Тема 2.

Модели стационарных временных рядов

Модели ARMA: методология Бокса-Дженкинса

Модели стационарных ВР - ARMA(p, q) AutoRegressive Moving Average models - модель авторегрессии – скользящего среднего

- Бокс и Дженкинс предложили данный класс моделей ко всем временным рядам (финансовым, так и к микроэкономическим) в 70-е.
- Было установлено, что практически все экономические процессы описываются моделями ARIMA с параметрами p и q не превышающими 2.
- Точность прогнозирования по моделям ARIMA оказалась выше.
- Были разработаны статистические пакеты для прогнозирования на основе ARIMA
- ARIMA используют в современных исследованиях

George Edward Pelham Box (18 October 1919 – 28 March 2013)

Gwilym Meirion Jenkins (1933 – 10 July 1982) was a Welsh statistician

Box, George and Jenkins, Gwilym (1970) Time series analysis: Forecasting and control, San Francisco: Holden-Day.

Модели ARMA в методологии Бокса-Дженкинса

1. **Идентификация модели (выбор параметров модели ARMA(p,q))** → необходимость знать характеристики процессов
2. **Оценивание**
3. **Тестирование и диагностика**
4. **Прогнозирование.**

Box, George and Jenkins, Gwilym (1970) Time series analysis: Forecasting and control, San Francisco: Holden-Day.

Класс ARIMA-моделей: примеры использования

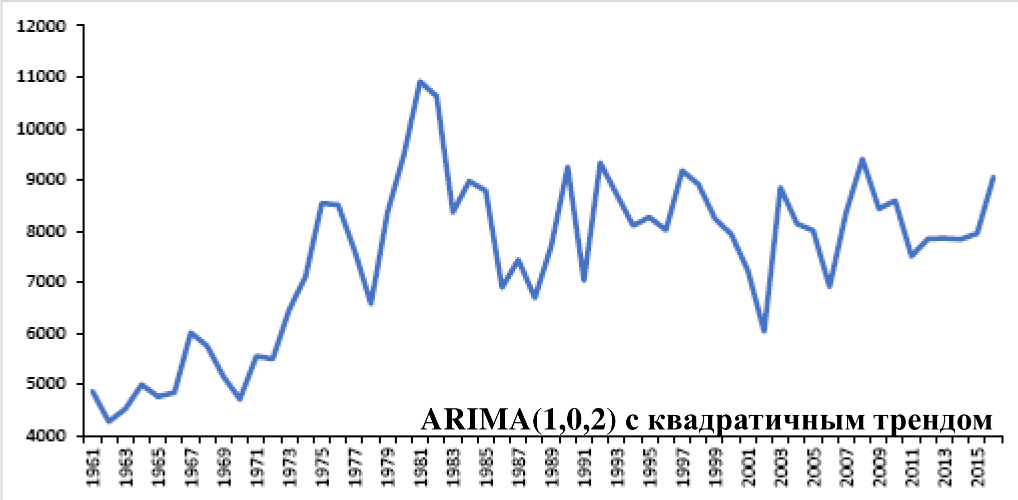
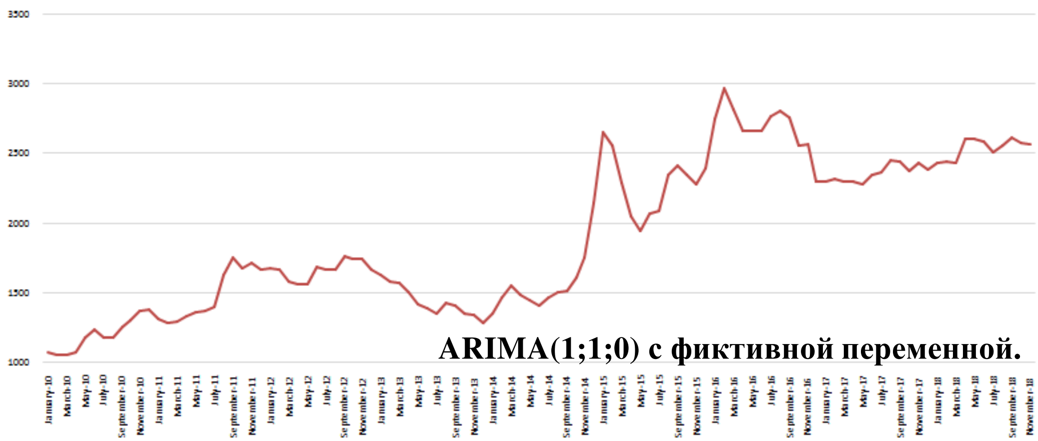
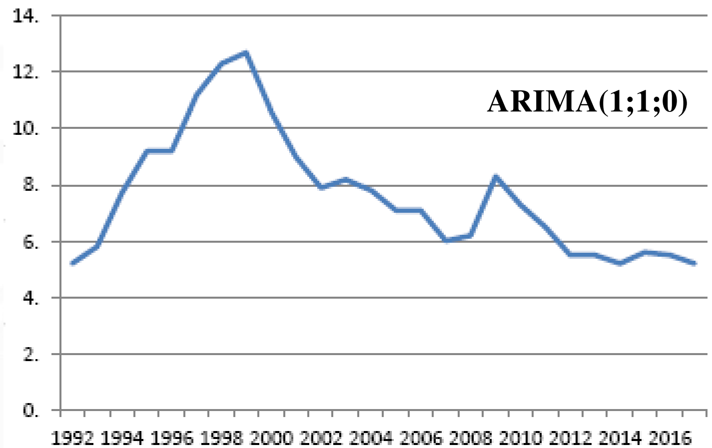


Рис. 1. Чистая стоимость производства пшеницы в США (в млн. постоянных долл. США 2004 - 2006) по годам 1961 - 2016
Источник: данные FAOSTAT (http://www.fao.org/faostat)

Значения реального ВВП Испании (млрд. евро)



Динамика цен на золото (руб./грамм) 2010 -2018



Уровень безработицы в России

Как выбрать параметры моделей?

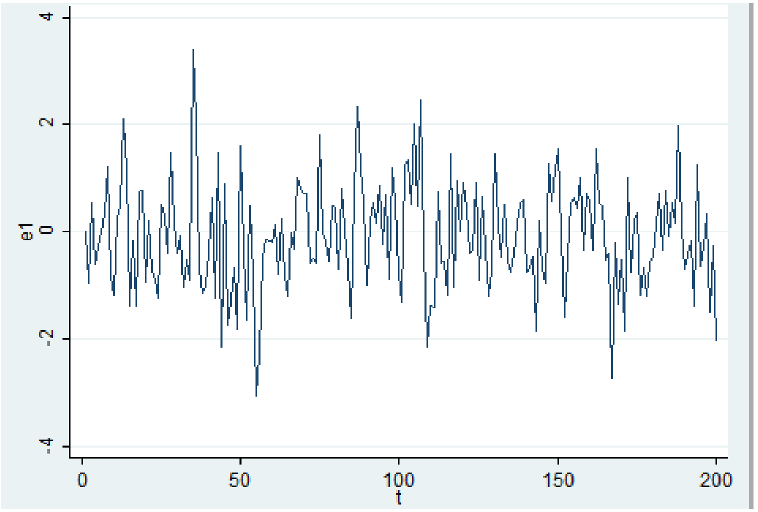
Стационарность временного ряда

Опр. 2.1. Случайный процесс наз. ***стационарным (слабо стационарным)* (weak stationary) (в широком смысле),** если он обладает постоянной средней и дисперсией, а ковариация зависит только от временного интервала между отдельными наблюдениями.

