)(1-\lambda_2 L)\ldots\left(1-\lambda_p L\right)y_p = \delta + \varepsilon_t$$

При $\left|\lambda_j\right|<1$

MA(1) — процесс скользящего среднего 1-го порядка

$$y_t = \delta + \varepsilon_t + \alpha\varepsilon_{t-1}$$

Докажем, что:

$$\mu = \delta,$$

$$\gamma_0 = (1 + \alpha^2)\sigma^2,$$

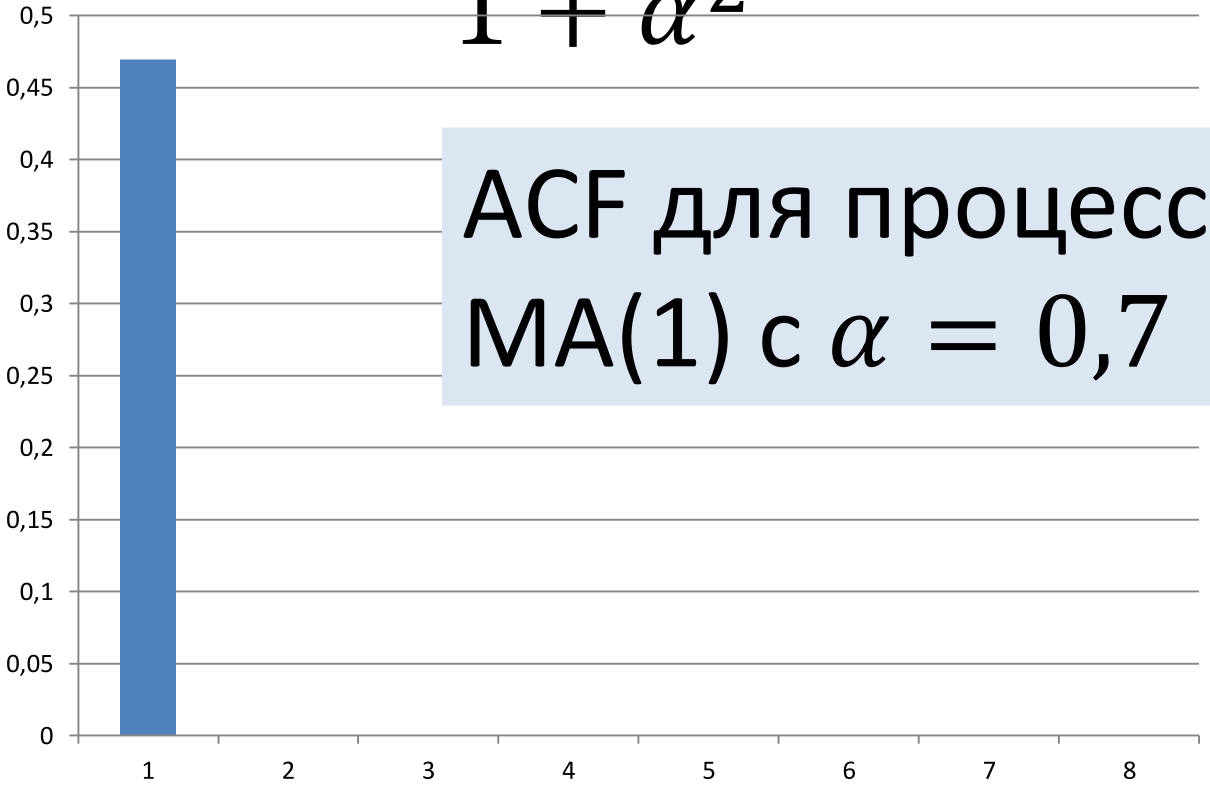
$$\gamma_1 = \alpha\sigma^2,$$

$$\gamma_2 = \gamma_3 = \cdots = 0$$

MA(1) — процесс скользящего среднего 1-го порядка

$$y_t = \delta + \varepsilon_t + \alpha\varepsilon_{t-1}$$

$$\rho_1 = \frac{\alpha}{1+\alpha^2}, \rho_2 = \rho_3 = \cdots = 0$$



ACF для процесса MA(1) с $\alpha = 0{,}7$

MA(q) — процесс скользящего среднего порядка q

$$y_t = \delta + \varepsilon_t + \alpha_1 \varepsilon_{t-1} + \cdots + \alpha_q \varepsilon_{t-q}$$

MA(q) — процесс скользящего среднего порядка q

$$y_t = \delta + \varepsilon_t + \alpha_1 \varepsilon_{t-1} + \cdots + \alpha_q \varepsilon_{t-q}$$

Запись с использованием лагового оператора:

$$y_t = \delta + \varepsilon_t + \alpha_1 L \varepsilon_t + \cdots + \alpha_q L^q \varepsilon_t$$

$$y_t = \delta + (1 + \alpha_1 L + \cdots + \alpha_q L^q) \varepsilon_t$$

MA(q) — процесс скользящего среднего порядка q

$$y_t = \delta + \varepsilon_t + \alpha_1 \varepsilon_{t-1} + \cdots + \alpha_q \varepsilon_{t-q}$$

Процесс стационарен при любых значениях $\alpha_1, \ldots \alpha_q$.

Почему?

Процесс ARMA(p, q)

Естественное обобщение подхода к моделированию: объединение AR и MA процессов в одну модель ARMA

$$y_t = \delta + \theta_1 y_{t-1} + \cdots + \theta_p y_{t-p} +$$
$$+\varepsilon_t + \alpha_1 \varepsilon_{t-1} + \cdots + \alpha_q \varepsilon_{t-q}$$

Разложение Вольда

Любой **стационарный** процесс может быть представлен в виде бесконечного ряда членов белого шума с коэффициентами, образующими абсолютно сходящийся числовой ряд.

Это означает, например, что стационарный AR(1) процесс можно представить как процесс MA(∞)

– *Покажите это.*

Таким образом нет фундаментальной разницы между AR и MA представлением ВР.

– Выбор между AR, MA или ARMA — вопрос удобства и лаконичности представления.

