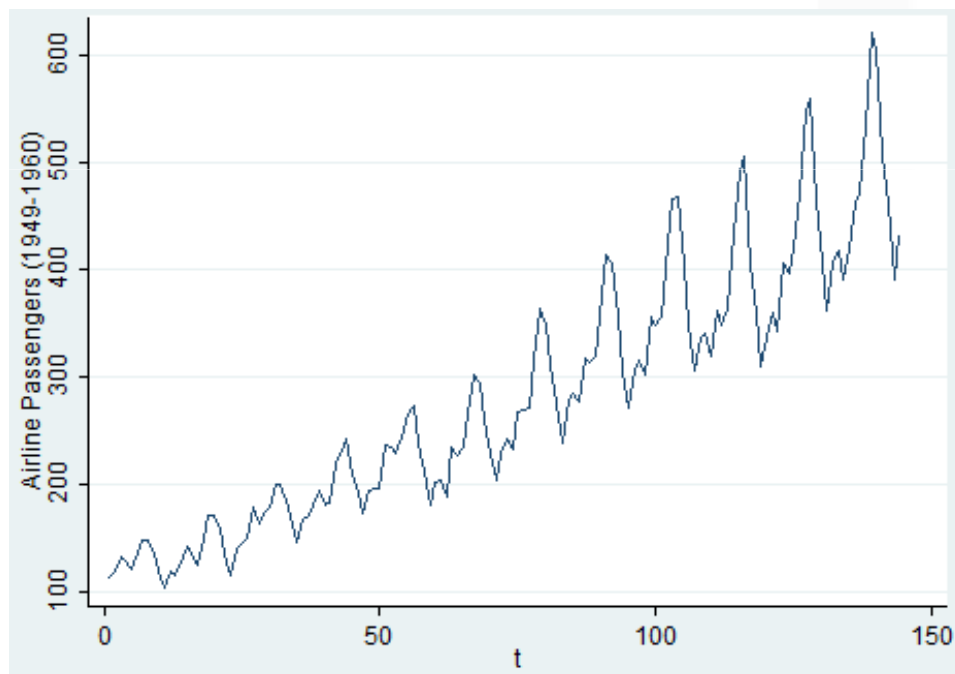


Тема 4. Анализ и моделирование сезонности во временных рядах





Генезис наблюдений, образующих ВР

4 типа факторов (Айвазян):

(А) Долговременные

(Б) Сезонные

(В) Циклические

(Г) Случайные

$$Y_t = \chi(A)f_{mp}(t) + \chi(B)\varphi(t) + \chi(B)\psi(t) + \varepsilon_t$$

Сезонные колебания – это повторяющиеся в каждом временном периоде колебания, связанные с изменением времени года (период колебания <1 года)



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ УПРАВЛЕНИЯ

Сезонность в экономике: примеры

Среднесуточная добыча нефти в России



<http://www.minenergo.gov.ru/activity/oilgas/>

Динамика производства древесины в России в янв 2010- дек2013



Динамика объема добычи нефти, включая газовый конденсат (декабрь 2006 года - 100%)



Источник: Минэкономразвития РФ.



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

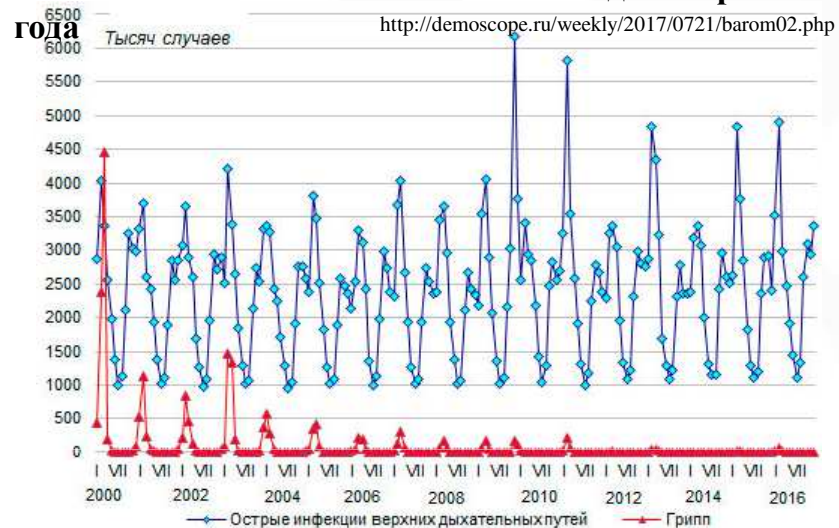
Анализ сезонной компоненты: примеры в демографии



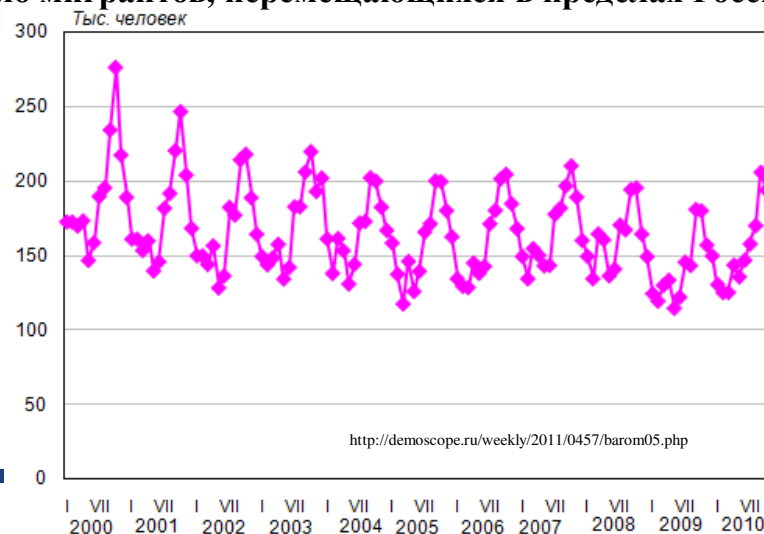
Родионова Л.А.

