

# Оценивание модели ARIMA

## Лекция 13



**БАЛАКОВСКИЙ ИНЖЕНЕРНО-ТЕХНОЛОГИЧЕСКИЙ ИНСТИТУТ**  
ФИЛИАЛ НАЦИОНАЛЬНОГО ИССЛЕДОВАТЕЛЬСКОГО ЯДЕРНОГО УНИВЕРСИТЕТА «МИФИ»



# 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)

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

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

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

Обозначим  $(\theta - 1) = b$ .  $\Delta y_t = b \cdot y_{t-1} + \varepsilon_t$

# Тестирование стационарности для AR(1)

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

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

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

Обозначим  $(\theta - 1) = b$ .  $\Delta y_t = b \cdot y_{t-1} + \varepsilon_t$

В этом случае:

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

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

# Тестирование стационарности для AR(1)

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

$H_0: b = 0$

$H_1: b < 0$

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

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

# Тестирование стационарности для AR(1)

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

$H_0: b = 0$

$H_1: b < 0$

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

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

**Проблема:** если верна гипотеза  $H_0$ , то эта статистика не будет иметь t-распределение Стьюдента  $\Rightarrow$  нужны другие критические значения.



# Тестирование стационарности для

AR(1):

тест Дики – Фуллера (DF)

Оцениваем уравнение:

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

$H_0: b = 0.$

$H_1: b < 0.$

Расчетное значение статистики:

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

Сравниваем расчетное значение с критическим значением из специальных таблиц Дики и Фуллера.

# Тестирование стационарности для

AR(1):

**тест Дики – Фуллера (ДФ)**

Вычисляем расчетную статистику  $\hat{t} = \frac{b}{se(\hat{b})}$

Сравниваем расчетное значение с критическим значением из специальных таблиц Дики и Фуллера.

**Если расчетное значение отрицательное и меньше критического (то есть по модулю больше!), то гипотеза  $H_0$  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

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

$H_0: \theta = 1$  – 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

# Тест Дики – Фуллера с константой

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

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

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

Обозначим  $(\theta - 1) = b$ .

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

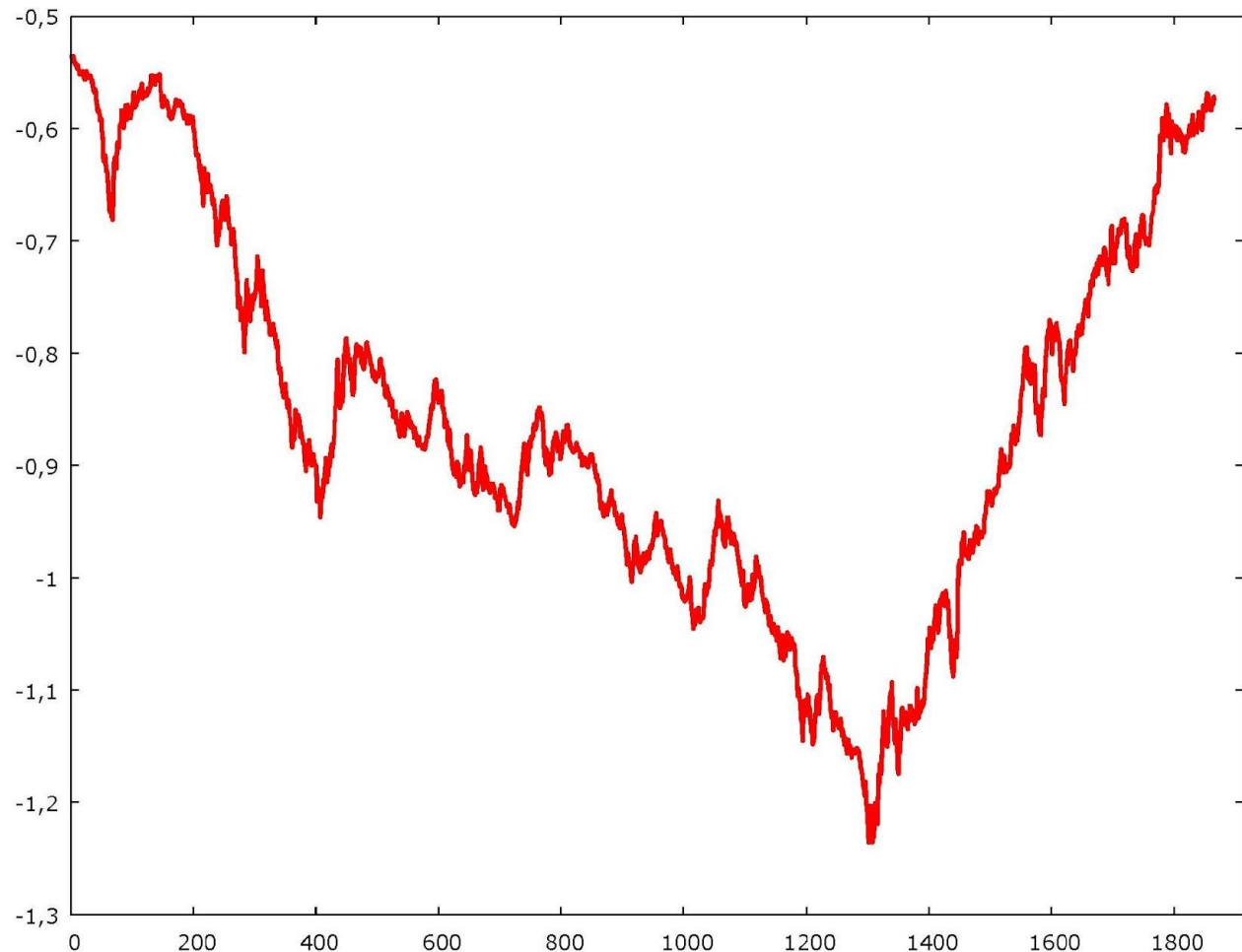
Расчетное значение статистики:

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

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

# Тест Дики — Фуллера с константой

Пример: Логарифм обменного курса доллара США к немецкой марке (2 января 1980 — 21 мая 1987)



# Тест Дики – Фуллера с константой

Пример: Логарифм обменного курса доллара США к немецкой марке (2 января 1980 – 21 мая 1987)

## Результаты оценивания в Gretl

Тест Дики-Фуллера для I\_DM

Объем выборки 1866

нулевая гипотеза единичного корня:  $a = 1$

тест с константой

модель:  $(1-L)y = b_0 + (a-1)*y(-1) + e$

коэф. Автокорреляции 1-го порядка для  $e$ : -0,059

оценка для  $(a-1)$ : -0,00125568

тестовая статистика:  $\tau_c(1) = -1,19626$

P-значение 0,6782

# Тест Дики – Фуллера с константой

Пример: Логарифм обменного курса доллара США к немецкой марке (2 января 1980 – 21 мая 1987)

## Результаты оценивания в Gretl

Тест Дики-Фуллера для I\_DM

Объем выборки 1866

нулевая гипотеза единичного корня:  $a = 1$

тест с константой

модель:  $(1-L)y = b_0 + (a-1)*y(-1) + e$

коэф. Автокорреляции 1-го порядка для  $e$ : -0,059

оценка для  $(a-1)$ : **-0,00125568**

тестовая статистика:  $\tau_c(1) = -1,19626$

P-значение 0,6782



# Тест Дики – Фуллера с константой

Пример: Логарифм обменного курса доллара США к немецкой марке (2 января 1980 – 21 мая 1987)

## Результаты оценивания в Gretl

Тест Дики-Фуллера для I\_DM

Объем выборки 1866

нулевая гипотеза единичного корня:  $a = 1$

тест с константой

модель:  $(1-L)y = b_0 + (a-1)*y(-1) + e$

коэф. Автокорреляции 1-го порядка для  $e$ : -0,059

оценка для  $(a-1)$ : -0,00125568

тестовая статистика:  **$\tau_c(1) = -1,19626$**

P-значение 0,6782

# Тест Дики – Фуллера с константой

Пример: Логарифм обменного курса доллара США к немецкой марке (2 января 1980 – 21 мая 1987)