Общие корни

Если AR и MA компоненты ARMA модели имеют общие корни, то представление ряда можно упростить. Например:

$\text{ARMA}(2,1)\ y_t = 0{,}3y_{t-1} - 0{,}02y_{t-2} + \varepsilon_t - 0{,}2\varepsilon_{t-1}$

Общие корни

Если AR и MA компоненты ARMA модели имеют общие корни, то представление ряда можно упростить. Например:

$\mathrm{ARMA}(2{,}1)\ y_t = 0{,}3y_{t-1} - 0{,}02y_{t-2} + \varepsilon_t - 0{,}2\varepsilon_{t-1}$

$$y_t - 0{,}3y_{t-1} + 0{,}02y_{t-2} = \varepsilon_t - 0{,}2\varepsilon_{t-1}$$

$$(1 - 0{,}3L + 0{,}02L^2)y_t = (1 - 0{,}2L)\varepsilon_t$$

$$(1 - 0{,}1L)(1 - 0{,}2L)y_t = (1 - 0{,}2L)\varepsilon_t$$

Общие корни

Если AR и MA компоненты ARMA модели имеют общие корни, то представление ряда можно упростить. Например:

$\mathrm{ARMA}(2,1)\ y_t = 0{,}3y_{t-1} - 0{,}02y_{t-2} + \varepsilon_t - 0{,}2\varepsilon_{t-1}$

$$y_t - 0{,}3y_{t-1} + 0{,}02y_{t-2} = \varepsilon_t - 0{,}2\varepsilon_{t-1}$$

$$(1 - 0{,}3L + 0{,}02L^2)y_t = (1 - 0{,}2L)\varepsilon_t$$

$$(1 - 0{,}1L)(1 - 0{,}2L)y_t = (1 - 0{,}2L)\varepsilon_t$$

$$(1 - 0{,}1L)y_t = \varepsilon_t$$

$$y_t - 0{,}1y_{t-1} = \varepsilon_t$$

$$y_t = 0{,}1y_{t-1} + \varepsilon_t \Rightarrow \mathrm{AR}(1)$$

Случайное блуждание (Random walk, RW)

$$y_t = y_{t-1} + \varepsilon_t$$

Стационарен ли этот процесс?

$$(1 - L)y_t = \varepsilon_t$$

Характеристическое уравнение:

$$1 - z \ = 0$$

$z = 1$. **Единичный корень**. Процесс не стационарен

Случайное блуждание (Random walk, RW)

$$y_t = y_{t-1} + \varepsilon_t$$

Покажите, что

$$E(y_t) = y_0 \text{ (если } y_0 = const)$$

$$Var(y_t) = t * \sigma^2$$

Случайное блуждание (Random walk, RW)

Пример RW: Обменный курс доллара США к немецкой марке (ежедневные данные, 2 января 1980 — 21 мая 1987)

Случайное блуждание (Random walk, RW)

Более общий случай — случайное блуждание с дрейфом:

$$y_t = \delta + y_{t-1} + \varepsilon_t$$

Покажите, что

$$E(y_t) = y_0 + t * \delta$$

$$Var(y_t) = t * \sigma^2$$

Случайное блуждание (Random walk, RW)

$$y_t = \delta + y_{t-1} + \varepsilon_t$$

Как устроена первая разность для этого процесса?

Она стационарна:

$$y_t - y_{t-1} = \delta + \varepsilon_t$$
$$\Delta y_t = \delta + \varepsilon_t$$

Случайное блуждание (Random walk, RW)

Процесс случайного блуждания не стационарен, однако его первая разность является стационарным процессом.

На практике многие финансовые или макроэкономические переменные также не стационарны сами по себе но стационарны в разностях.

Это приводит нас к полезному определению…

Порядок интегрированности ВР

- Если процесс не стационарен, а его первые разности стационарны, то он называется интегрированным процессом 1-го порядка, $I(1)$
- Если процесс y_t не стационарен, его первые разности Δy_t тоже не стационарны, но его вторые разности
$$\Delta^2 y_t = (\Delta y_t - \Delta y_{t-1})$$
стационарны, то процесс называется интегрированным 2-го порядка, $I(2)$

Порядок интегрированности ВР

- Аналогично определяют интегрированный процесс k-го порядка, $I(k)$
- На практике почти не встречается процессов с порядком интегрированности выше второго.
- Стационарный процесс часто обозначают $I(0)$ и называют интегрированным процессом нулевого порядка.

Процесс ARIMA(p, k, q)

Если

- процесс является интегрированным k-го порядка,
- а его разность k-го порядка описывается процессом ARMA(p,q),

то исходный процесс называется интегрированным процессом авторегрессии со скользящим средним в остатках ARIMA(p,k,q)

Процесс ARIMA(p, k, q)

Пример: $\boldsymbol{y_t = 0,9y_{t-1} + 0,1y_{t-2} + \varepsilon_t}$

p = ?, k = ?, q = ?

$$y_t - 0{,}9y_{t-1} - 0{,}1y_{t-2} = \varepsilon_t$$
$$(1 - 0{,}9L - 0{,}1L^2)y_t = \varepsilon_t$$
$$(1 - L)(1 + 0{,}1L)y_t = \varepsilon_t$$

Процесс содержит единичный корень, то есть нестационарен.

Рассмотрим первые разности процесса

Процесс ARIMA(p, k, q)

Пример: $\boldsymbol{y_t = 0,9y_{t-1} + 0,1y_{t-2} + \varepsilon_t}$

Рассмотрим первые разности процесса:

$$y_t - y_{t-1} = 0{,}9y_{t-1} - y_{t-1} + 0{,}1y_{t-2} + \varepsilon_t$$

$$\Delta y_t = -0{,}1\Delta y_{t-1} + \varepsilon_t$$

Первые разности описываются стационарным процессом AR(1) или, что тоже самое ARMA(1,0).

Значит исходный процесс описывается моделью ARIMA(1,1,0)

Прогнозирование в рамках модели ARIMA(p, k, q)

Главная цель построения ARIMA моделей — прогнозирование будущих значений экономических переменных.

- Хорошая новость: ARIMA модели при своей относительной простоте часто справляются с задачей прогнозирования лучше, чем сложные структурные модели.
- Плохая новость: При этом у ARIMA моделей может не быть содержательной экономической интерпретации.

Прогнозирование в рамках модели ARIMA(p, k, q)

- Предположим, что мы находимся в моменте времени T. Нам доступна информация о y_T, y_{T-1}, y_{T-2} … (эти величины для нас уже не являются случайными)
- Нас интересует предсказание y_{T+h}, то есть предсказание на h шагов вперед.

Прогнозирование в рамках модели ARIMA(p, k, q)

- Естественно использовать в качестве прогноза случайной величины y_{t+h} ее условное математическое ожидание при условии доступной нам информации:

$$\widehat{y_{T+h}} = E(y_{T+h} \ \ |y_T, y_{T-1}, y_{T-2} \dots)$$

- Такой прогноз является наилучшим с точки зрения минимизации средней квадратической ошибки прогноза

Пример: прогнозирование для процесса AR(1)

$$y_t = \theta y_{t-1} + \varepsilon_t$$

Известно, что в момент T моделируемая переменная приняла значение y_T.