2019

Заболевания населения ОРЗ в холодное время



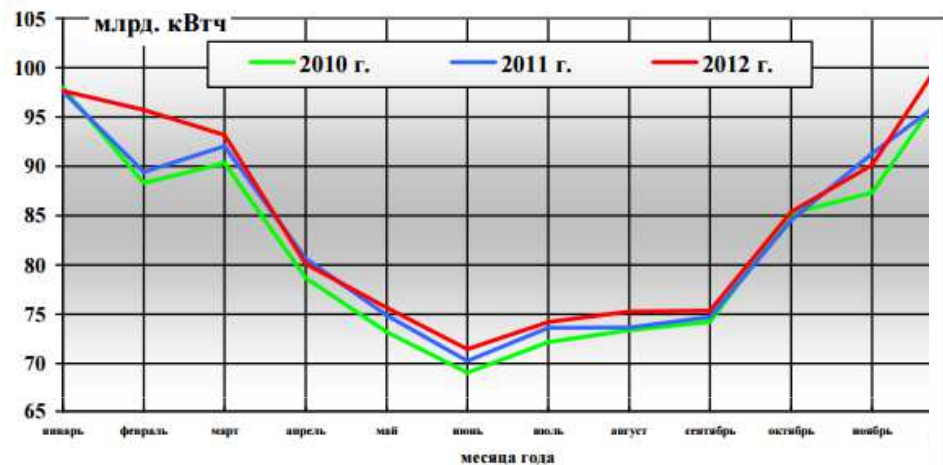
Число мигрантов, перемещающихся в пределах России





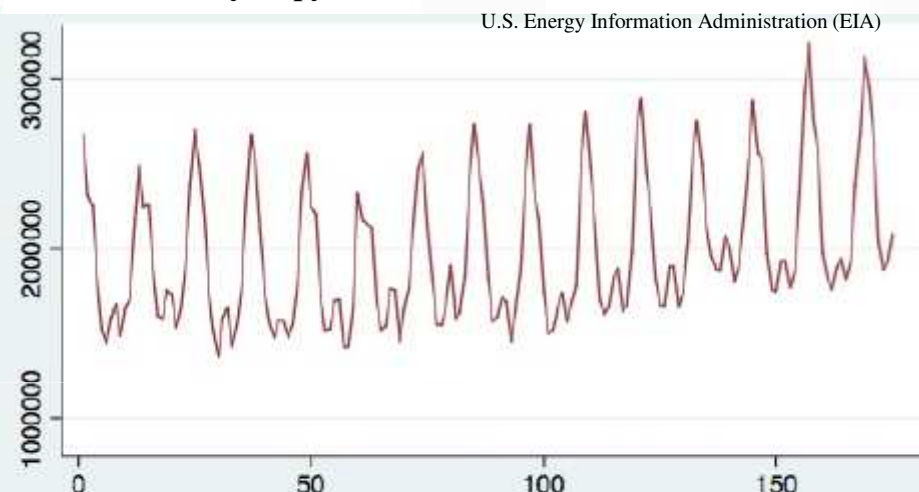
Сезонность в потреблении товаров и услуг: примеры

Динамика потребления электроэнергии в ЕЭС России по мес 2010–2012 гг.



http://so-ups.ru/fileadmin/files/company/reports/disclosure/2013/ues_rep2012.pdf

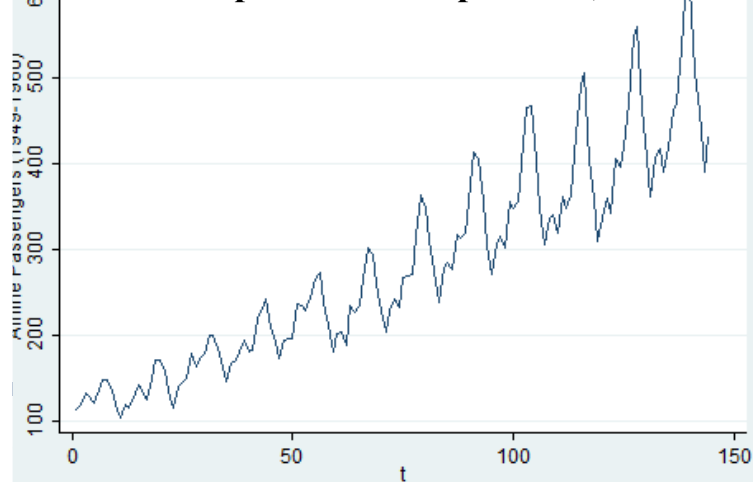
Месячное потребление газа в США, 2001-2015 гг. (в млн. куб. фут)



Пассажирские авиаперевозки в России, млн чел



Пассажирские авиаперевозки, 1949-1960 гг.

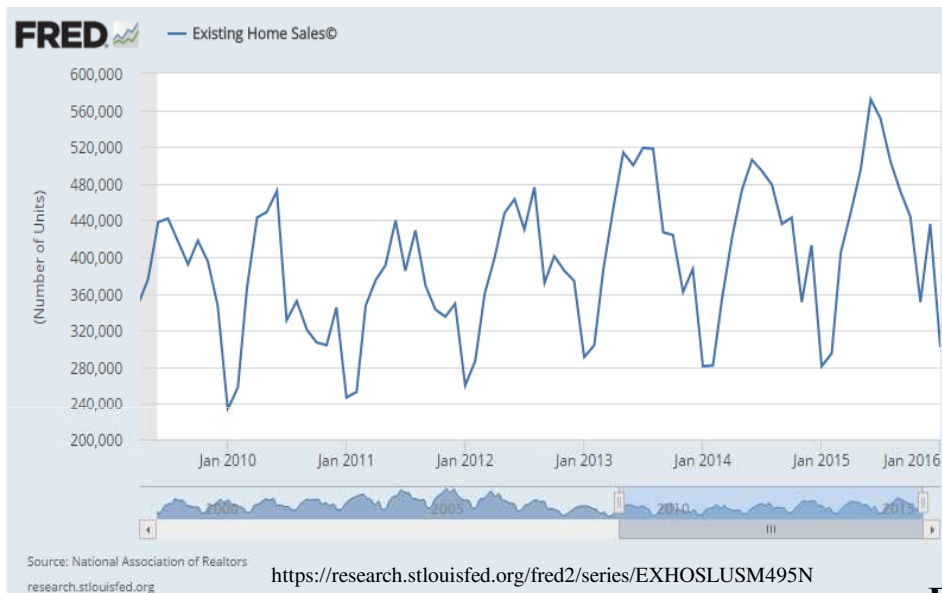




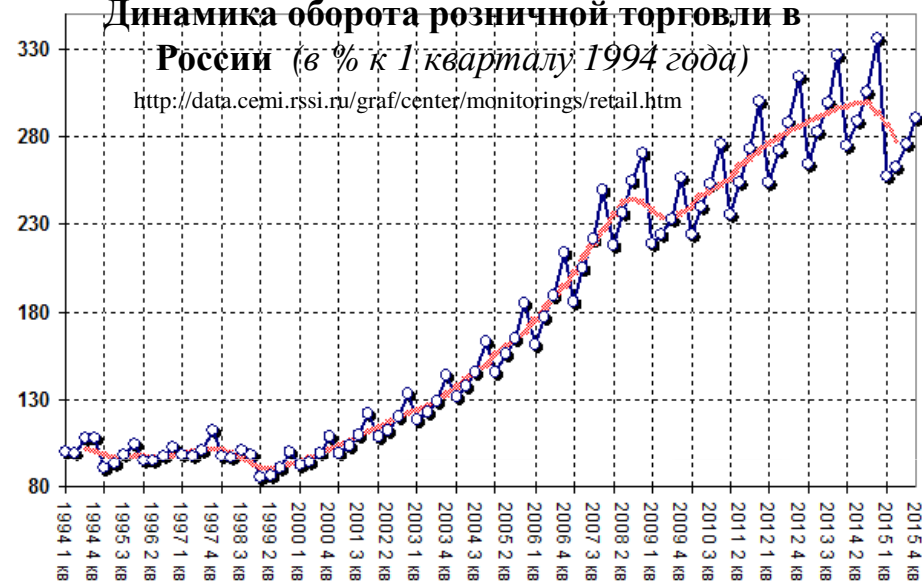
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ЦЕНТР «ЕВРАЗИЙСКАЯ ШКОЛА ЭКОНОМИКИ»
(NRC «Eurasian School of Economics»)