## Результаты оценивания в Gretl

Тест Дики-Фуллера для I\_DM

Объем выборки 1866

нулевая гипотеза единичного корня:  $a = 1$

тест с константой

модель:  $(1-L)y = b_0 + (a-1)*y(-1) + e$

коэф. Автокорреляции 1-го порядка для  $e$ : -0,059

оценка для  $(a-1)$ : -0,00125568

тестовая статистика:  $\tau_c(1) = -1,19626$

**P-значение 0,6782 => нестационарность**

# Тест Дики – Фуллера с константой и трендом

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

**H<sub>0</sub>:  $\theta = 1$**  – ряд является нестационарным, описывается процессом случайного блуждания с дрейфом

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

**H<sub>1</sub>:  $|\theta| < 1$**  – ряд является стационарным.

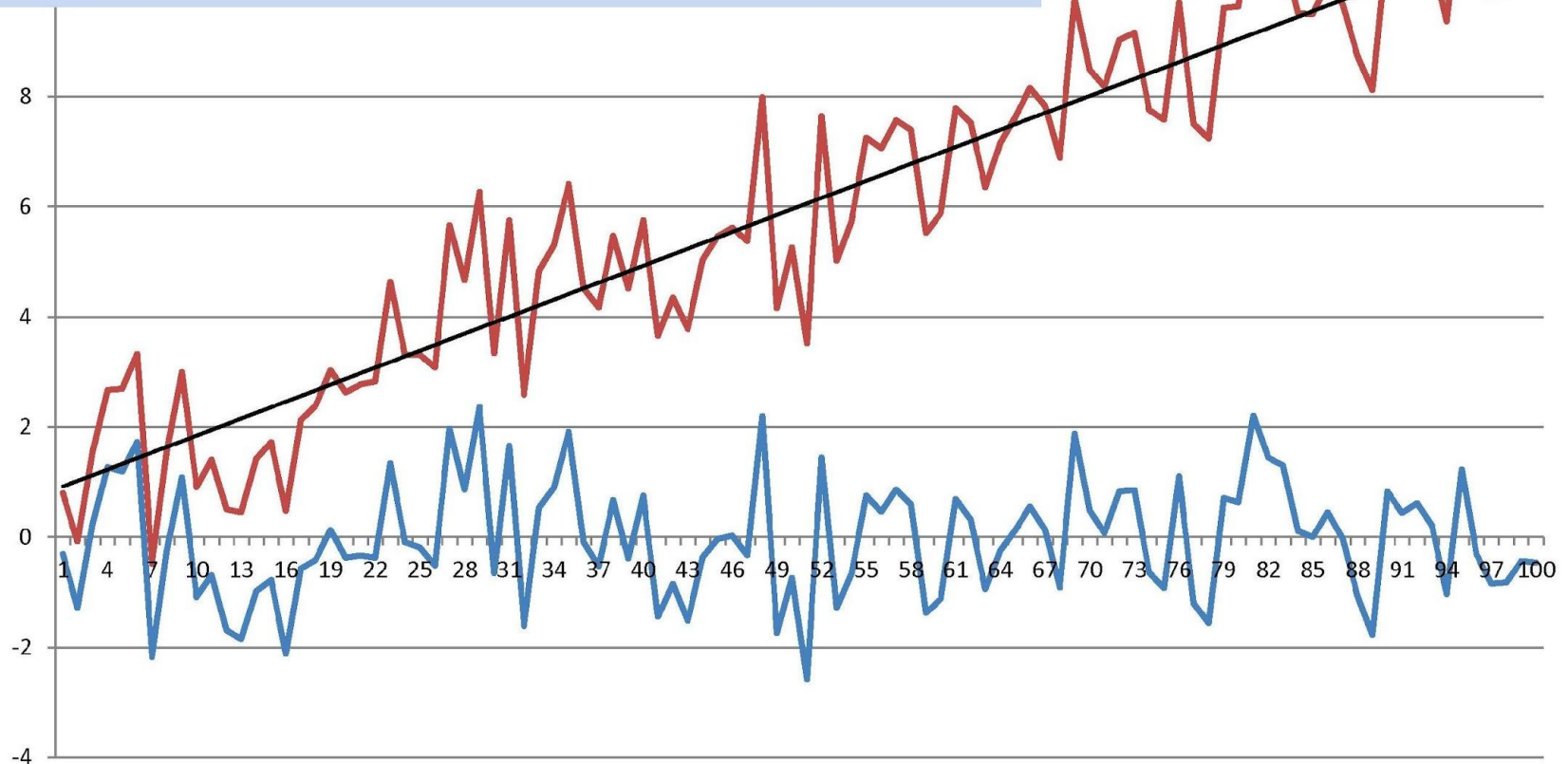
При  $|\theta| < 1$  и  $\varphi \neq 0$  ряд называется стационарным относительно линейного тренда (тренд-стационарным, trend-stationary).