Прогноз на 1 шаг вперед:

$$\widehat{y_{T+1}} = E(y_{T+1}|y_T) =$$

$$= E(\theta y_T + \varepsilon_{T+1}|y_T) =$$

$$= \theta y_T + E(\varepsilon_{T+1}|y_T) = \theta y_T$$

Пример: прогнозирование для процесса AR(1)

$$y_t = \theta y_{t-1} + \varepsilon_t$$

Известно, что в момент T моделируемая переменная приняла значение y_T.

Прогноз на 2 шага вперед:

$$\widehat{y_{T+2}} = E(y_{T+2}|y_T) =$$
$$= E(\theta y_{T+1} + \varepsilon_{T+2}|y_T) =$$
$$= \theta E(y_{T+1}|y_T) + E(\varepsilon_{T+2}|y_T) =$$
$$= \theta \widehat{y_{T+1}} + 0 = \theta^2 y_T$$

Пример: прогнозирование для процесса AR(1)

$$y_t = \theta y_{t-1} + \varepsilon_t$$

Известно, что в момент T моделируемая переменная приняла значение y_T.

Прогнозы на 1, 2 и h шагов вперед:

$$\widehat{y_{T+1}} = \theta y_T, \qquad \widehat{y_{T+2}} = \theta^2 y_T,$$

$$\widehat{y_{T+h}} = \theta^h y_T$$

Точность прогноза

Ошибка прогноза: $y_{T+h} - \widehat{y_{T+h}}$

Выбранный нами способ построения прогноза минимизирует математическое ожидание квадрата ошибки прогноза (дисперсии ошибки прогноза):

$$C_h$$
$$= E(y_{T+h} - \widehat{y_{T+h}})^2 | y_T, y_{T-1}, y_{T-2}, \dots)$$
$$== V(y_{T+h} \, | y_T, y_{T-1}, y_{T-2}, \dots)$$

Пример: прогнозирование для процесса AR(1)

$$y_t = \theta y_{t-1} + \varepsilon_t$$

Известно, что в момент T моделируемая переменная приняла значение y_T.

Покажите, что дисперсии ошибок прогноза:

$$C_1 = \sigma^2, \qquad C_2 = (1 + \theta^2)\sigma^2,$$

$$C_h = (1 + \theta^2 + \dots + \theta^{2(h-1)})\sigma^2$$

Точность прогноза

- Если предположить, что ε_t — не просто белый шум, а имеет нормальное распределение, то можно построить доверительный интервал для прогноза.
- Например, 95-процентный доверительный интервал для прогноза на h шагов вперед будет иметь вид:

$$\left(\widehat{y_{T+h}} - 1{,}96\sqrt{C_h}, \qquad \widehat{y_{T+h}} + 1{,}96\sqrt{C_h}\right)$$

Как меняется ширина доверительного интервала с увеличением горизонта прогнозирования?

Точность прогноза

- Мы рассматривали построение прогноза в условиях, когда параметры ARIMA модели **точно известны**
- В этом случае есть единственный источник неточности прогноза — воздействие случайных ошибок
- На практике параметры моделей ARIMA **не известны** и их приходится оценивать, это создает дополнительный источник неточности прогноза $\Rightarrow$ на практике доверительные интервалы прогнозов будут шире
- Асимптотически дисперсии оценок коэффициентов сходятся к нулю и второй источник неточности исчезает

Стационарность временных рядов

Стационарен ли временной ряд?

Для принятия решения полезно:

1. Смотреть на график временного ряда
2. Использовать формальные статистические тесты

Тестирование стационарности для AR(1)

$$y_t = \theta y_{t-1} + \varepsilon_t$$

$H_0: \theta = 1$ — ряд является нестационарным

- содержит единичный корень,
- описывается процессом случайного блуждания

$H_1: |\theta| < 1$ — ряд является стационарным

- не содержит единичный корень,
- описывается стационарным авторегрессионным процессом первого порядка

Тестирование стационарности для AR(1)

$$y_t = \theta y_{t-1} + \varepsilon_t$$

$$y_t - y_{t-1} = \theta y_{t-1} - y_{t-1} + \varepsilon_t$$

Тестирование стационарности для AR(1)

$$y_t = \theta y_{t-1} + \varepsilon_t$$

$$y_t - y_{t-1} = \theta y_{t-1} - y_{t-1} + \varepsilon_t$$

$$\Delta y_t = (\theta - 1) * y_{t-1} + \varepsilon_t$$

Обозначим $\theta - 1 = b$. $\Delta \boldsymbol{y_t} = \boldsymbol{b} * \boldsymbol{y_{t-1}} + \boldsymbol{\varepsilon_t}$

Тестирование стационарности для AR(1)

$$y_t = \theta y_{t-1} + \varepsilon_t$$

$$y_t - y_{t-1} = \theta y_{t-1} - y_{t-1} + \varepsilon_t$$

$$\Delta y_t = (\theta - 1) * y_{t-1} + \varepsilon_t$$

Обозначим $\theta - 1 = b$. $\Delta \boldsymbol{y_t} = \boldsymbol{b} * \boldsymbol{y_{t-1}} + \boldsymbol{\varepsilon_t}$

В этом случае:

H_0: $\theta = 1 \Rightarrow b = 0$. Если ряд содержит единичный корень, то коэффициент $\boldsymbol{b}$ должен быть незначимым.

H_1: $|\theta| < 1 \Rightarrow b < 0$. Если ряд стационарен, то коэффициент $\boldsymbol{b}$ должен быть значимым и отрицательным.

Тестирование стационарности для AR(1)

$$\Delta \boldsymbol{y_t} = \boldsymbol{b} * \boldsymbol{y_{t-1}} + \boldsymbol{\varepsilon_t}$$

$H_0: b = 0.$ $H_1: b < 0.$

Идея теста: давайте оценим уравнение обычным МНК и проверим значимость коэффициента b при помощи обычной t-статистики:

$$\frac{\hat{b}}{se(\hat{b})}$$

Тестирование стационарности для AR(1)

$$\Delta \boldsymbol{y_t} = \boldsymbol{b} * \boldsymbol{y_{t-1}} + \boldsymbol{\varepsilon_t}$$

$H_0: b = 0.$ $H_1: b < 0.$

Идея теста: давайте оценим уравнение обычным МНК и проверим значимость коэффициента b при помощи обычной t-статистики:

$$\frac{\hat{b}}{se(\hat{b})}$$

Проблема: если верна гипотеза H_0, то эта статистика не будет иметь t-распределение Стьюдента ⇒ нужны другие критич. значения

Тестирование стационарности для AR(1): тест Дики — Фуллера (DF)