$$1.\,E(Y_t) = \mu$$

$$2.\,V(Y_t) = \sigma^2$$

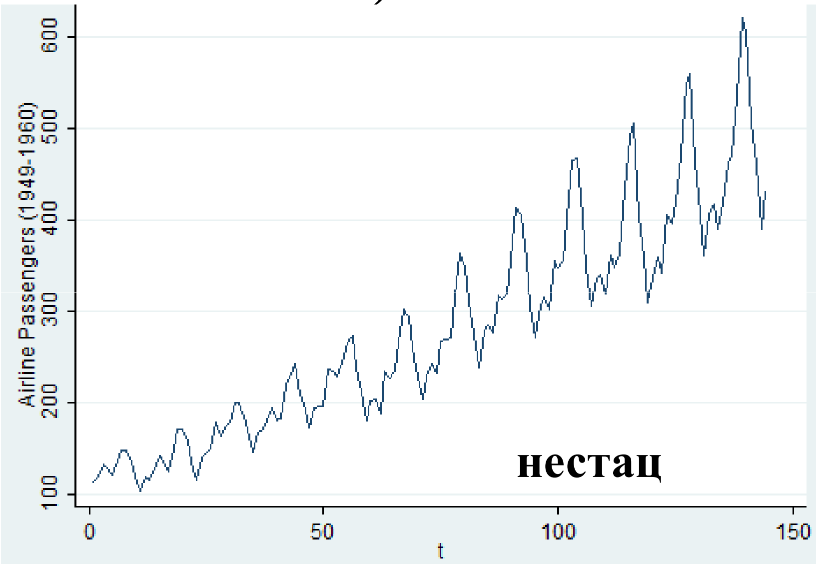
$$3.\,Cov(Y_t, Y_{t+\tau}) = M\left[(Y_t - \mu)(Y_{t+\tau} - \mu)\right] = \gamma(\tau)$$



Проверка временного ряда на стационарность

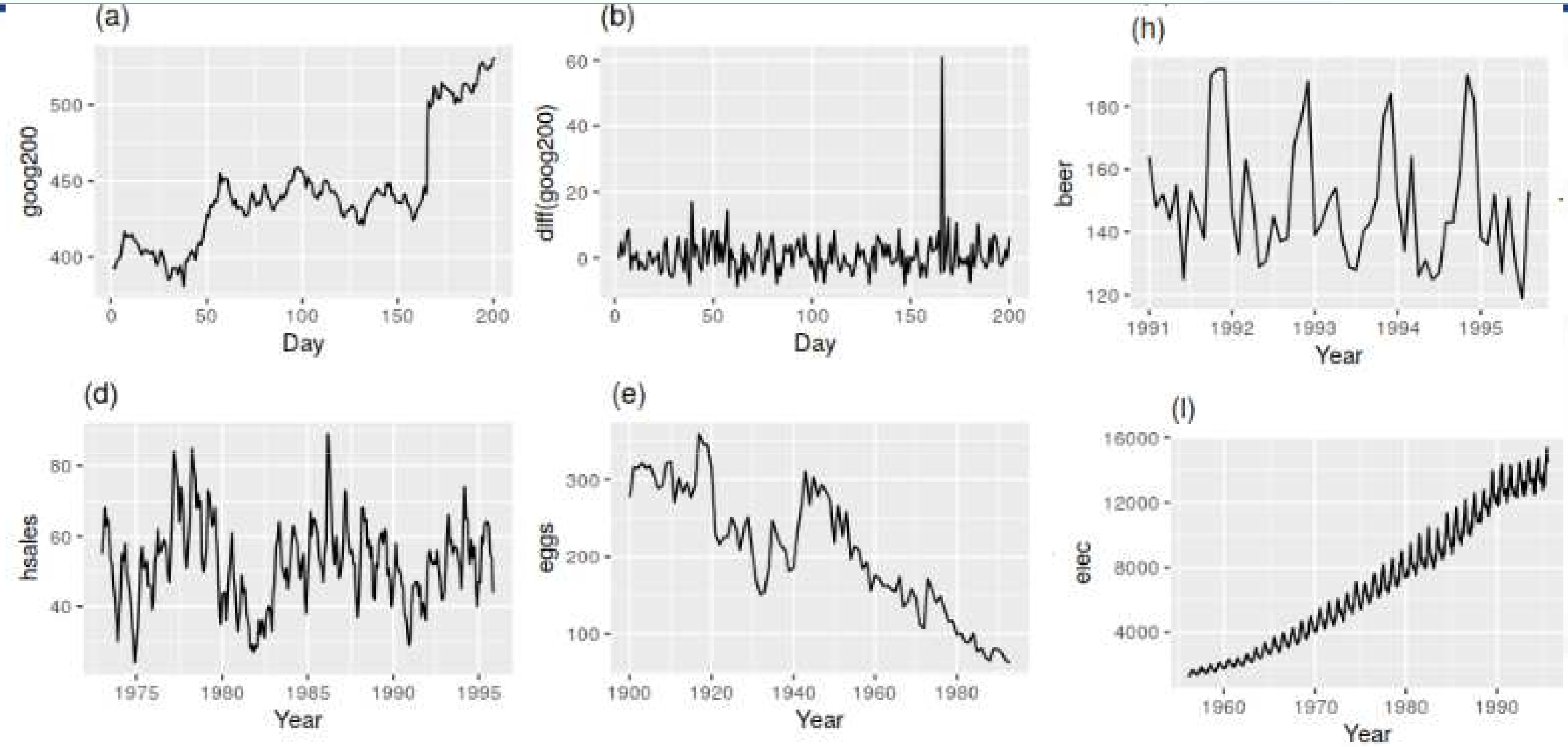
- Графический анализ (наличие тренда и анализ вариации показателя)

Как различить стационарный и нестационарный ряд?



- Анализ автокорреляционной и частной автокорреляционной функции,
- Тесты на стационарность

Какие из этих ВР являются стационарными?



(a) цена акций Google в течение 200 дней подряд; (b) Ежедневное изменение цены акций Google в течение 200 дней подряд; d) ежемесячные продажи новых домов на одну семью, продаваемых в США; e) годовая цена дюжины яиц в США (в постоянных долларах);

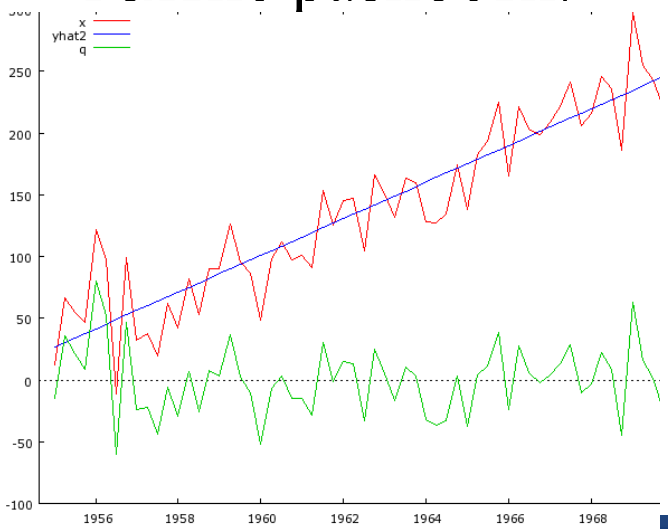
h) ежемесячное производство австралийского пива; (i) Ежемесячное производство электроэнергии в Австралии.

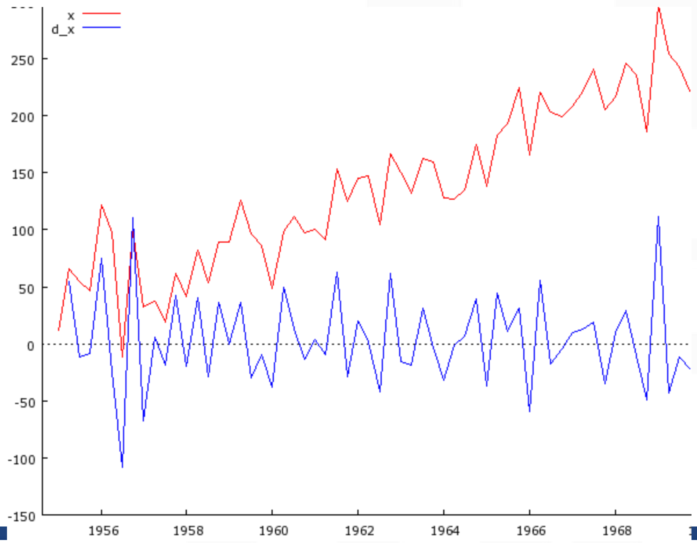
Hyndman, R.J., & Athanasopoulos, G. (2018) *Forecasting: principles and practice*, 2nd edition, OTexts: Melbourne, Australia. OTexts.com/fpp2.

! Большинство экономических показателей являются нестационарными.