**Также в этом случае говорят, что ряд содержит только детерминистический 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**

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

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

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

Обозначим  $(\theta - 1) = b$ .

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

Расчетное значение статистики  $\hat{t} = \frac{b}{se(\hat{b})}$

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:

$$y_t = \theta_1 \cdot y_{t-1} + \dots + \theta_p \cdot y_{t-p} + \varepsilon_t$$

$H_0$ : ряд является нестационарным, содержит единичный корень

$H_1$ : ряд является стационарным процессом AR(p)

# Расширенный тест Дики – Фуллера

Оцениваем уравнение:

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

Расчетное значение статистики:

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

- 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 (Т штук):

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

4. Вычисляем 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 статистики меньше критического значения, равного **0,146**, то нулевая гипотеза принимается. 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**.

# Методология Бокса-Дженкинса

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 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 автокорреляции.

$$ACF(k) = \hat{\rho}_k = \overline{Corr}(y_t, y_{t-k})$$

## 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  $\theta_k$  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 АСФ и РАСФ для рассматриваемого временного ряда.

Далее описано поведение типичных графиков для разных видов временных рядов.



## Шаг 2. Анализ автокорреляционной функции и частной автокорреляционной функции

### Случай А. Процесс $AR(p)$

























1. АСФ бесконечна по протяженности и только в пределе при  $k \rightarrow \infty$  сходится к нулю
2. РАСФ равна (или близка) к нулю для лагов, больших, чем  $p$

### Случай Б. Процесс $MA(q)$

1. АСФ равна (или близка) к нулю для лагов, больших, чем  $q$
2. РАСФ бесконечна по протяженности и только в пределе при  $k \rightarrow \infty$  сходится к нулю

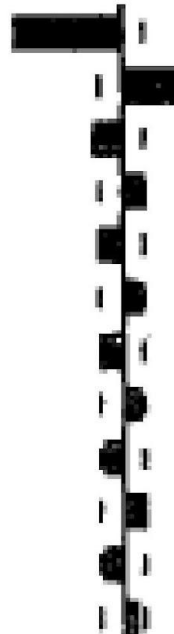
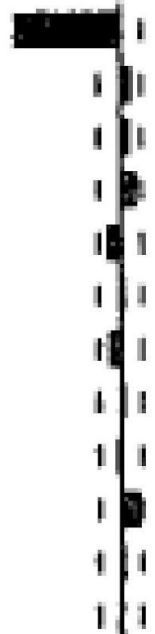
### Случай В. Если не А и не Б, то у вас $ARMA(p,q)$