Оцениваем уравнение:

$$\Delta \boldsymbol{y_t} = \boldsymbol{b} * \boldsymbol{y_{t-1}} + \boldsymbol{\varepsilon_t}$$

$H_0: b = 0.$ $H_1: b < 0.$

Расчетное значение статистики:

$$\hat{\tau} = \frac{\widehat{b}}{se(\widehat{b})}$$

Сравниваем расчетное значение с критическим значением из специальных таблиц Дики и Фуллера (см., например, табл. 8.1. в Вербике)

Тестирование стационарности для AR(1): тест Дики — Фуллера (DF)

Вычисляем критическую статистику: $\hat{\tau} = \dfrac{\hat{b}}{se(\hat{b})}$

Сравниваем расчетное значение с критическим значением из специальных таблиц Дики и Фуллера (см., например, табл. 8.1. в Вербике)

Если расчетное значение отрицательное и меньше критического (то есть по модулю больше!), то гипотеза $\boldsymbol{H_0}$ отвергается $\Rightarrow$ делаем вывод о том, что ряд стационарен.

В остальных модификациях теста процедура принятия решения будет аналогичной

Тест Дики — Фуллера (DF) и его обобщения

Мы рассмотрели самый простой случай, когда тестируется стационарность AR(1) процесса без константы. В прикладных исследованиях важны и более общие случаи, которые будут рассмотрены далее:

- Тест Дики — Фуллера с константой
- Тест Дики — Фуллера с константой и трендом
- Расширенный тест Дики — Фуллера (augmented DF, ADF)

Тест Дики — Фуллера с константой

$$y_t = \delta + \theta y_{t-1} + \varepsilon_t$$

$H_0\colon \theta = 1$ — ряд является нестационарным

- содержит единичный корень,
- описывается процессом случайного блуждания с дрейфом

$H_1\colon |\theta| < 1$ — ряд является стационарным

- не содержит единичный корень,
- описывается стационарным авторегрессионным процессом первого порядка (с, вообще говоря, ненулевым математическим ожиданием)

Тест Дики — Фуллера с константой

$$y_t = \delta + \theta y_{t-1} + \varepsilon_t$$

$$y_t - y_{t-1} = \delta + \theta y_{t-1} - y_{t-1} + \varepsilon_t$$

$$\Delta y_t = \delta + (\theta - 1) * y_{t-1} + \varepsilon_t$$

Обозначим $\theta - 1 = b$.

Оцениваем уравнение $\Delta \boldsymbol{y_t} = \boldsymbol{\delta} + \boldsymbol{b} * \boldsymbol{y_{t-1}} + \boldsymbol{\varepsilon_t}$

Расчетное значение статистики:

$$\hat{\tau} = \frac{\hat{b}}{se(\hat{b})}$$

Сравниваем расчетное значение с критическим значением из специальных таблиц Дики и Фуллера (для теста с константой)

Тест Дики — Фуллера с константой и трендом

$$y_t = \delta + \theta y_{t-1} + \varphi t + \varepsilon_t$$

$\boldsymbol{H_0}$: $\boldsymbol{\theta = 1}$ — ряд является нестационарным описывается процессом случайного блуждания с дрейфом

Также в этом случае говорят, что ряд содержит *стохастический* тренд.

$\boldsymbol{H_1}$: $\boldsymbol{|\theta| < 1}$ — ряд является стационарным.

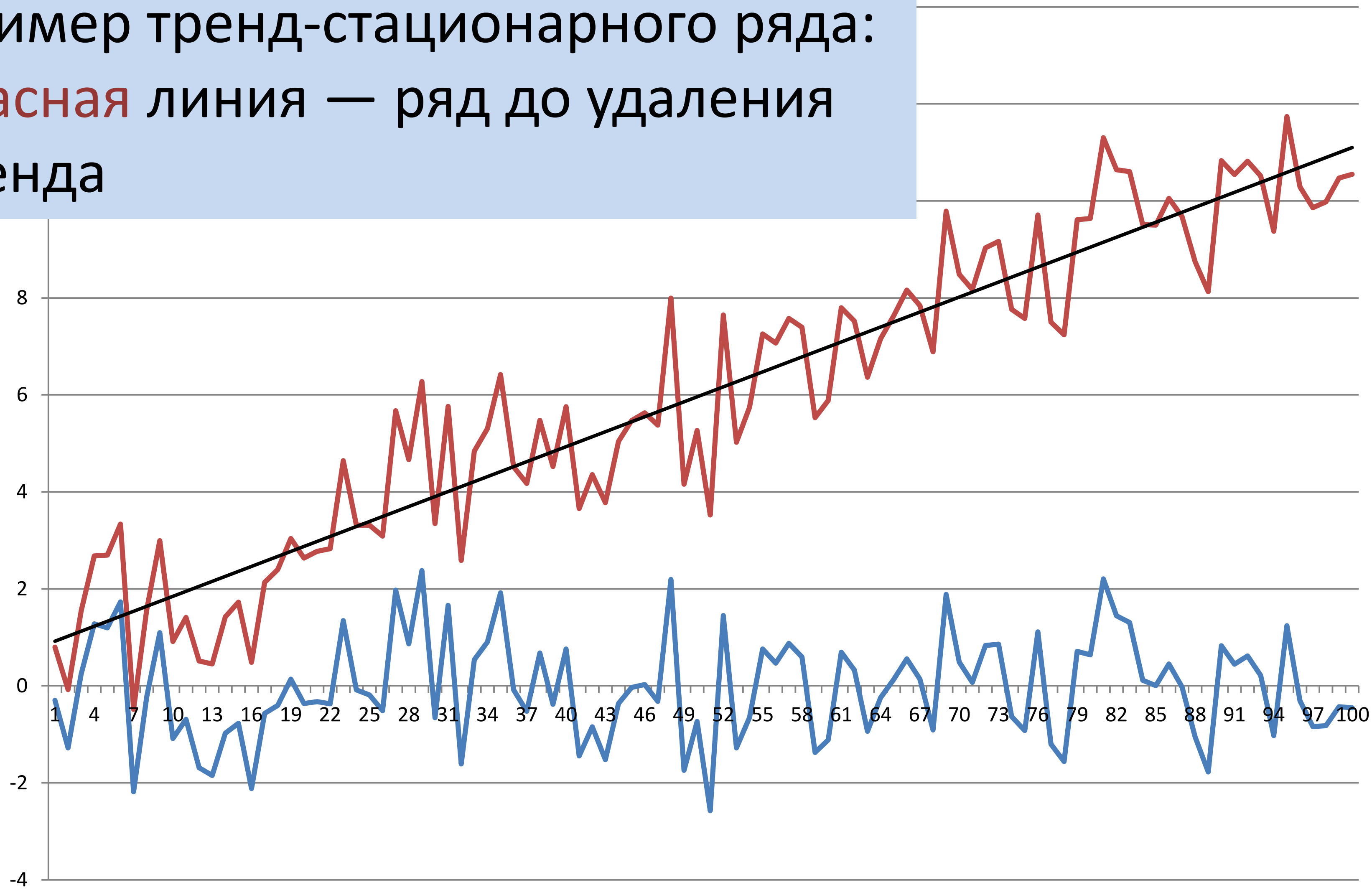
При $|\theta| < 1$ и $\varphi \neq 0$ ряд y_t называется стационарным относительно линейного тренда (тренд-стационарным, trend-stationary)

Также в этом случае говорят, что ряд содержит только *детерминированный* тренд.