Сезонность в сфере торговли: примеры

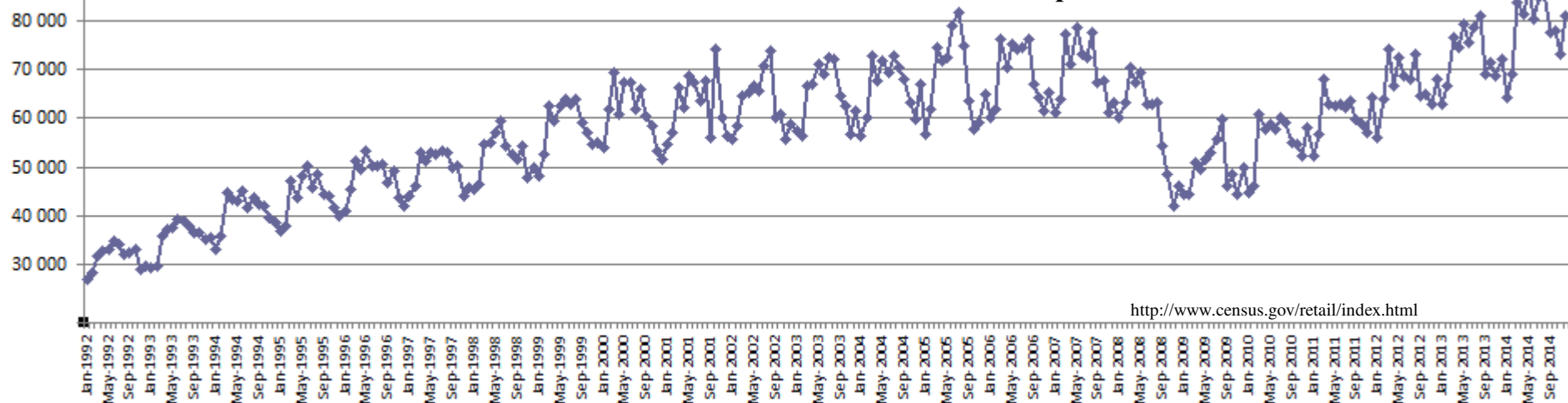
Динамика ежемесячной продажи домов (Сент-Луис)



Динамика оборота розничной торговли в России (в % к 1 кварталу 1994 года)



Ежегодные объемы продаж автомобилей в США

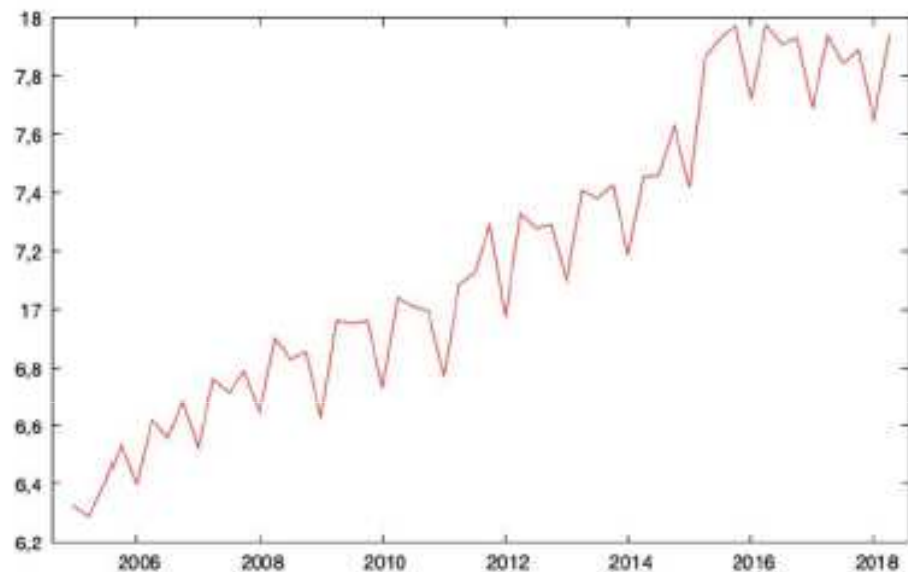




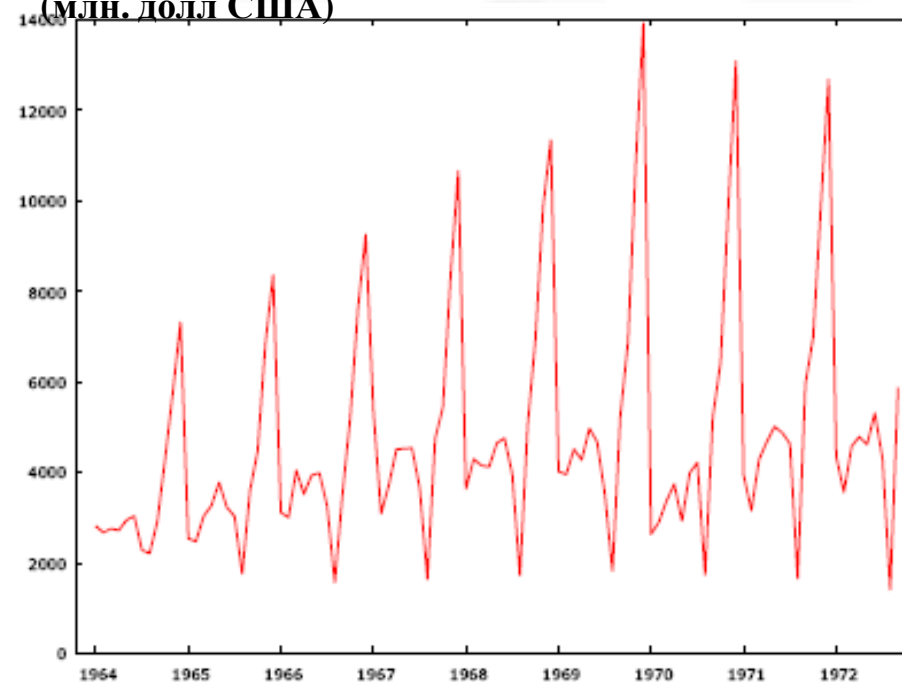
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