Методы сведения ВР с стационарному:

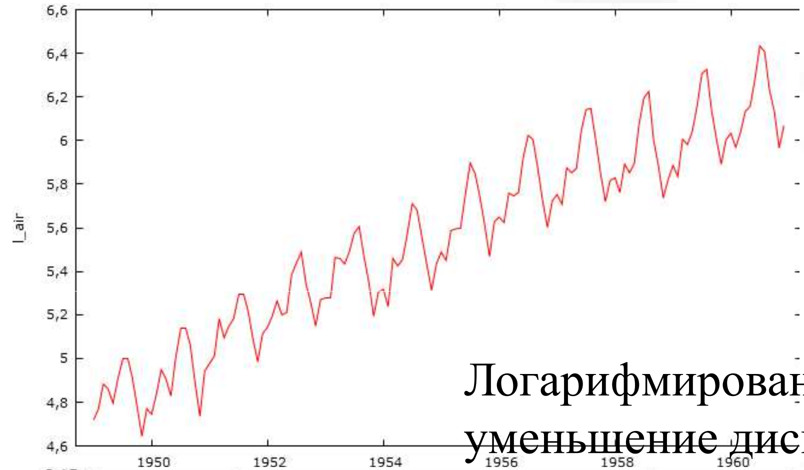
- выделение детерминированного тренда,
- фильтрация сезонной компоненты
- взятие разности.

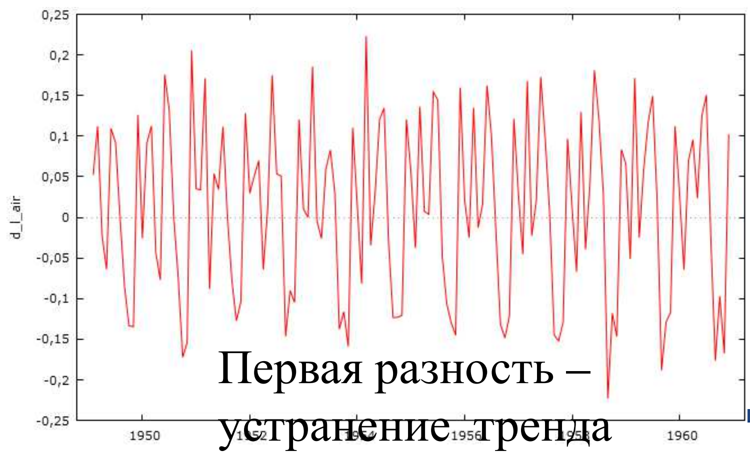


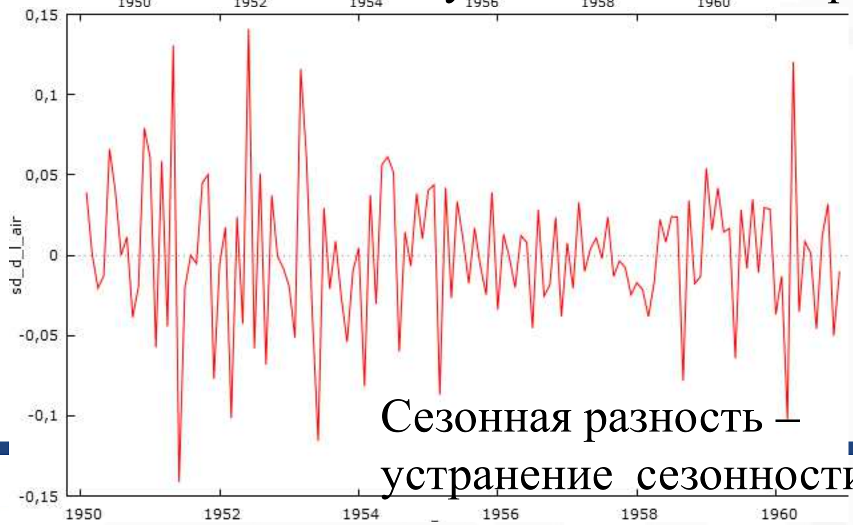


Что делать с нестационарными ВР?

Методы сведения ВР с стационарному (сезонность): выделение детерминированного тренда, фильтрация сезонной компоненты; взятие разности (обычной и сезонной).

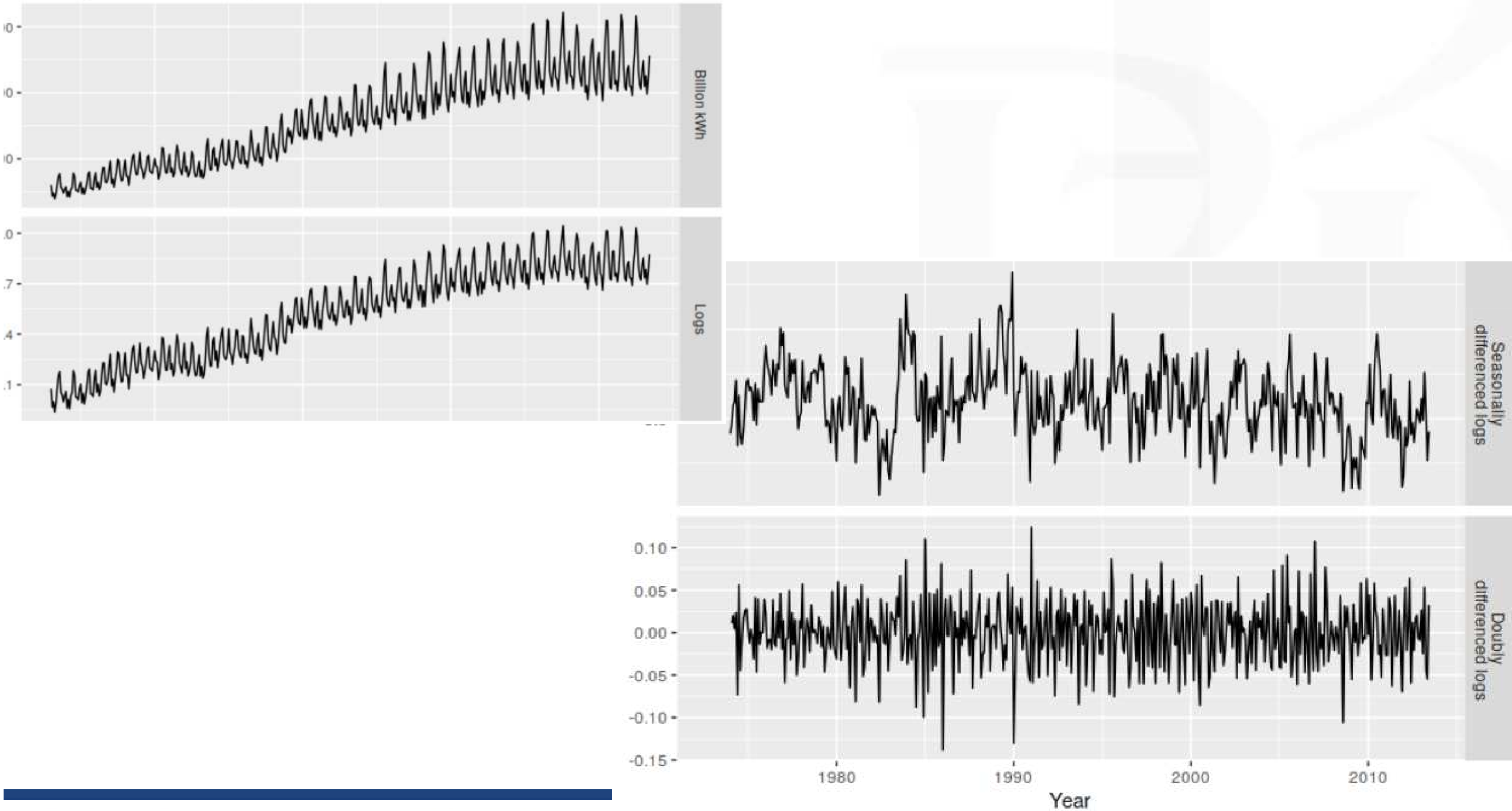






Сведение к стационарному ВР

Производство электроэнергии в США (млрд. КВтч).