В этом случае ряд $z_t = y_t - \varphi t$ стационарен

Тест Дики — Фуллера с константой и трендом

Пример тренд-стационарного ряда: красная линия — ряд до удаления тренда



Тест Дики — Фуллера с константой и трендом

$$y_t = \delta + \theta y_{t-1} + \varphi t + \varepsilon_t$$

$$y_t - y_{t-1} = \delta + \theta y_{t-1} - y_{t-1} + \varphi t + \varepsilon_t$$

$$\Delta y_t = \delta + (\theta - 1) * y_{t-1} + \varphi t + \varepsilon_t$$

Обозначим $\theta - 1 = b$.

Оцениваем уравнение $\Delta \boldsymbol{y_t} = \boldsymbol{\delta} + \boldsymbol{b} * \boldsymbol{y_{t-1}} + \boldsymbol{\varphi t} + \boldsymbol{\varepsilon_t}$

Расчетное значение статистики:

$$\hat{\tau} = \frac{\widehat{b}}{se(\widehat{b})}$$

Сравниваем расчетное значение с критическим значением из специальных таблиц Дики и Фуллера (для теста с константой и трендом)

Расширенный тест Дики — Фуллера

(Augmented DF-test, ADF-test)

Рассмотрим более общий случай авторегрессионного процесса

$$y_t = \theta_1 * y_{t-1} + \cdots + \theta_p * y_{t-p} + \varepsilon_t$$

H_0: ряд является нестационарным, содержит единичный корень

H_1: ряд является стационарным процессом AR(p).

Расширенный тест Дики — Фуллера

Оцениваем уравнение

$$\Delta \boldsymbol{y_t} = \boldsymbol{b y_{t-1}} + \boldsymbol{c_1} \Delta \boldsymbol{y_{t-1}} + \cdots + \boldsymbol{c_{p-1}} \Delta \boldsymbol{y_{t-p+1}} + \boldsymbol{\varepsilon_t}$$

Расчетное значение статистики:

$$\hat{\tau} = \frac{\hat{b}}{se(\hat{b})}$$

- Аналогично можно осуществлять ADF-тест с добавлением константы и тренда.
- Порядок лага для ADF-теста можно выбирать при помощи информационного критерия Шварца, который мы обсудим на следующей лекции

Тест Kwiatkowski, Phillips, Schmidt, Shin (KPSS)

Альтернативным тестом для проверки стационарности является KPSS-тест

H_0: Ряд является тренд-стационарным
H_1: Ряд является нестационарным

Обратите внимание, что в этом тесте нулевая гипотеза (в отличие от нулевой гипотезы ADF-теста) соответствует стационарности

Тест Kwiatkowski, Phillips, Schmidt, Shin (KPSS)

1. Оцениваем регрессию: $y_t = \delta + \varphi t + \varepsilon_t$

2. Вычисляем остатки $e_1, e_2, \dots e_T$

3. Вычисляем вспомогательные суммы (T штук):

$$S_t = \sum_{m=1}^{T} e_m$$

4. Вычисляем расчетное значение статистики:

$$KPSS = \sum_{t=1}^{T} \frac{S_t^2}{\widehat{\sigma^2}},$$

где $\widehat{\sigma^2}$ — оценка дисперсии случайной ошибки

Тест Kwiatkowski, Phillips, Schmidt, Shin (KPSS)

5. Если расчетное значение статистки меньше критического значения, равного **0,146**, то нулевая гипотеза не отвергается при уровне значимости 5%. Можно сделать вывод о стационарности ряда.

Замечание: если нулевой гипотезой является стационарность (а не тренд-стационарность), то процедура теста аналогична, только на первом шаге оценивается уравнение $y_t = \delta + \varepsilon_t$, а критическое значение равно **0,463**

Оценка моделей ARIMA

Методология Бокса — Дженкинса

На этой лекции мы рассмотрим решение следующей задачи:

Имеется T наблюдений временного ряда:

$$y_1, y_2, \dots, y_T$$

Необходимо подобрать ***ARIMA(p,d,q)*** модель, которая хорошо описывает динамику этого временного ряда.

Методология Бокса — Дженкинса

Шаг 1. Определение порядка интегрированности ряда и переход к стационарным разностям

Шаг 2. Анализ автокорреляционной функции и частной автокорреляционной функции

Шаг 3. Оценивание и проверка адекватности модели

Шаг 4. Прогнозирование

Шаг 1. Определение порядка интегрированности ряда и переход к стационарным разностям

1. Тестируем ряд на стационарность, используя тесты, которые мы обсудили ранее
2. Если ряд оказался стационарным, то переходим к шагу 2. Если нет — то переходим к разностям ряда и снова тестируем стационарность
3. И так до тех пор, пока не получим стационарный ряд
4. Таким образом, на этом шаге определяется параметр ***d*** модели ***ARIMA(p,d,q)***, то есть порядок интегрированности ряда
5. Далее в рамках шагов 2 и 3 следует работать со стационарными разностями ряда

Шаг 2. Анализ автокорреляционной функции и частной автокорреляционной функции

Эмпирическая автокорреляционная функция временного ряда (**ACF**) — выборочный аналог теоретической автокорреляционной функции, рассчитывается на основе выборочных коэффициентов корреляции:

$$ACF(k) = \widehat{\rho_k} = \widehat{Corr}(y_t, y_{t-k})$$

Шаг 2. Анализ автокорреляционной функции и частной автокорреляционной функции

Эмпирическая **частная** автокорреляционная функция временного ряда (**PACF**) рассчитывается на основе выборочных частных коэффициентов корреляции.