Сезонность : примеры

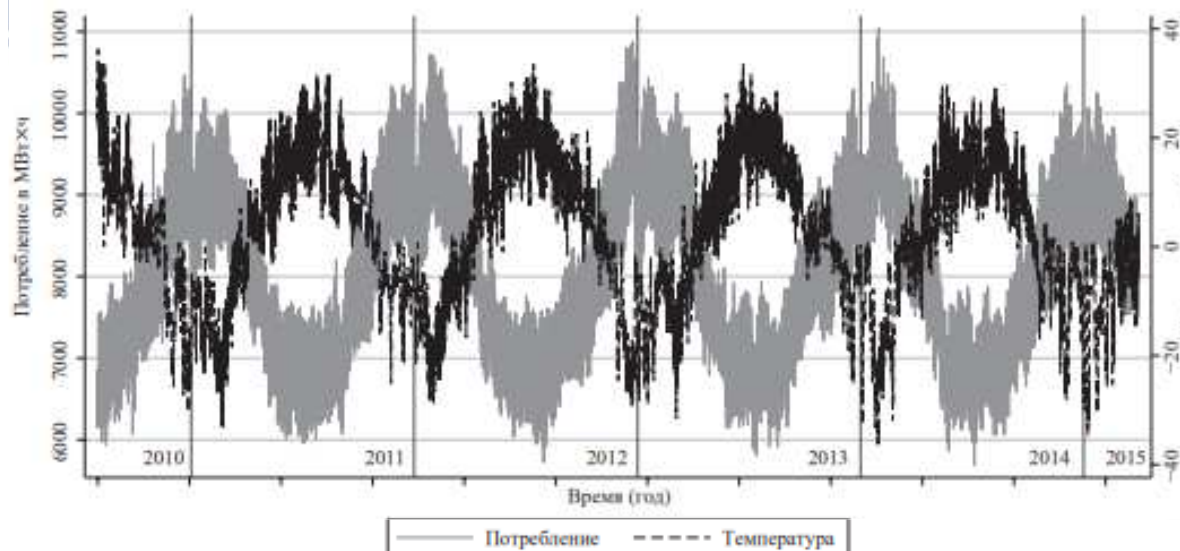
Ежеквартальные поступления по ОСАГО



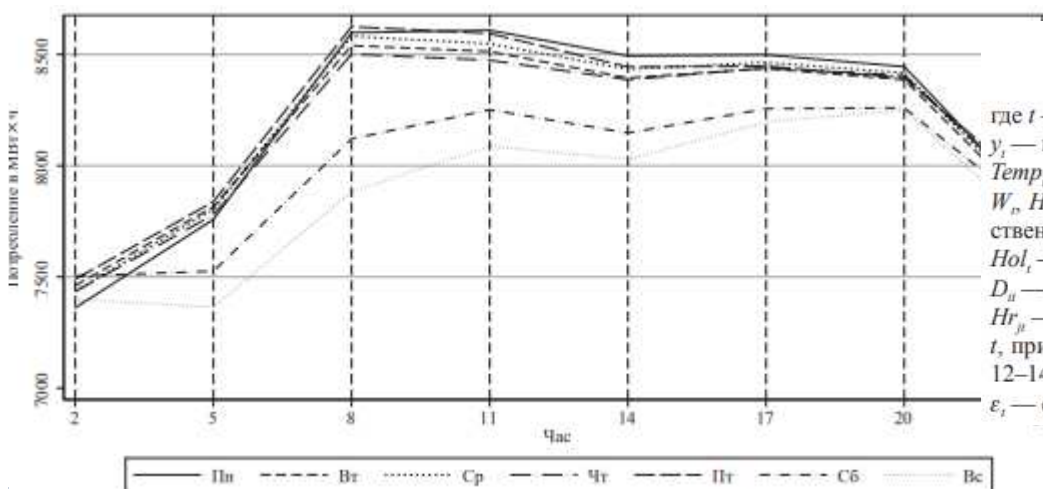
Ежемесячные продажи шампанского Perrin Freres (млн. долл США)



Сезонность : примеры



Динамика потребления электроэнергии и температуры с 2010 по 2015 г. в Пермском крае



$$y_t = \beta_0 + \theta(Temp_t) + \beta_1 W_t + \beta_2 H_t + \beta_3 t + \beta_4 Hol_t + \sum_{i=1}^6 \gamma_i D_{it} + \sum_{j=1}^7 \mu_j Hr_{jt} + \varepsilon_t$$

где t — номер наблюдения, $t = 1, \dots, 13631$;
 y_t — потребление электроэнергии в Пермском крае для наблюдения t ;
 $Temp_t$ — температура окружающей среды для наблюдения t ;
 W_t, H_t — скорость ветра и относительная влажность воздуха для наблюдения t соответственно;
 Hol_t — индикаторная переменная государственных праздников для наблюдения t ;
 D_{it} — индикаторная переменная дня недели i для наблюдения t ;
 Hr_{jt} — индикаторная переменная трехчасового интервала времени суток j для наблюдения t , при этом интервалы времени суток в выборке — это промежутки $\{0-2, 3-5, 6-8, 9-11, 12-14, 15-17, 18-20, 21-23\}$ часов, $j = 1, \dots, 7$;
 ε_t — случайная ошибка.