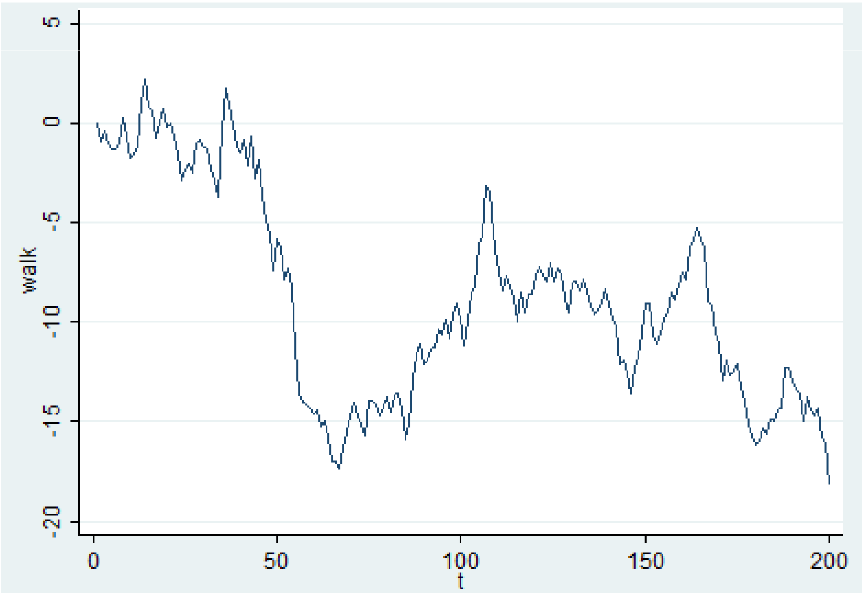


Простейшие примеры временных рядов

Пример 1. случайное блуждание $y_t = y_{t-1} + \varepsilon_t$

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

Что произойдет?



Простейшие примеры временных рядов

Пример 2. Процесс с линейным трендом $Z_t = \alpha + \beta t + \varepsilon_t$

Пример 3. Процесс с квадратичным трендом $Z_t = \alpha + \beta t + \gamma t^2 + \varepsilon_t$

Разложение Вольда

(Wold decomposition, 1938)

Недетерминированный стационарный случайный процесс можно представить в виде:

(*) $$y_t = Y_t - \mu = \sum_{\tau=0}^{\infty} \Psi_\tau \varepsilon_{t-\tau},$$ - линейной комбинации белых шумов с разными весовыми коэффициентами

$$\sum_{\tau=0}^{\infty} |\Psi_\tau| < \infty$$ (условие сходимости)

μ - математическое ожидание; $\varepsilon_t \sim WN(0,\sigma^2)$ - белый шум

Замечание. Выражение (*) наз. также **линейным фильтром.**

Модель авторегрессии – скользящего среднего ARMA(p, q) AutoRegressive Moving Average models

ARMA(p,q) процесс:

$$y_t = \underbrace{\alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \ldots + \alpha_p y_{t-p}}_{AR(p)} + \varepsilon_t + \underbrace{\theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta_q \varepsilon_{t-q}}_{MA(q)}$$

p – порядок авторегрессии,

q – порядок скользящего среднего.

2.1. Модель скользящего среднего MA(q) (Moving Average - MA(q) models)

Модель скользящего среднего MA(q) (Moving Average - MA(q) models)

Процесс скользящего среднего *q*-ого порядка:

$$y_t = \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \ldots + \theta_q\varepsilon_{t-q}$$

- значение ряда в момент времени *t* зависит от случайной компоненты в текущий и предыдущие *q* моментов времени.
- Частный случай ***разложения Вольда***
- «***скользящее среднее***» - текущее значение случайного процесса определяется взвешенным средним предыдущих значений БШ.
- MA(q) **стационарен** при любых q и θ

Модель скользящего среднего MA(q) (частные случаи)

Процесс скользящего среднего ***q*-ого** порядка **MA(q)** :

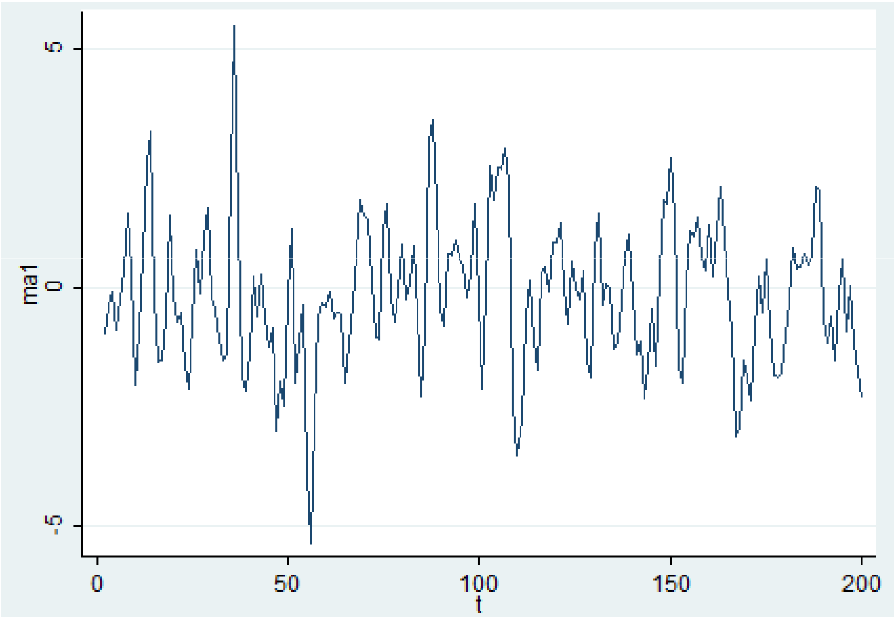
$$y_t = \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \ldots + \theta_q\varepsilon_{t-q}$$

MA(1): $y_t = \varepsilon_t + \theta_1\varepsilon_{t-1}$

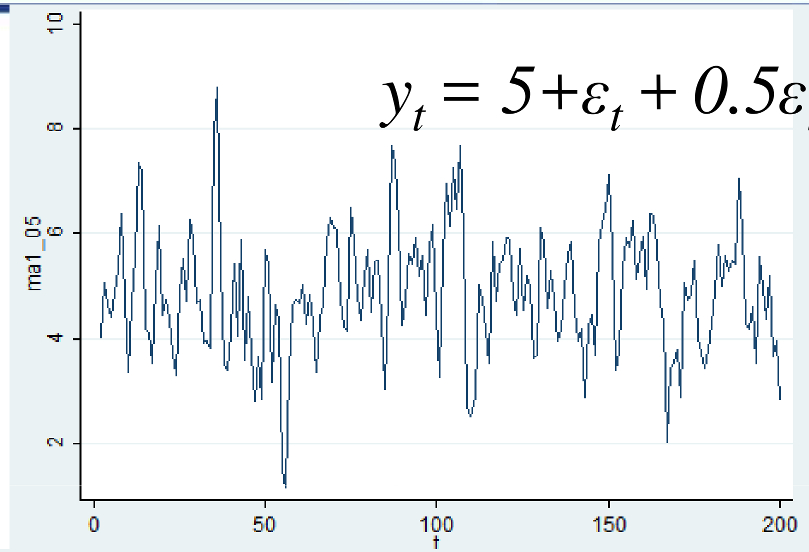
MA(2): $y_t = \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2}$

MA(1): примеры

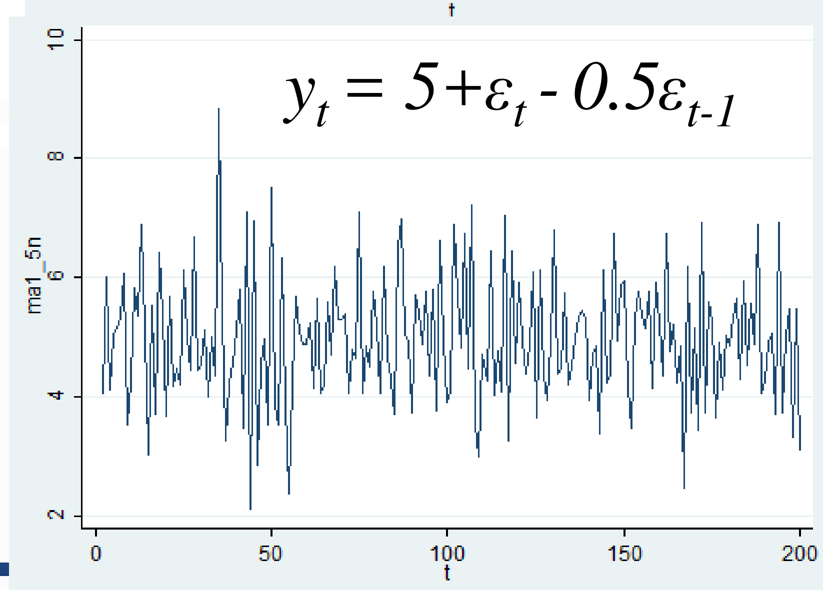
MA(1): $y_t = \theta_0 + \varepsilon_t + \theta_1 \varepsilon_{t-1}$