Определим выборочный частный коэффициент корреляции k-го порядка как МНК-оценку для θ_k в модели AR(k)

$$PACF(k) = \widehat{\theta_k}$$

Шаг 2. Анализ автокорреляционной функции и частной автокорреляционной функции

На шаге 2 следует построить и проанализировать графики ACF и PACF для рассматриваемого временного ряда.

Далее описано поведение типичных графиков ACF и PACF для разных видов временных рядов.

Шаг 2. Анализ автокорреляционной функции и частной автокорреляционной функции

Случай А. Процесс AR($\boldsymbol{p}$)

1. ACF бесконечна по протяженности и только в пределе при $k \to \infty$ сходится к нулю
2. PACF равна (или близка) к нулю для лагов, больших, чем $\boldsymbol{p}$.

Случай Б. Процесс MA($\boldsymbol{q}$)

1. ACF равна (или близка) к нулю для лагов, больших, чем $\boldsymbol{q}$.
2. PACF бесконечна по протяженности и только в пределе при $k \to \infty$ сходится к нулю

Случай В. Если не А и не Б, то у вас ARMA(p,q)

Шаг 2. Анализ автокорреляционной функции и частной автокорреляционной функции

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.539	0.539	116.40	0.000
		2	0.319	0.041	157.37	0.000
		3	0.190	0.004	171.91	0.000
		4	0.092	-0.029	175.35	0.000
		5	0.014	-0.044	175.43	0.000
		6	0.012	0.033	175.50	0.000
		7	-0.013	-0.026	175.56	0.000
		8	0.025	0.059	175.81	0.000
		9	0.042	0.018	176.52	0.000
		10	0.069	0.042	178.47	0.000
		11	0.027	-0.051	178.78	0.000
		12	0.036	0.028	179.32	0.000

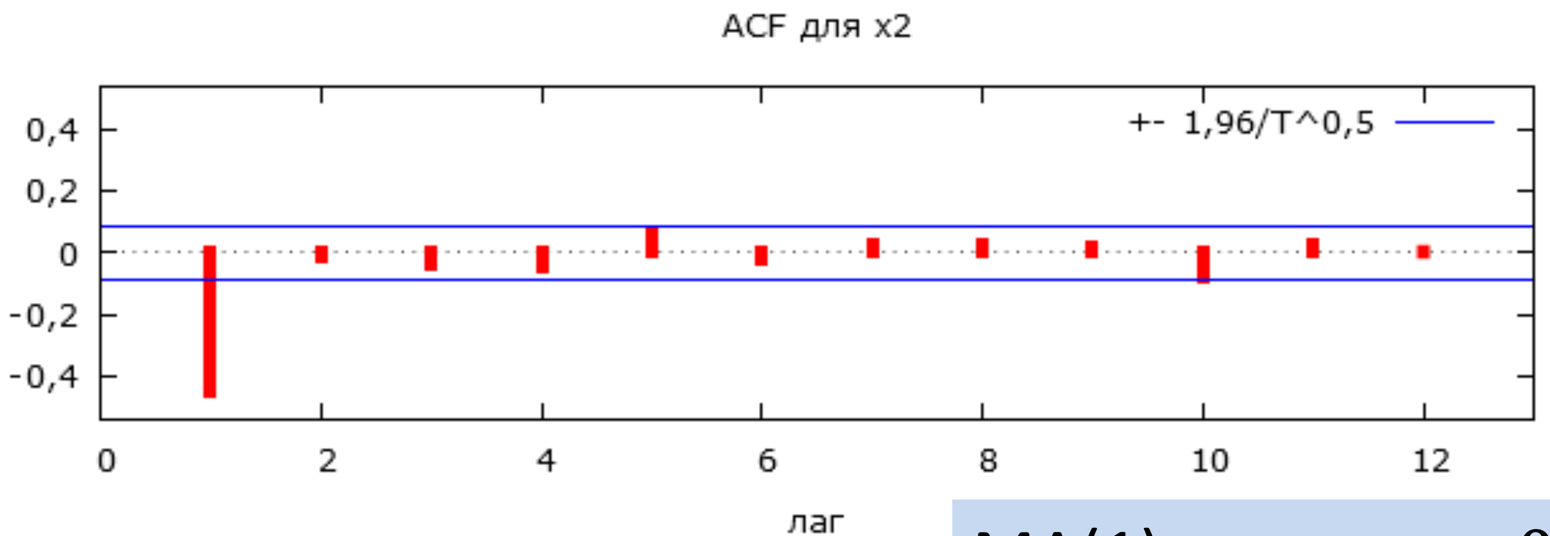
Рис. 11.9. AR(1). $Y_t = 0.5Y_{t-1} + \varepsilon_t$. Корень $\mu = 2$

Шаг 2. Анализ автокорреляционной функции и частной автокорреляционной функции

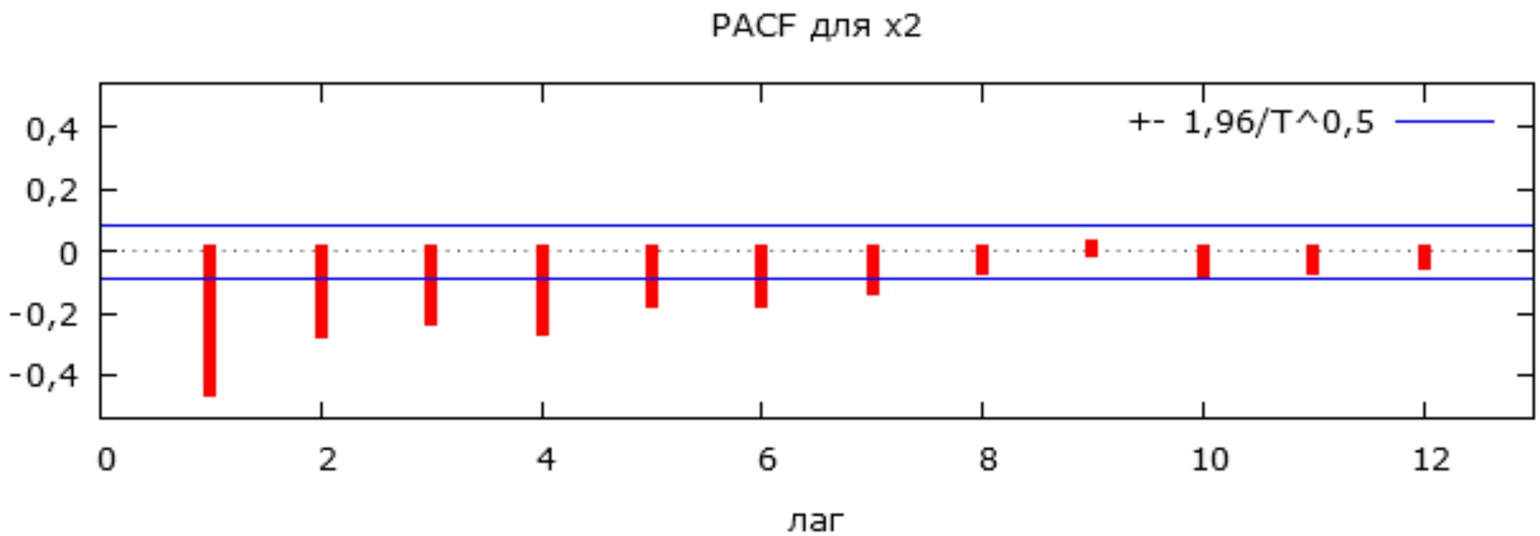
Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.500	-0.500	100.19	0.000
		2	0.281	0.041	131.88	0.000
		3	-0.125	0.041	138.15	0.000
		4	0.104	0.063	142.49	0.000
		5	-0.106	-0.049	147.01	0.000
		6	0.090	0.009	150.33	0.000
		7	-0.096	-0.043	154.11	0.000
		8	0.080	0.011	156.70	0.000
		9	-0.068	-0.010	158.57	0.000
		10	0.103	0.074	162.91	0.000
		11	-0.081	0.009	165.60	0.000
		12	0.063	-0.002	167.23	0.000

