

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 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**

❖ **Схема согласования результатов тестирования временного ряда с помощью тестов на принадлежность к TS и DS рядам (на примере тестов ADF и KPSS)**

<div> <div>KPSS</div> <div>ADF</div> </div>	Нулевая гипотеза	Альтернативная гипотеза
	H0: TS не отвергается	H1: TS отвергается
H0: DS не отвергается	Исход 1	Исход 2
H1: DS отвергается	Исход 3	Исход 4

- (1) если наблюдается исход 1, то это можно объяснить низкой мощностью обоих критериев;
- (2) если наблюдается исход 2, то это говорит в пользу DS-гипотезы;
- (3) если наблюдается исход 3, то это говорит в пользу TS-гипотезы;
- (4) если наблюдается исход 4, то это может говорить о том, что процесс порождения данных (DGP) не описывается DS или TS моделями, а может быть, например, дробно-интегрированным процессом или процессом с нелинейным трендом.

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 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 $AR(p)$

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

Случай Б. Процесс $MA(q)$

1. ACF равна (или близка) к нулю для лагов, больших, чем q .
2. PACF 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

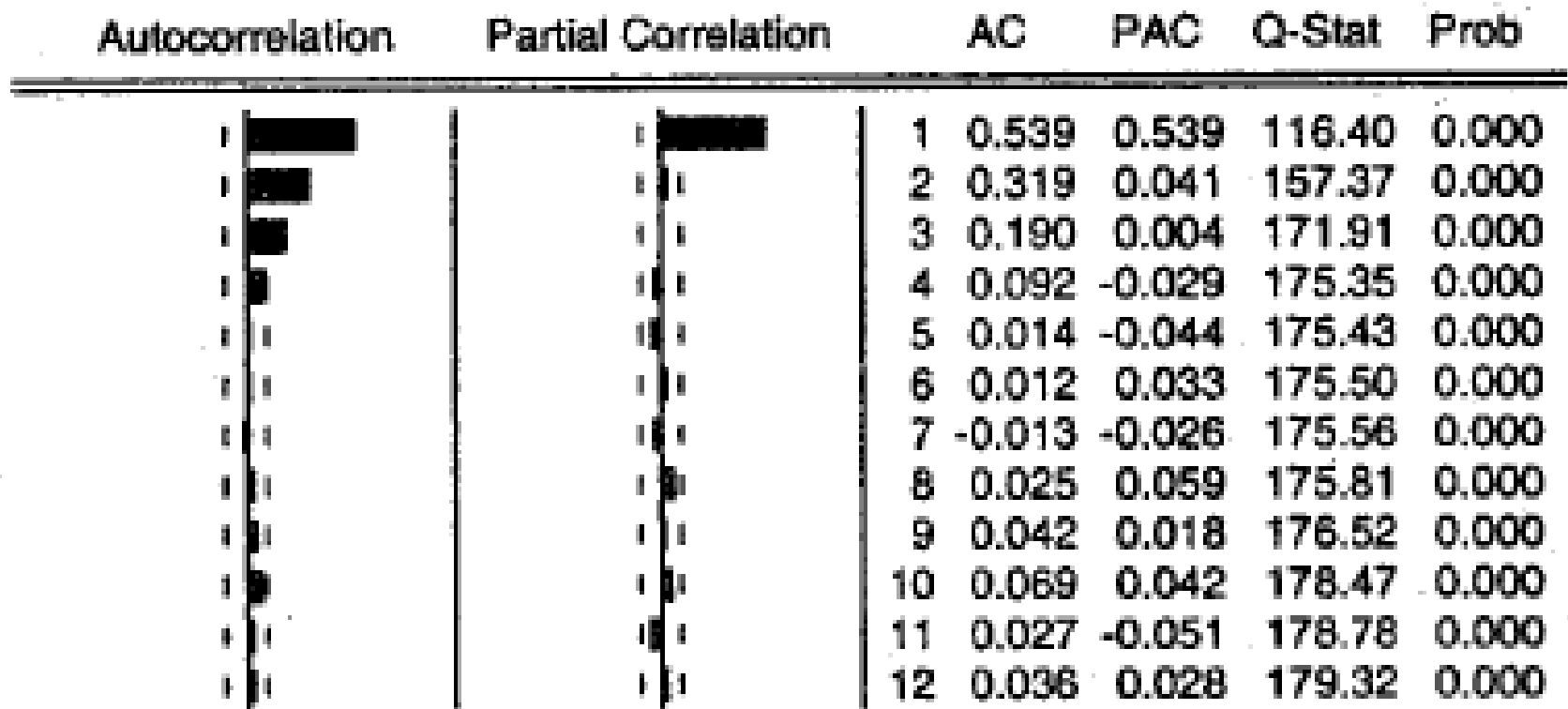


Рис. 11.9. AR(1). $Y_t = 0.5Y_{t-1} + \varepsilon_t$. Корень $\mu = 2$

Шаг 2. Анализ автокорреляционной функции и частной автокорреляционной функции

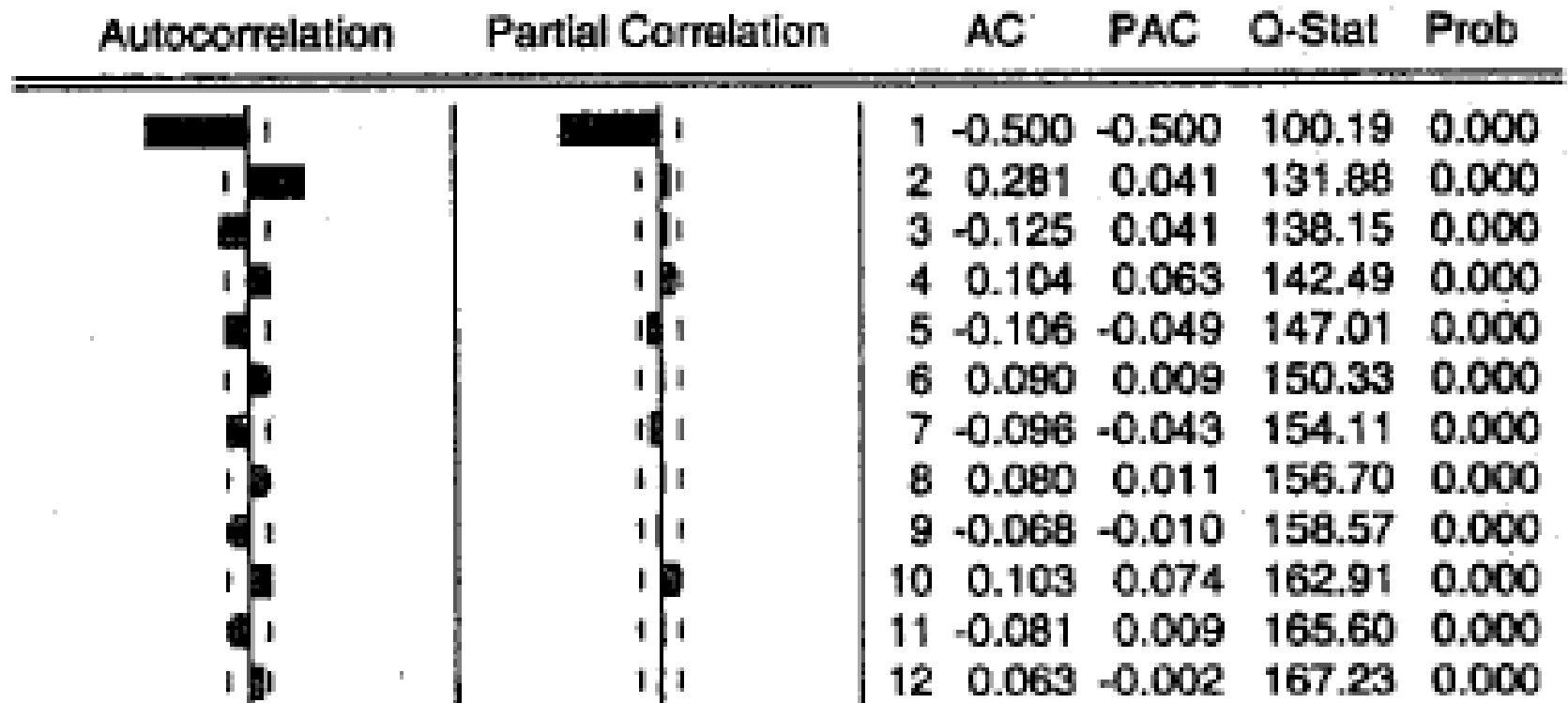
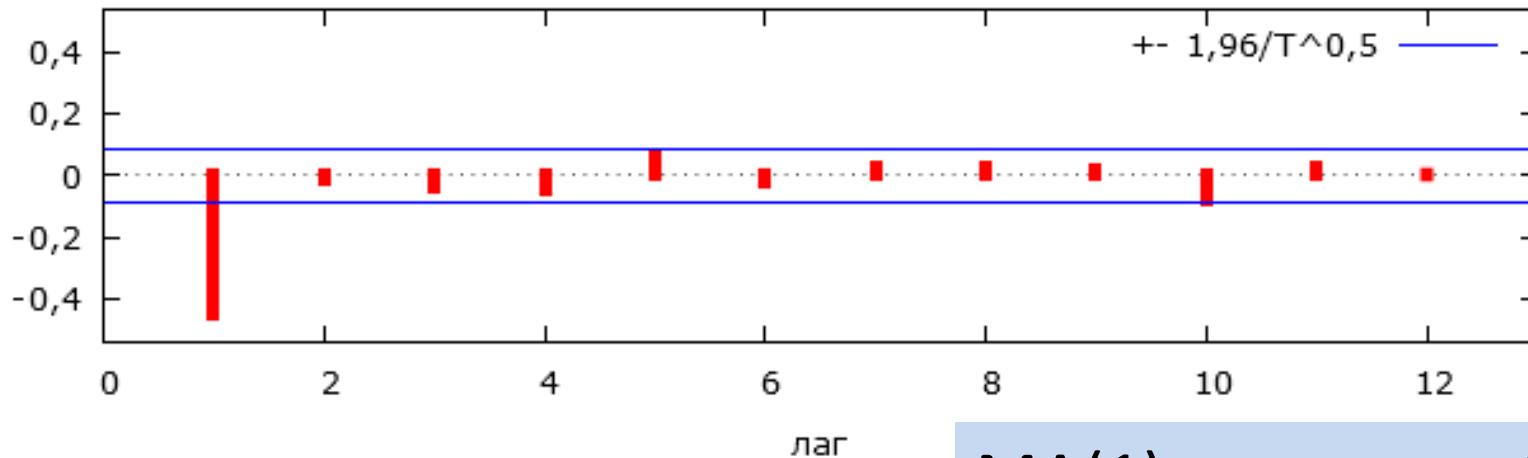