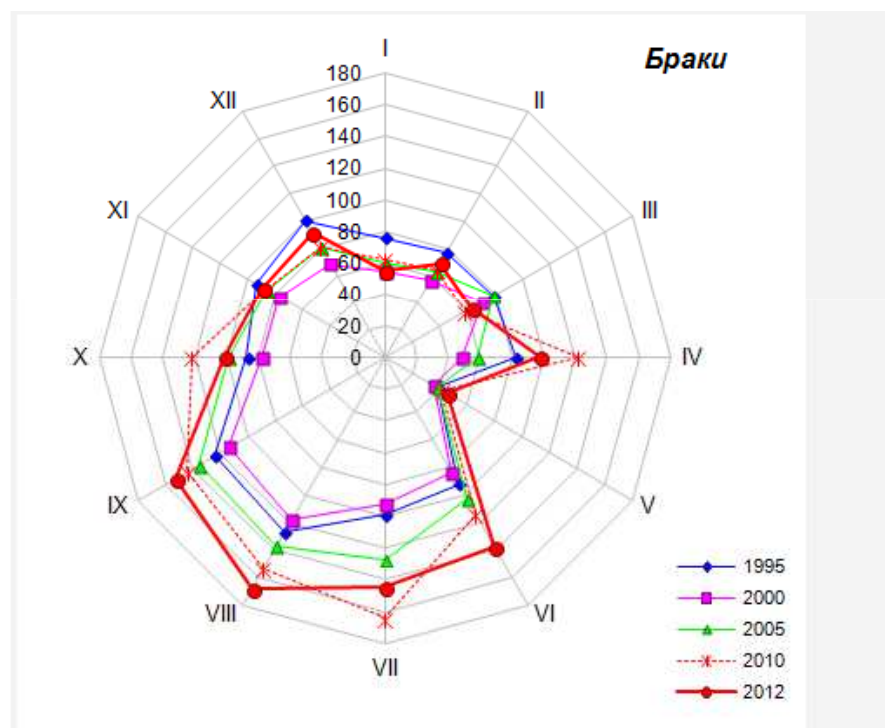


Анализ сезонной компоненты во временных рядах

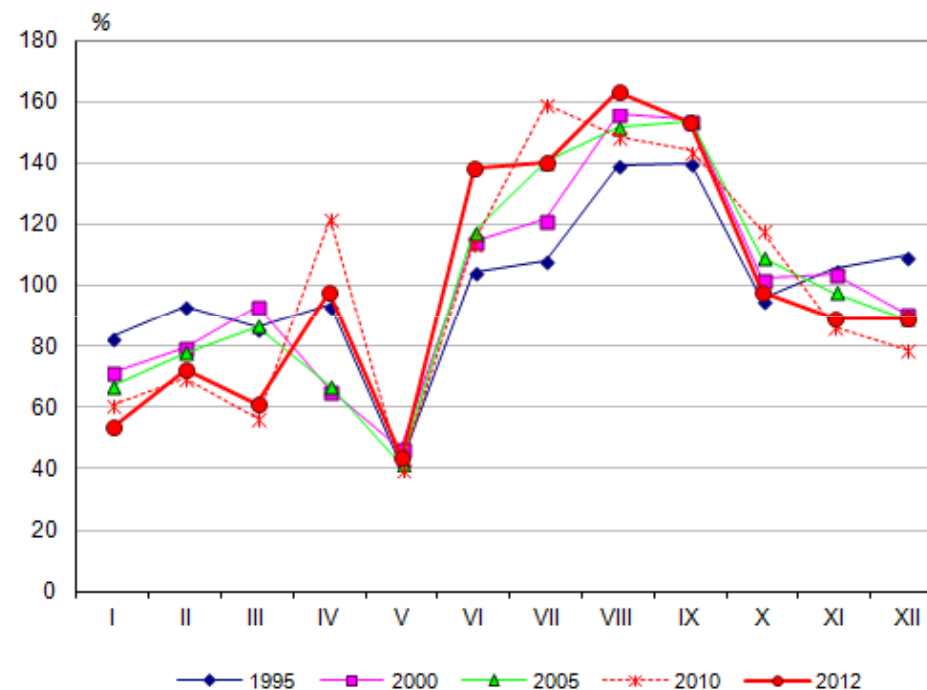
- Графический анализ, сезонная волна
- Анализ автокорреляционной и частной автокорреляционной функции
- Спектральный анализ

Графический анализ сезонных колебаний

Индекс сезонности:
$$I_s = \frac{\tilde{y}_i}{\bar{y}} \cdot 100\%$$



Число браков, зарегистрированных в России, по месяцам 1995, 2000, 2005, 2010 и 2012 годов, тысяч



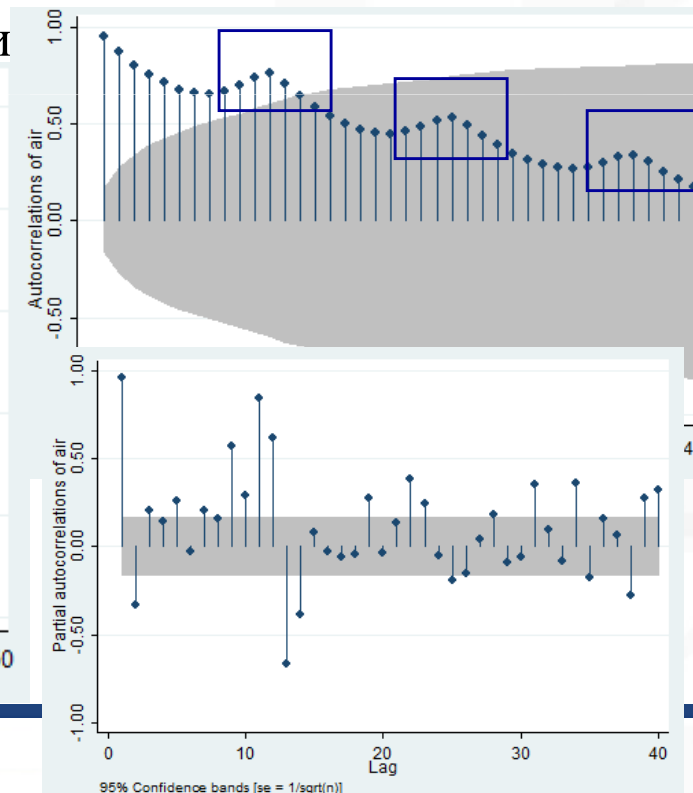
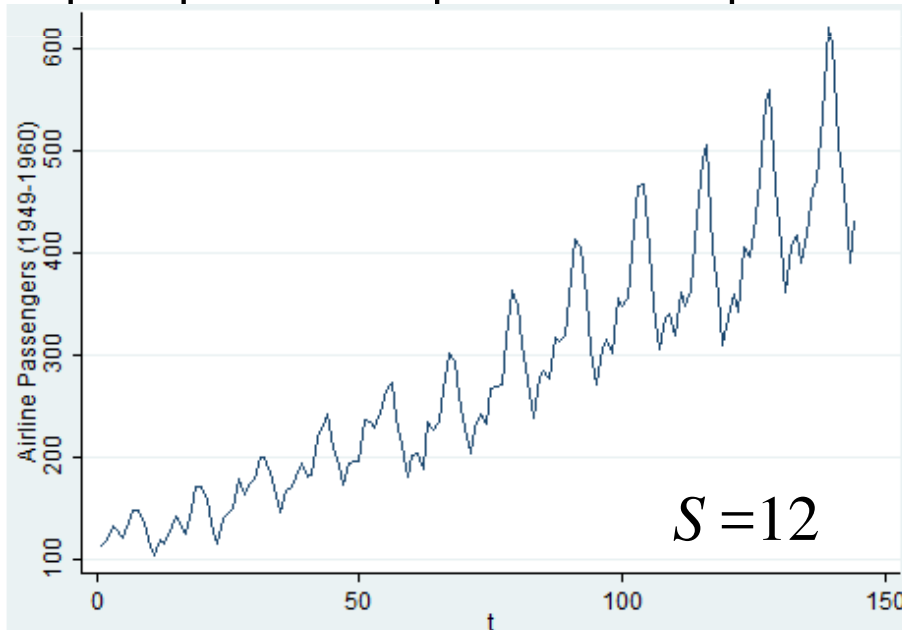
Сезонные отклонения помесичных чисел заключенных браков от среднегодовых значений в России, 1995, 2000, 2005, 2010 и 2012 годы, %