Рис. 11.10. AR(1). $Y_t = -0.5Y_{t-1} + \varepsilon_t$. Корень $\mu = -2$

Шаг 2. Анализ автокорреляционной функции и частной автокорреляционной функции



MA(1): $y_t = \varepsilon_t - 0{,}8\varepsilon_{t-1}$



Шаг 2. Анализ автокорреляционной функции и частной автокорреляционной функции

- Анализ коррелограмм на втором шаге позволяет сделать предварительные предположения о возможных порядках авторегрессии $\boldsymbol{p}$ и скользящего среднего $\boldsymbol{q}$.
- Эмпирические ACF и PACF не обязаны в точности совпадать с теоретическими, но должны быть похожи на них.
- По возможности рекомендуется использовать экономичные модели: $p + q \leq 3$ (если нет сезонной компоненты).

Шаг 3. Оценивание и проверка адекватности модели

- Для каждой из выбранных на втором шаге моделей оцениваются их параметры.
- Обычно оценивание производится при помощи ММП. Для AR моделей состоятельные оценки также дает обычный МНК.
- Каждая из моделей проверяется на адекватность на основе критериев, представленных далее.
- Наилучшая из моделей выбирается в качестве итоговой для использования на четвертом шаге.

Шаг 3. Оценивание и проверка адекватности модели

Критерии адекватности ARMA модели

1. Значимость коэффициентов модели

2. Анализ остатков модели

Остатки должны быть белым шумом => должны иметь нулевую автокорреляцию => все элементы ACF для ряда остатков должны незначимо отличатся от нуля

2. Информационные критерии

Шаг 3. Оценивание и проверка адекватности модели

Тестирование отсутствия автокорреляции: тестирование гипотезы о равенстве нулю отдельного коэффициента автокорреляции

H_0: $\rho_k = 0$

Тестовая статистика: $\widehat{\rho_k} \sim N(0, \frac{1}{T})$

Если $|\widehat{\rho_k}| < \frac{1{,}96}{\sqrt{T}}$, то при уровне значимости 5% гипотеза H_0 принимается (не отклоняется)

Шаг 3. Оценивание и проверка адекватности модели

Тестирование отсутствия автокорреляции:

Q-статистика Бокса — Пирса

$H_0: \rho_1 = \rho_2 = \dots = \rho_K = 0$

$$Q = T * \sum_{i=1}^{K} \widehat{\rho_i}^2 \sim \chi^2(K - p - q)$$

p и q — параметры ARIMA модели

Шаг 3. Оценивание и проверка адекватности модели

Тестирование отсутствия автокорреляции:

Тест Льюинга — Бокса

$H_0: \rho_1 = \rho_2 = \cdots = \rho_K = 0$

$$\tilde{Q} = T(T+2) * \sum_{i=1}^{K} \frac{\widehat{\rho_i}^2}{T-i} \sim \chi^2(K-p-q)$$

p и q — параметры ARIMA модели

Шаг 3. Оценивание и проверка адекватности модели

Информационный критерий Шварца

Schwarz information criterion (**SIC**)

Так же называется Байесовским информационным критерием

Bayes information criterion (**BIC**)

$$SIC = lnT\frac{p+q}{T} + \ln\left(\frac{\sum e_t^2}{T}\right)$$

p и q — параметры ARIMA модели, если в модель включена константа, то вместо $p+q$ следует использовать $p+q+1$

Шаг 3. Оценивание и проверка адекватности модели

Информационный критерий Шварца

- Можно использовать для сравнения разных моделей с одинаковой зависимой переменной
- Следует выбирать модель с наименьшим значением критерия.
- Можно использовать не только для ARIMA, но и для любых других моделей временных рядов, в этом случае вместо $p + q$ следует поставить k —число оцениваемых коэффициентов в модели (считая константу):

$$SIC = lnT * \frac{k}{T} + \ln\left(\frac{\sum e_t^2}{T}\right)$$

Шаг 3. Оценивание и проверка адекватности модели

Информационный критерий Акаике

$$AIC = 2\frac{p+q}{T} + \ln\left(\frac{\sum e_t^2}{T}\right)$$

Работает аналогично критерию Шварца, однако используется реже, так как асимптотически критерий Акаике приводит к выбору перепараметризированных моделей

Шаг 4. Прогнозирование

После выбора наилучшей модели можно использовать ее для прогнозирования

(см. первые лекции по временным рядам)

Для анализа качества прогноза можно использовать процедуру псевдо вневыборочного прогнозирования — ***Pseudo Out-of-Sample Forecasts***

Pseudo Out-of-Sample Forecasts

1. Выберите P наблюдений, которые будут использоваться для вневыборочного прогноза (обычно — 10–15% выборки)
2. Оцените параметры модели, используя наблюдения $y_1, y_2, \ldots, y_{T-P}$.
3. Вычислите прогноз на один шаг вперед $\hat{y}_{T-P+1}$. Вычислите ошибку прогноза $y_{T-P+1} - \hat{y}_{T-P+1}$.
4. Повторяйте шаги 2–4, каждый раз увеличивая выборку на одно наблюдение.