## 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 |

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

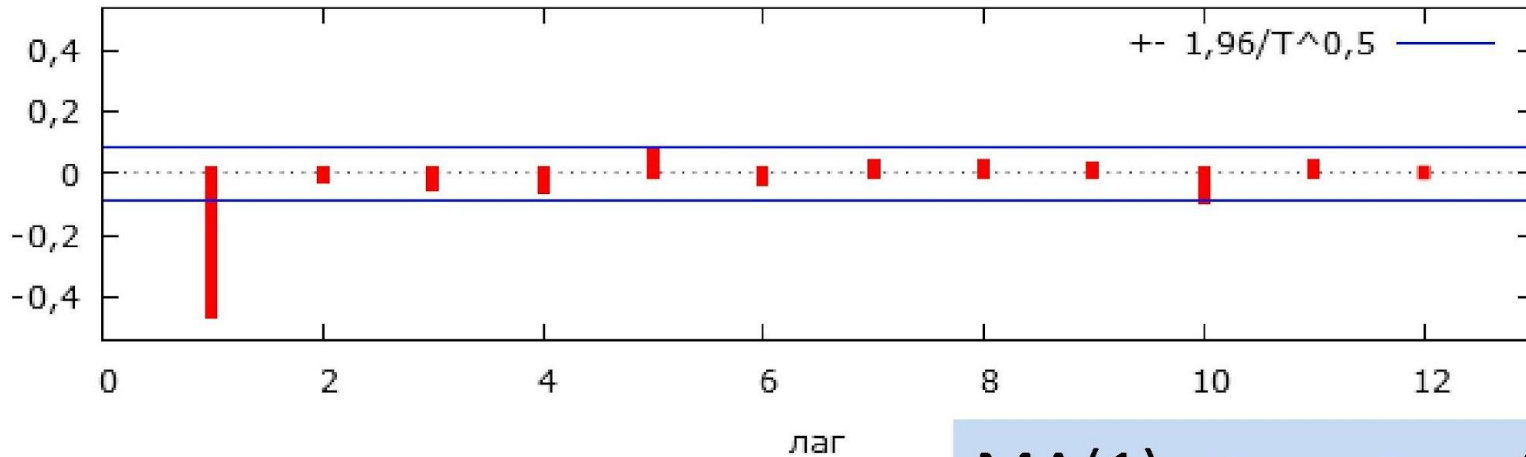
## Шаг 2. Анализ автокорреляционной функции и частной автокорреляционной функции

| 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 |

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

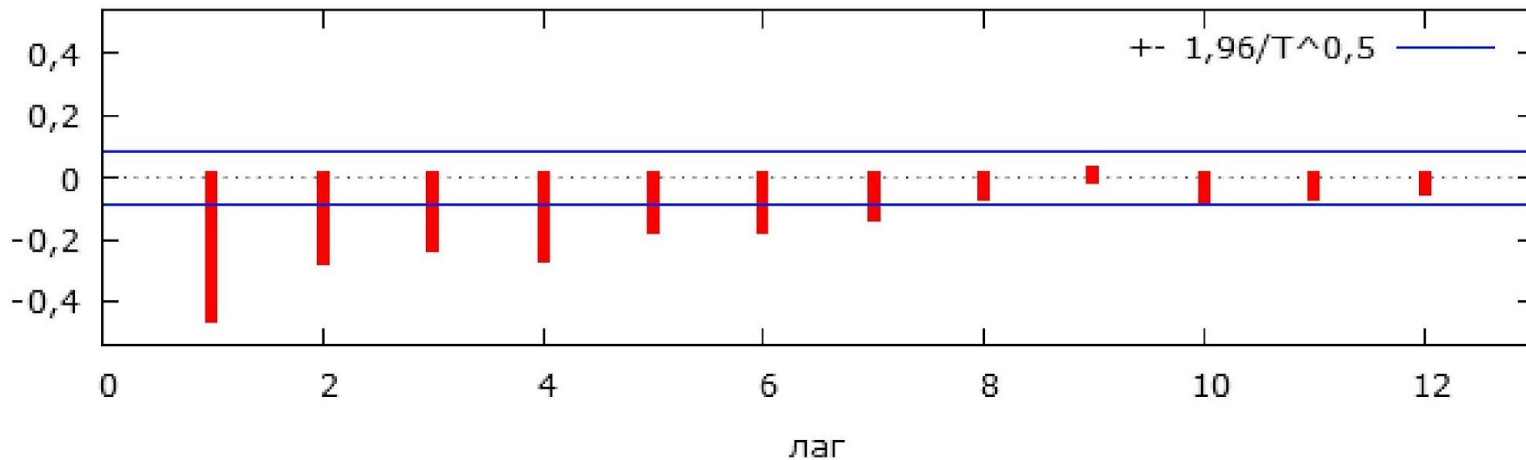
## Шаг 2. Анализ автокорреляционной функции и частной автокорреляционной функции

ACF для x2



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

PACF для x2



## Шаг 2. Анализ автокорреляционной функции и частной автокорреляционной функции

- Анализ коррелограмм на втором шаге позволяет сделать предварительные предположения о возможных порядках авторегрессии  $p$  и скользящего среднего  $q$ .
- Эмпирические АСФ и РАСФ 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)