Анализ автокорреляционной и частной автокорреляционной функции

$$\hat{\rho}(\tau) = \frac{\hat{\gamma}(\tau)}{\hat{\gamma}(0)}$$

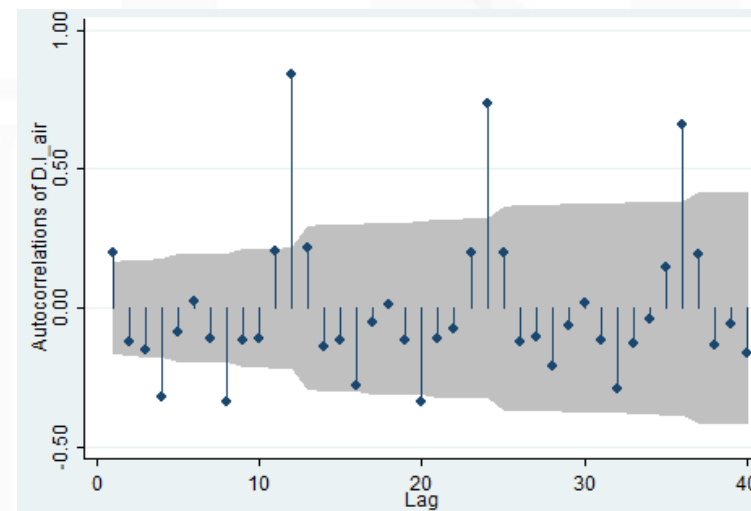
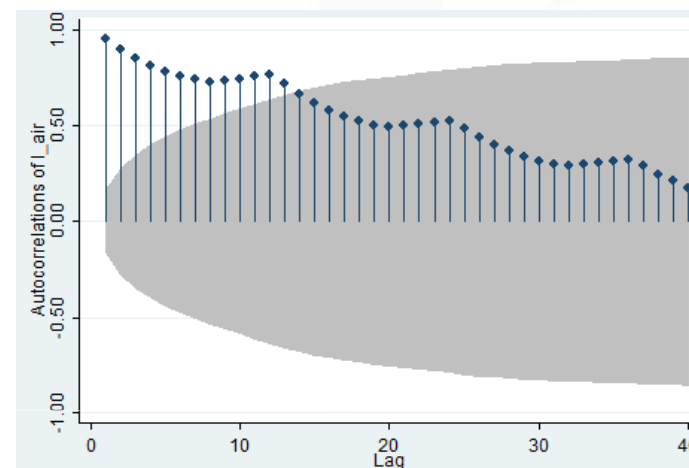
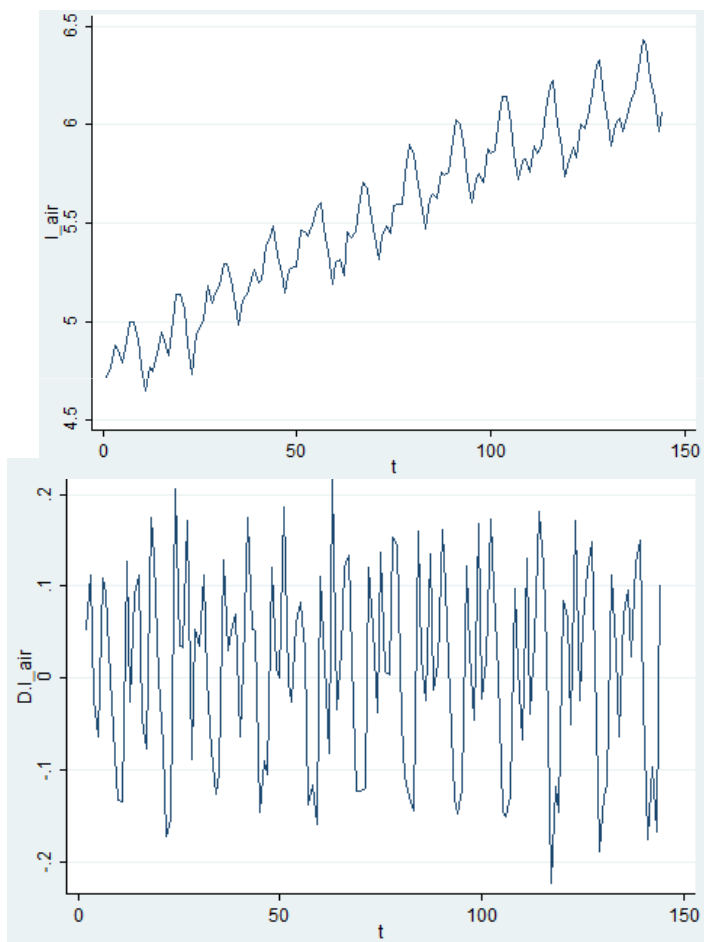
Kendall(1976): периодическая зависимость м.б. определена как корреляционная зависимость пор-ка τ между i -м элементом ряда и $(i + \tau)$ -м элементом.

Пример: Пассажирские авиаперевозки



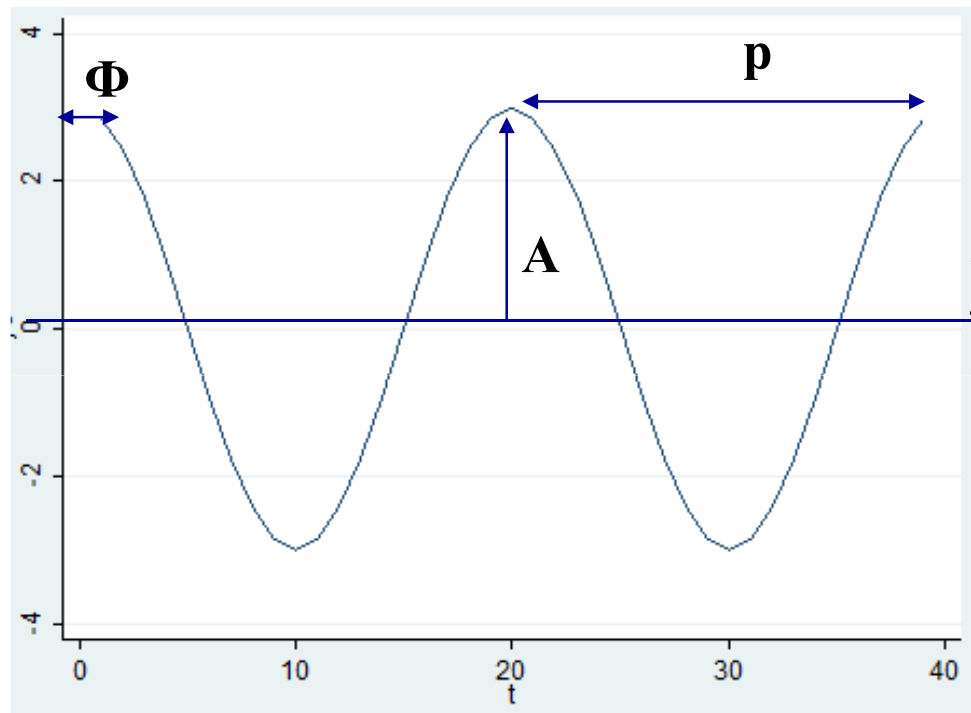
Анализ автокорреляционной и частной автокорреляционной функции