Рис. 11.10. AR(1). $Y_t = -0.5Y_{t-1} + \epsilon_t$. Корень $\mu = -2$

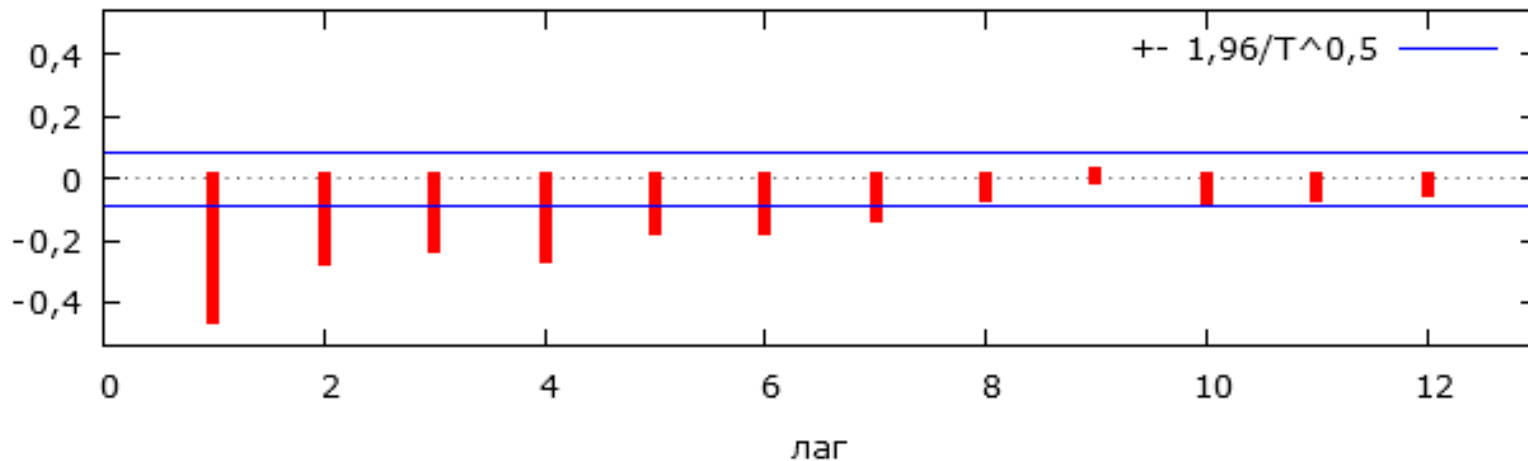
Шаг 2. Анализ автокорреляционной функции и частной автокорреляционной функции

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 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

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 отличаться от нуля

3. Информационные критерии

Шаг 3. Оценивание и проверка адекватности модели

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

$$H_0: \rho_k = 0$$

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

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:

Тест Льюинга — Бокса

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

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

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**)

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

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):

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

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 перепараметризованных моделей

Шаг 4. Прогнозирование

После выбора наилучшей модели можно использовать ее для прогнозирования в соответствии с тем, как мы обсуждали это выше.