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



$y_t = 5 + \varepsilon_t + 0.5\varepsilon_{t-1}$

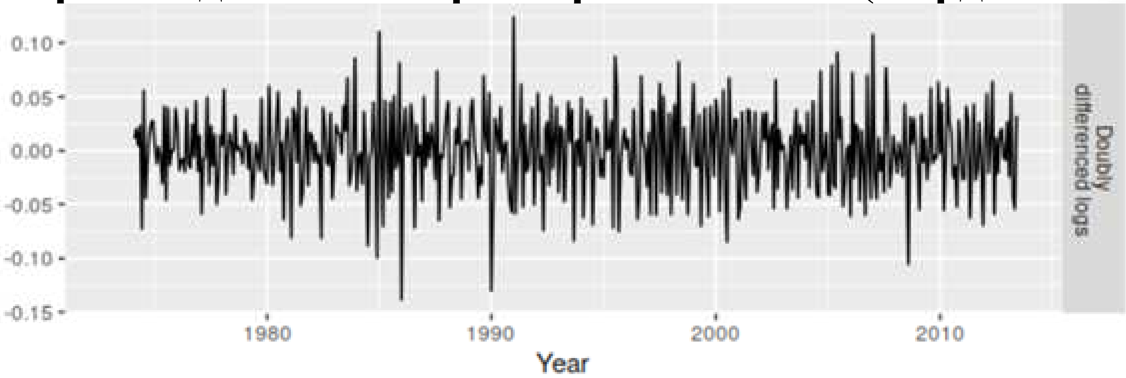


$y_t = 5 + \varepsilon_t - 0.5\varepsilon_{t-1}$

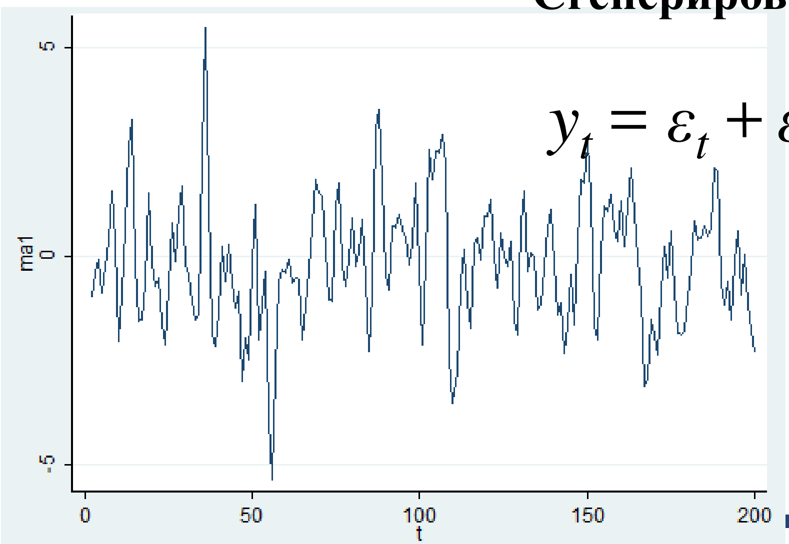
MA(1): примеры

MA(1): $y_t = \theta_0 + \varepsilon_t + \theta_1 \varepsilon_{t-1}$

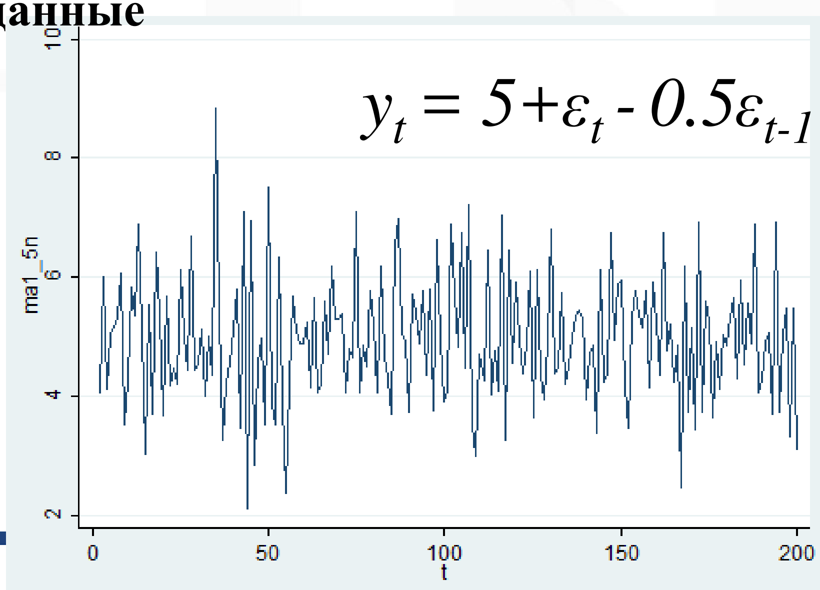
Производство электроэнергии в США (млрд. КВтч).



Сгенерированные данные



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



$y_t = 5 + \varepsilon_t - 0.5\varepsilon_{t-1}$

Модель скользящего среднего MA(1)

MA(1): $y_t = \varepsilon_t + \theta\varepsilon_{t-1}$

Числовые характеристики: $E(y_t) = 0$, $V(y_t) = (1+\theta^2)\sigma^2$,

$$\gamma(k) = \operatorname{cov}(y_t, y_{t+k}) = \begin{cases} (1+\theta^2)\sigma^2, \ k = 0 \\ \theta\sigma^2, \ k = 1 \\ 0, \ k > 1 \end{cases} \qquad \rho(k) = \frac{\operatorname{cov}(y_t, y_{t+k})}{V(y_t)} = \begin{cases} 1, \ k = 0 \\ \theta/(1+\theta^2), \ k = 1 \\ 0, \ k > 1 \end{cases}$$

MA(1): св-ва ACF

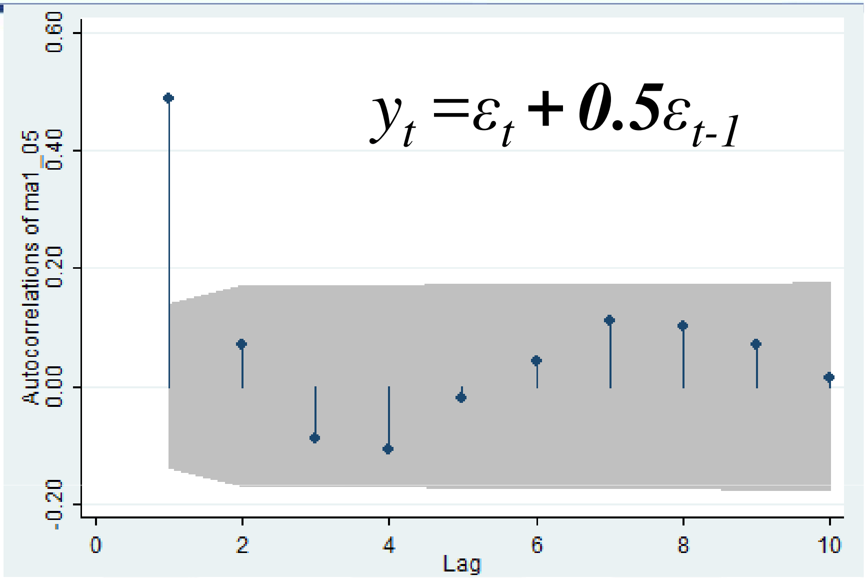
MA(1):

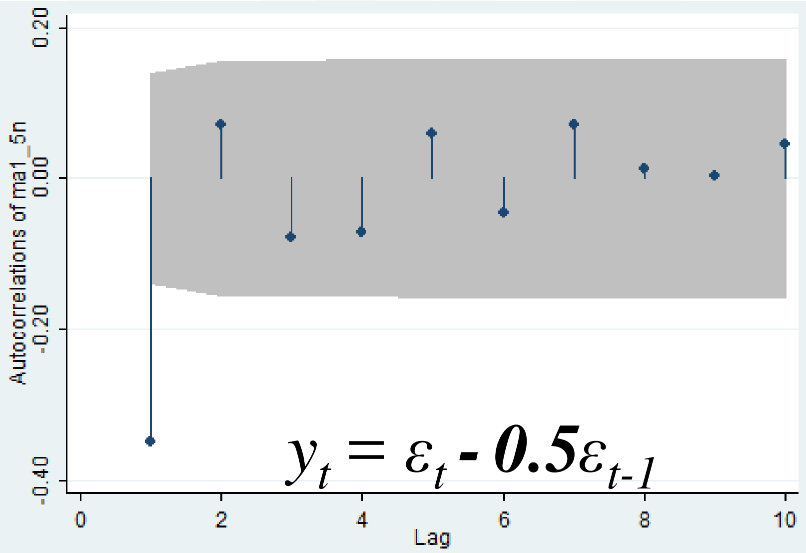
$$\rho(k)=\begin{cases}1,\ k=0\\ \theta/(1+\theta^2),\ k=1\\ 0,\ k>1\end{cases}$$

Поведение ACF: Корреляция только между соседними наблюдениями.