Пример: Пассажирские авиаперевозки



Введение в спектральный анализ

- гармоническое представление



$$y_t = \bar{y} + A \cos(\omega t + \Phi) =$$

$$= \bar{y} + A \cos(2\pi f t + \Phi)$$

Период (p) - минимальный интервал времени, необходимый для того, чтобы значения y_t начали повторяться.

Амплитуда (A) – максимальное смещение от среднего уровня.

Начальная фаза (Φ) – расстояние между началом отсчета ($t=0$) и ближайшим пиковым значением.

Фаза колебания ($\omega t + \Phi$)

Циклическая частота (ω) - на сколько радиан (градусов) изменяется фаза колебания за единицу времени.

Частота (f) - величина, обратная периоду. $f = \frac{1}{p}$ (Число повторений циклов в единицу времени)



Введение в спектральный анализ

$$\begin{aligned}y_t &= \bar{y} + A \cos(2\pi f t + \Phi) = \\&= \bar{y} + A \cos(2\pi f) \cos(\Phi) - A \sin(2\pi f) \sin(\Phi) = \\&= \bar{y} + a \cos(2\pi f) + b \sin(2\pi f)\end{aligned}$$

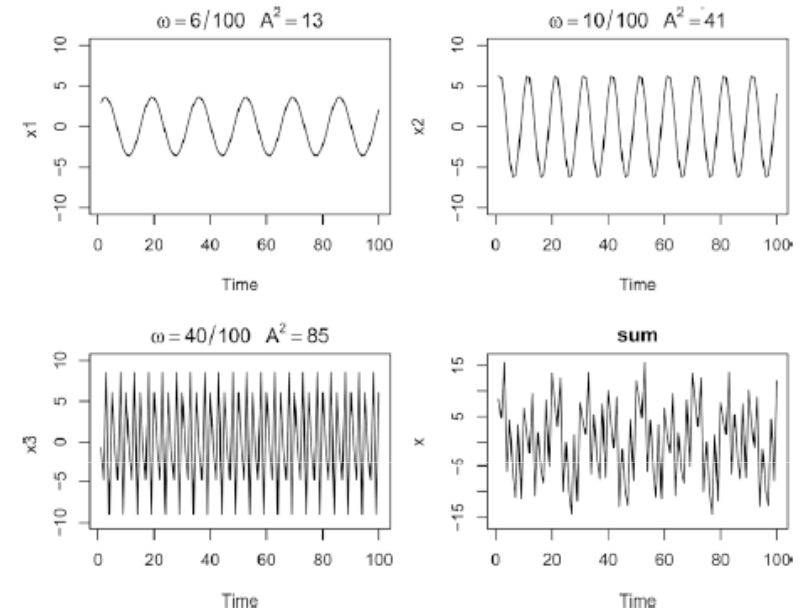
- Разложение в ряд Фурье

$$y_t = a_0 + \sum_{i=1}^q a_i \cos \omega_i t + \sum_{i=1}^q b_i \sin \omega_i t + \varepsilon_t$$

$$\omega_i = 2\pi f_i, \quad f_i = i/N; \quad q = N/2$$

$$МНК: a_0 = \bar{y}, \quad a_i = \frac{2}{N} \sum_{t=1}^N y_t \cos \omega_i t;$$

$$b_i = \frac{2}{N} \sum_{t=1}^N y_t \sin \omega_i t.$$



$$\begin{aligned}x_{t1} &= 2 \cos(2\pi t 6/100) + 3 \sin(2\pi t 6/100) \\x_{t2} &= 4 \cos(2\pi t 10/100) + 5 \sin(2\pi t 10/100) \\x_{t3} &= 6 \cos(2\pi t 40/100) + 7 \sin(2\pi t 40/100)\end{aligned}$$

$$A = \sqrt{a^2 + b^2}$$

$$\Phi = \arctg(-b / a)$$

Фурье



Жан-Батист Жозеф Фурье

(1768-1830) (62г)

фр. Jean-Baptiste Joseph Fourier

Жозеф Фурье был министром продовольствия в правительстве Наполеона, свое открытие он сделал во время Египетской кампании Наполеона (1798-1801), рассчитывал оптимальную глубину винных погребов (в условиях климата Египта французское вино нужно было хранить на другой глубине).

В результате родился метод Фурье для изучения и объяснения механизмов теплопроводности — распространения тепла в твердых телах.

Фурье предположил, что изначальное нерегулярное распределение тепловой волны можно разложить на простейшие синусоиды, каждая из которых будет иметь свой температурный минимум и максимум, а также свою фазу. При этом каждая такая компонента будет измеряться от минимума к максимуму и обратно.

<https://www.hse.ru/news/community/204320600.html>



Введение в спектральный анализ: периодограмма

$$a_0 = \bar{y}, \quad a_i = \frac{2}{N} \sum_{t=1}^N y_t \cos \omega_i t; \quad b_i = \frac{2}{N} \sum_{t=1}^N y_t \sin \omega_i t.$$

Периодограмма:

$$I(f_i) = \frac{N}{2} (a_i^2 + b_i^2)$$

Основная идея: Если ВР содержит циклическую компоненту с частотой f_i , то $I(f_i)$ будет иметь тенденцию к увеличению около f_i



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ ИТМО

Периодограмма: пример

Бокс, Дженкинс

Среднемесечные температуры в центральной Англии в 1964 г.

t	1	2	3	4	5	6	7	8	9	10	11	12
z_t	3,4	4,5	4,3	8,7	13,3	13,8	16,1	15,5	14,1	8,9	7,4	3,6

$$a_i = \frac{2}{N} \sum_{t=1}^N y_t \cos \omega_i t; \quad b_i = \frac{2}{N} \sum_{t=1}^N y_t \sin \omega_i t. \quad I(f_i) = \frac{N}{2} (a_i^2 + b_i^2)$$

Таблица дисперсионного анализа температурного ряда

i	Частота f_i	Период	Периодо- грамма $I(f_i)$
1	$1/12$	12	254,96
2	$1/6$	6	0,19
3	$1/4$	4	1,56
4	$1/3$	3	3,22
5	$5/12$	$12/5$	2,09
6	$1/2$	2	1,08