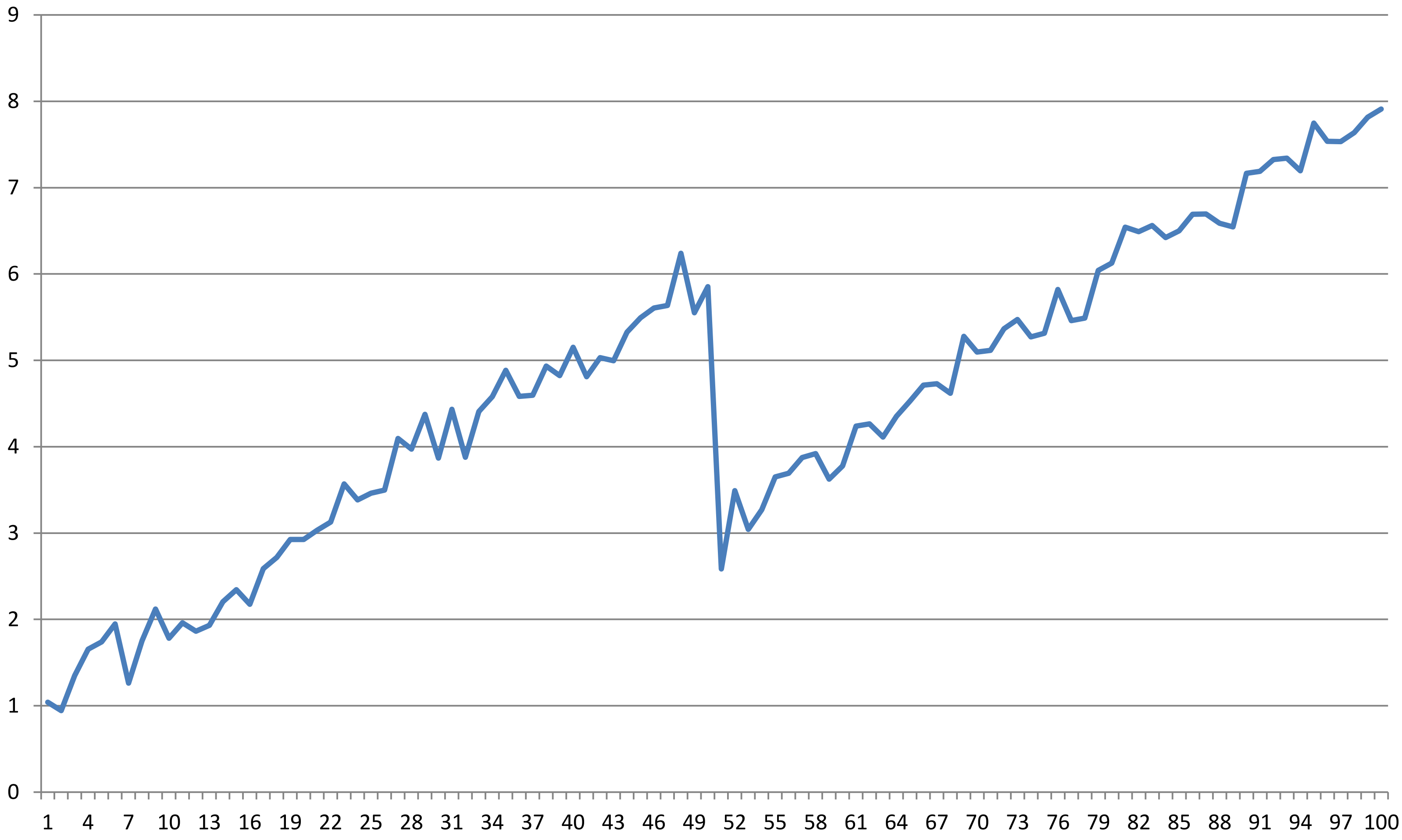
Таким образом, вы получите P ошибок прогноза. Проанализировав их поведение, вы можете оценить прогностическое качество модели

Pseudo Out-of-Sample Forecasts

- В хорошей модели средняя ошибка прогноза должна быть близка к нулю (не должно быть систематического занижения или завышения прогноза)
- Ошибки прогноза должны вести себя как стационарный временной ряд
- Средняя квадратичная ошибка прогноза должна быть не слишком велика.

Качество прогнозирования можно использовать как дополнительный критерий выбора наилучшей модели среди нескольких кандидатов

Структурные сдвиги (Breaks)



Под структурным сдвигом будем понимать изменение значений параметров модели, описывающей динамику ВР

Структурные сдвиги (Breaks)

Тестирование структурных сдвигов:

Случай 1. Дата структурного сдвига известна

Случай 2. Дата структурного сдвига **не**известна

Тестирование гипотезы о структурном сдвиге, когда дата потенциального сдвига известна

На примере AR(1) модели:

$$y_t = \beta_1 + \beta_2 y_{t-1} + \varepsilon_t, \qquad t < t_0$$

$$y_t = (\beta_1 + \alpha_1) + (\beta_2 + \alpha_2) y_{t-1} + \varepsilon_t, \qquad t \geq t_0$$

Тест на линейное ограничение:

$$y_t = \beta_1 + \alpha_1 d_t + \beta_2 y_{t-1} + \alpha_2 d_t y_{t-1} + \varepsilon_t$$

$$d_t = \begin{cases} 0, & \text{если } t < t_0 \\ 1, & \text{если } t \geq t_0 \end{cases}$$

$$H_0\colon \alpha_1 = \alpha_2 = 0$$

Обозначим $F(t_0)$ — расчетное значение соответствующей F-статистики для тестирования гипотезы о сдвиге в момент времени t_0

Тестирование гипотезы о структурном сдвиге, когда дата потенциального сдвига неизвестна

1. Берем значения $t_0 = 0{,}15T$ и $t_1 = 0{,}85T$ (с округлением до ближайшего целого)

2. Вычисляем значения $F(t_0), F(t_0+1), \dots, F(t_1)$ (как на предыдущем слайде)

3. Вычисляем **Quandt likelihood ratio (QLR) statistic** (or sup-Wald statistic)

$$QLR = \max(F(t_0), F(t_0+1), \dots, F(t_1))$$

4. Критические значения QLR-статистики можно найти в таблице 14.6. учебника [Stock, Watson] или в статье [Quandt, 1960]

Тестирование гипотезы о структурном сдвиге, когда дата потенциального сдвига неизвестна

1. Отвержение нулевой гипотезы в QLR-тесте может указывать на единственный резкий сдвиг, множественные резкие сдвиги или на постепенное изменение коэффициентов в модели

2. Если речь идет о единственном сдвиге, то его дату указывает максимальное значение $F(t)$

3. QLR-тест можно использовать и для многомерных моделей временных рядов, которые будут обсуждаться на будущих лекциях (например, для ADL моделей)