ρ > 0, если $\theta > 0$,

ρ < 0, если $\theta < 0$.





Вывод: *ACF*- Равна нулю после лага 1

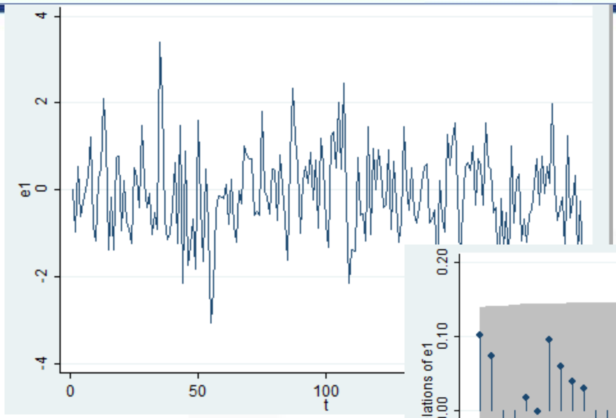
Критерий Бокса-Пирса (1970)

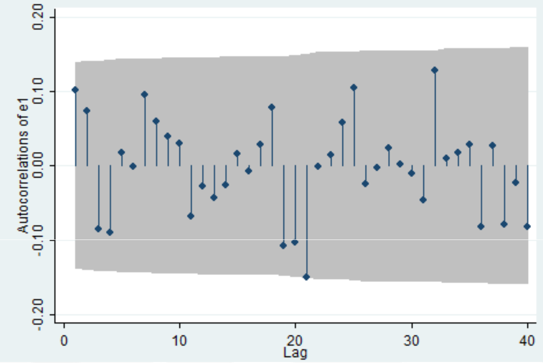
H$_0$: ρ(1)=ρ (2)=ρ (m)=0

ряд является белым шумом
(все автокорреляции до лага m равны 0

Q-статистика Бокса-Пирса:

$$Q = N\sum_{j=1}^{m} \hat{\rho}_j^{\,2} \rightarrow \chi^2_m$$





corrgram eps

LAG	AC	PAC	Q	Prob>Q
1	0.0978	0.0983	.98545	0.3209
2	-0.0279	-0.0388	1.0667	0.5866
3	-0.1341	-0.1326	2.9575	0.3982
4	-0.1212	-0.0987	4.5175	0.3405
5	0.1514	0.1792	6.9785	0.2222
6	0.0063	-0.0566	6.9829	0.3224

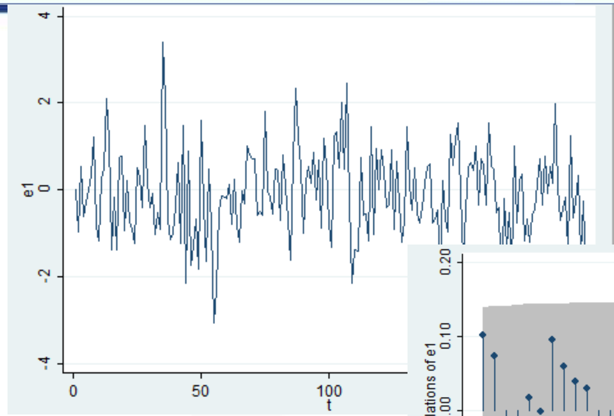
Box-Pierce' Q statistic tests

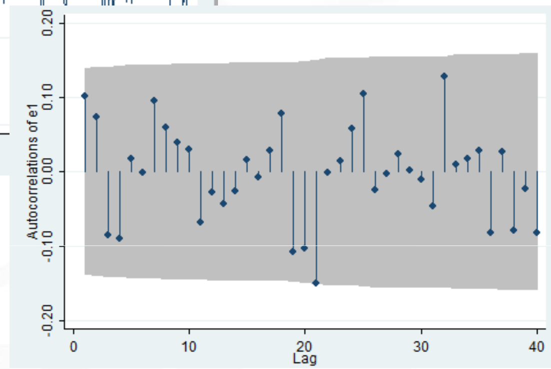
Box, G. E. P., and D. A. Pierce. 1970. Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. Journal of the American Statistical Association 65: 1509–1526.

Критерий Льюинга-Бокса (1978)

H_0: $\rho(1)=\rho(2)=\rho(m)=0$

ряд является белым шумом
(все автокорреляции до лага m равны 0





Q-статистика Льюинга- Бокса :

$$Q = N(N+2)\sum_{j=1}^{m}\frac{1}{n-j}\hat{\rho}_j^{\,2} \to \chi_m^2$$

Достоинства: при малом количестве наблюдений демонстрирует соответствие асимптотическому распределению

Ljung- Box Q statistic tests

```
. wntestq eps

Portmanteau test for white noise
---------------------------------------------
 Portmanteau (Q) statistic =     47.8316
 Prob > chi2(40)           =      0.1847
```

Ljung, G. M., and G. E. P. Box. 1978. On a measure of lack of fit in time series models. Biometrika 65: 297–303.

Задание (самостоятельно). Для процессов

$$y_t = 6 + \varepsilon_t + 0.8\ \varepsilon_{t-1}$$

$$y_t = 6 + \varepsilon_t - 0.8\ \varepsilon_{t-1}\ .$$

Найти

- $E(y_t)$, $V(y_t)$, $\rho(0)$, $\rho(1)$, $\rho(2)$. Как константа влияет на значения характеристик?

-Построить схематично графики ACF.

- Как различаются характеристики процессов?

Модель скользящего среднего MA(2)

Процесс скользящего среднего **2-ого** порядка **MA(2)**:

$$y_t = \theta_0 + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}$$

Числовые характеристики: $E(y_t) = \theta_0$, $V(y_t) = (1+\theta_1^2+\theta_2^2)\sigma^2$

$$\gamma(\tau) = \operatorname{cov}(y_t, y_{t+\tau}) = \begin{cases} (1+\theta_2)\theta_1 \sigma^2, & \tau = 1 \\ \theta_2 \sigma^2, & \tau = 2 \\ 0, & \tau > 2 \end{cases}$$

$$\rho(\tau) = \frac{\gamma(\tau)}{V(y_t)}$$

Вывести!

MA(2): задание

Задание (самостоятельно)

Даны два процесса MA(2):

$$y_t = \theta_0 + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}$$

$$y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}$$

Вычислить и сравнить *ковариации* процессов.

MA(2): пример

Пример MA(2): $y_t = \theta_0 + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}$

Пусть $\theta_1 = -0.5$, $\theta_2 = 0.25$, $T = 200$, рассчитать ACF и PACF по формулам, проверить их статистическую значимость, построить графики. Проверить расчеты в стат.пакете.

Решение:

$$\gamma(\tau) = \operatorname{cov}(y_t, y_{t+\tau}) = \begin{cases} (1+\theta_2)\theta_1 \sigma^2, & \tau = 1 \\ \theta_2 \sigma^2, & \tau = 2 \\ 0, & \tau > 2 \end{cases}$$

$$\rho(\tau) = \frac{\gamma(\tau)}{V(y_t)}$$

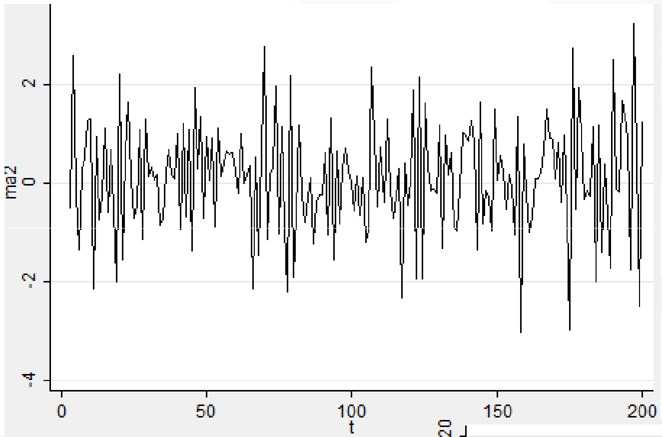
Задание МА(2): $y_t = \theta_0 + \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2}$

Пусть θ_1*=-0.5,* θ_2*=0.25*, рассчитать ACF и PACF по формулам, построить графики. Проверить расчеты в стат.пакете.

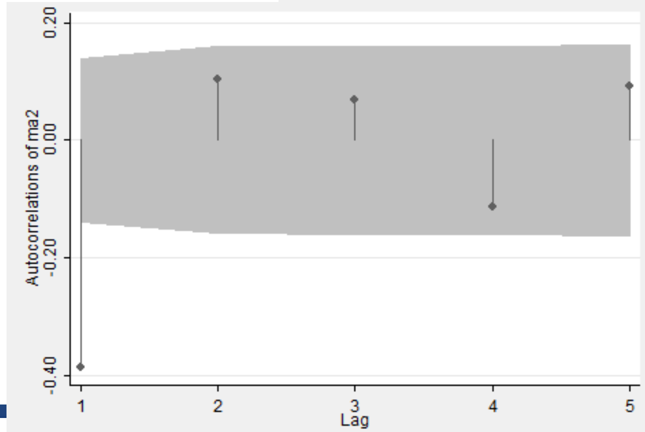
Решение:

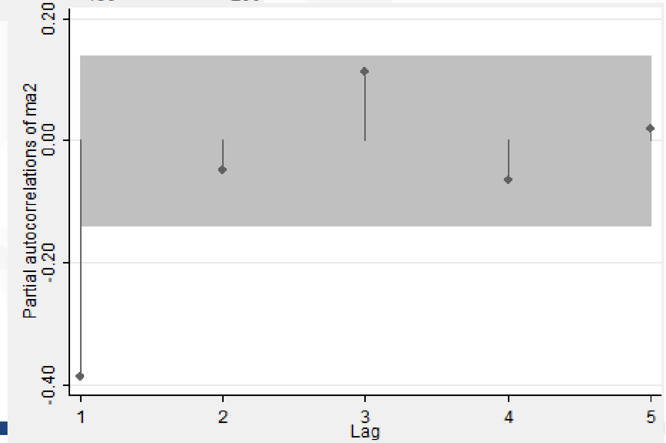
$$\gamma(\tau) = \mathrm{cov}(y_t, y_{t+\tau}) = \begin{cases} (1+\theta_2)\theta_1\sigma^2, & \tau = 0 \\ \theta_2\sigma^2, & \tau = 1 \\ 0, & \tau > 1 \end{cases}$$

$$\rho(\tau) = \frac{\gamma(\tau)}{V(y_t)}$$



LAG	AC	PAC	Q	Prob>Q
1	−0.3851	−0.3872	29.817	0.0000
2	0.1026	−0.0481	31.945	0.0000
3	0.0695	0.1121	32.926	0.0000
4	−0.1124	−0.0639	35.505	0.0000





Модель скользящего среднего MA(q) (Moving Average - MA(q) models)

Процесс **MA(*q*)**

$$y_t = \theta_0 + \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \ldots + \theta_q\varepsilon_{t-q}$$

Математическое ожидание: $E(y_t)=\theta_0$,

Дисперсия:

$$V(y_t) = \sigma^2 \sum_{i=1}^{q} \theta_i^2$$

Автоковариация:

$$\gamma(\tau) = \operatorname{cov}(y_t, y_{t+\tau}) = \begin{cases} \sigma^2 \sum_{i=1}^{q-\tau} \theta_i \theta_{i+\tau}, & \tau = \overline{1,q} \\ 0, & \tau > q \end{cases}$$

Автокорреляция:

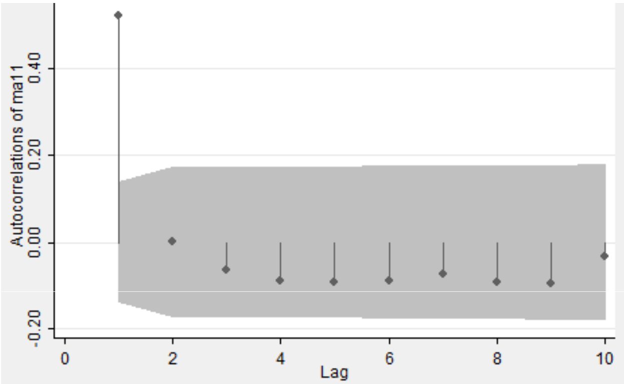
$$\rho(\tau) = \begin{cases} \left(\sum_{i=1}^{q-\tau} \theta_i \theta_{i+\tau} \right) \Big/ \left(\sum_{i=1}^{q} \theta_i^2 \right), & \tau = \overline{1,q} \\ 0, & \tau > q \end{cases}$$

Вывод: Глубина статистической «памяти» процесса определятся q

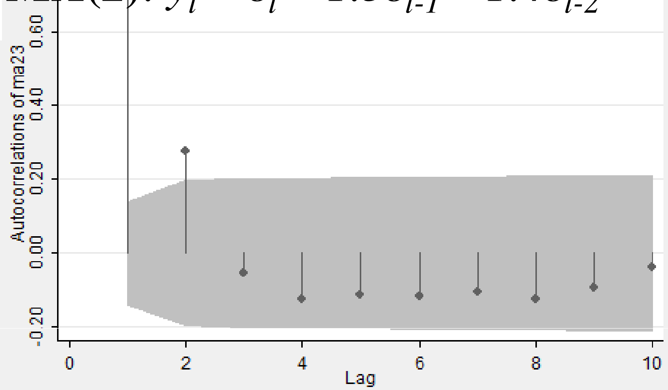
Модель скользящего среднего MA(q) (Moving Average - MA(q) models)

Глубина статистической «памяти» процесса определятся q

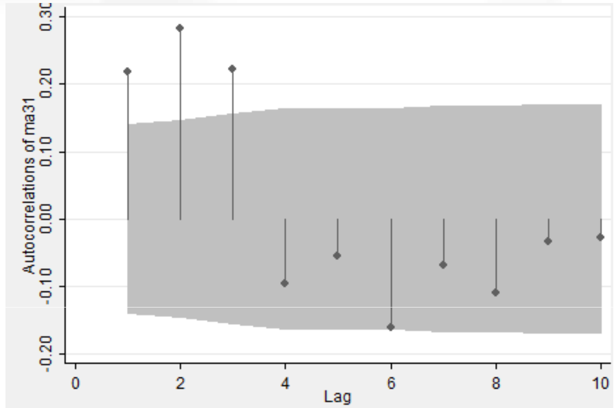
MA(1): $y_t = \varepsilon_t + \varepsilon_{t-1}$



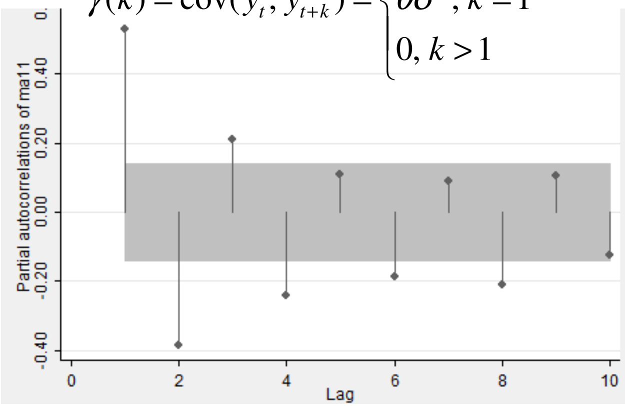
MA(2): $y_t = \varepsilon_t + 1.3\varepsilon_{t-1} + 1.4\varepsilon_{t-2}$



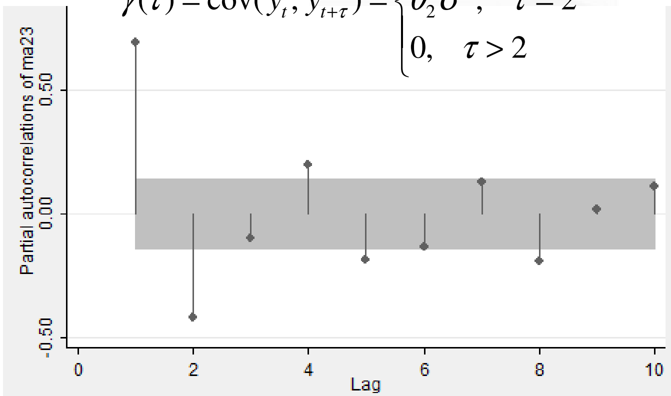
MA(3): $y_t = \varepsilon_t + \varepsilon_{t-1} + 0.1\varepsilon_{t-2} + 3\varepsilon_{t-3}$

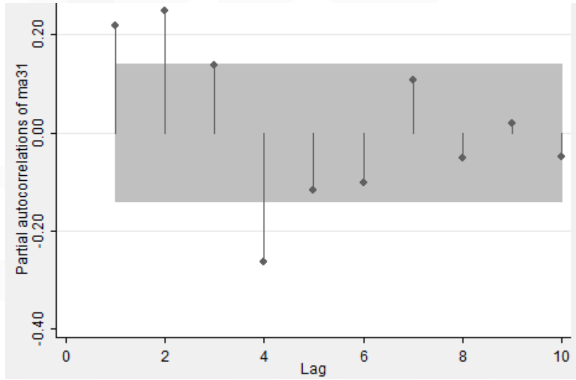


$$\gamma(k) = \operatorname{cov}(y_t, y_{t+k}) = \begin{cases} (1+\theta^2)\sigma^2, \ k=0 \\ \theta\sigma^2, \ k=1 \\ 0, \ k>1 \end{cases}$$



$$\gamma(\tau) = \operatorname{cov}(y_t, y_{t+\tau}) = \begin{cases} (1+\theta_2)\theta_1\sigma^2, \quad \tau=1 \\ \theta_2\sigma^2, \quad \tau=2 \\ 0, \quad \tau>2 \end{cases}$$





Модель скользящего среднего MA(q): альтернативные формы записи

$$y_t = \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \ldots + \theta_q\varepsilon_{t-q}$$

(1) $$y_t = \varepsilon_t + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i}$$

(2) $$y_t = \theta_q(L)\varepsilon_t, \quad \theta_q(L) = \sum_{i=0}^{q} \theta_i L^i = 1 + \theta_1 L + \ldots + \theta_q L^q$$

Свойства лаговых многочленов

$$Lx_t = x_{t-1}, \quad t = 2,...n \qquad L^m x_t = x_{t-m}, \quad m - целое$$

$$\theta_q(L) = \sum_{i=0}^{q} \theta_i L^i = 1 + \theta_1 L + ... + \theta_q L^q$$

Св1. Если существуют 2 операторных полинома, то имеют смысл арифметические операции сложения, вычитания и умножения на число.

Св2. Обратный оператор: $\varphi(L)^{-1}\varphi(L) = 1$

Свойства лаговых многочленов: примеры

Св1. $y_t = \varepsilon_t - 0.5\varepsilon_{t-1} = (1-0.5L)\varepsilon_t = \theta(L)\varepsilon_t,$

$y_t = \varepsilon_t + 0.3\varepsilon_{t-1} = (1+0.3L)\varepsilon_t = \varphi(L)\varepsilon_t,$

$\theta(L) + 2\varphi(L) = \qquad \theta(L) * \varphi(L) =$

Св2. Обратный оператор: 1. $y_t = (1-0.5L)\varepsilon_t = \theta(L)\varepsilon_t \rightarrow \theta^{-1}(L) = ?$

2. $y_t = (1-1.3L+0.4L^2)\varepsilon_t = \theta(L)\varepsilon_t \rightarrow \theta^{-1}(L) =$ '

Свойства лаговых многочленов

$$Lx_t = x_{t-1}, \quad t = 2,...n \qquad L^m x_t = x_{t-m}, \; m - целое \qquad \theta_q(L) = \sum_{i=0}^{q} \theta_i L^i = 1 + \theta_1 L + ... + \theta_q L^q$$

Св3. *Основная теорема алгебры:*

Любой полином степени q с действительными коэффициентами имеет q корней, среди которых м.б. равные по величине, и его можно представить в виде:

$$(1) \quad 1 + \theta_1 L + ... + \theta_q L^q = \theta_q \prod_{i=1}^{q} (L - z_i)$$

где z_i – корни характеристического уравнения

$$1 + \theta_1 z + \theta_2 z^2 + ... + \theta_q z^q = 0$$

$$(2) \quad 1 + \theta_1 L + ... + \theta_q L^q = \prod_{i=1}^{q} (1 - \lambda_i L)$$

где λ_i – корни характеристического уравнения

$$\lambda^q + \theta_1 \lambda^{q-1} + \theta_2 \lambda^{q-2} + ... + \theta_q = 0, \; \lambda_i = 1 / z_i$$

MA(q): пример

Пример $$\mathrm{y}_{\mathrm{t}} = \varepsilon_{\mathrm{t}} - 1.3\varepsilon_{\mathrm{t-1}} + 0.4\,\varepsilon_{\mathrm{t-2}} \rightarrow \mathrm{y}_{\mathrm{t}} = (1 - 1.3\mathrm{L} + 0.4\mathrm{L}^2)\,\varepsilon_{\mathrm{t}}$$

Задание: разложить лаговый многочлен на множители, используя два представления характеристического уравнения.

Разложение 1: $$1 + \theta_1 L + \ldots + \theta_q L^q = \theta_q \prod_{i=1}^{q} (L - z_i)$$

$$1 + \theta_1 z + \theta_2 z^2 + \ldots + \theta_q z^q = 0$$

Разложение 2: $$1 + \theta_1 L + \ldots + \theta_q L^q = \prod_{i=1}^{q} (1 - \lambda_i L),$$

$$\lambda_i = 1/z_i$$

$$\lambda_q + \theta_1 \lambda^{q-1} + \theta_2 \lambda^{q-2} + \ldots + \theta_q = 0,\ \lambda_i = 1/z_i$$

Обратимость (invertibility) MA(q) процесса

Обратимость процесса MA(q) – это возможность его представления в виде AR-процесса. $y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \ldots + \alpha_p y_{t-p} + \varepsilon_t$

Как выразить ε_t через y_t (MA(q) представить в виде AR(∞))?

$$y_t = \theta_q(L)\varepsilon_t, \quad \theta_q(L) = \sum_{i=0}^{q} \theta_i L^i = 1 + \theta_1 L + \ldots + \theta_q L^q \quad \to \varepsilon_t = \left[\theta_q(L)\right]^{-1} y_t$$

$$\left[\theta_q(L)\right]^{-1} = \left[\prod_{i=1}^{q}(1-\lambda_i L)\right]^{-1} = \left[\theta_q \prod_{i=1}^{q}(L - z_i)\right]^{-1}$$

Когда существует обратный оператор?

•***Условие обратимости***: корни характеристического уравнения

(представление1)

$$1 + \theta_1 z + \theta_2 z^2 \ldots + \theta_q z^q = 0, \quad |z_i| > 1$$

(представление2)

$$\lambda^q + \theta_1 \lambda^{q-1} + \theta_2 \lambda^{q-2} \ldots + \theta_q = 0, \quad |\lambda_i| < 1$$

Обратимость (invertibility) MA-процессов: примеры

Пример. 1. **MA(1)**: $y_t = \varepsilon_t - \theta\varepsilon_{t-1}$ $AR(\infty): y_t = -\theta y_{t-1} - \theta^2 y_{t-2} - \ldots + \varepsilon_t$

Условие обратимости MA(1) :|θ|<1.

2. $y_t = (1 - 1.3L + 0.4L^2)\,\varepsilon_t$ Проверить обратимость процесса и представить в виде AR-процесса.

Задание (самостоятельно)

$$y_t = 5 + (1 + 0.5L + 0.1L^2)\,\varepsilon_t$$

1. Вычислить математическое ожидание, дисперсию и ковариации процесса.

2. Разложить лаговый многочлен на множители, используя два представления характеристического уравнения.

3. Проверить обратимость процесса и представить MA в виде AR-процесса.