Амплитуды синусоидальных и косинусоидальных компонент для температурных данных

t	a_i	b_i
1	-5,30	-3,82
2	0,05	0,17
3	0,10	0,50
4	0,52	-0,52
5	0,09	-0,58
6	-0,30	-



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

Анализ периодограммы

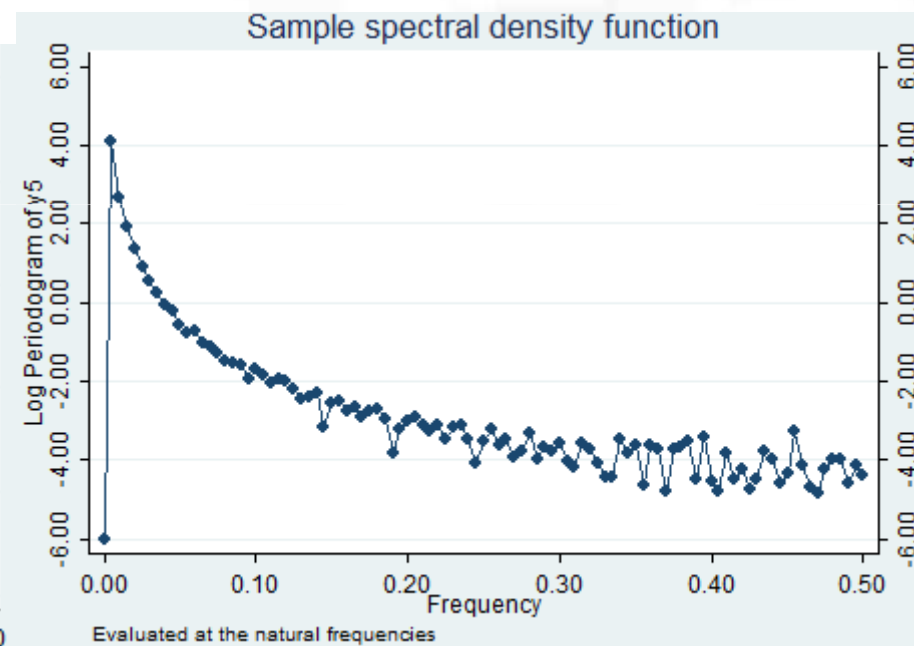
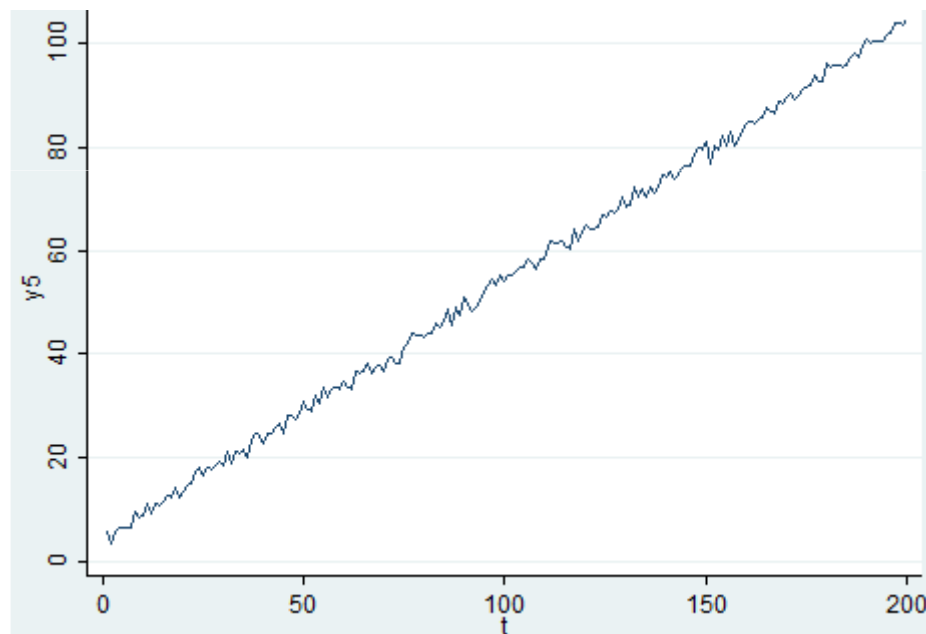
- Анализ пиков периодограммы.
- Тренд: скачок на нулевой частоте
- Большой пик в области частоты w_k указывает на то, что в спектральном разложении присутствует соответствующая *гармоническая компонента*.
- Для случайных данных функция спектральной плотности не имеет доминирующих пиков



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

Анализ периодограммы

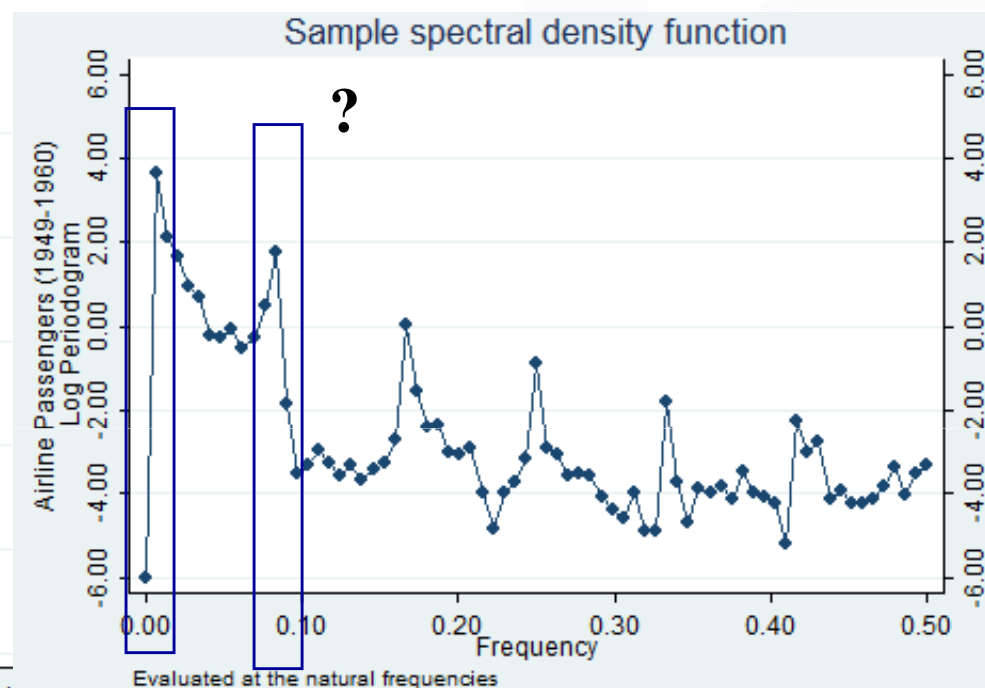
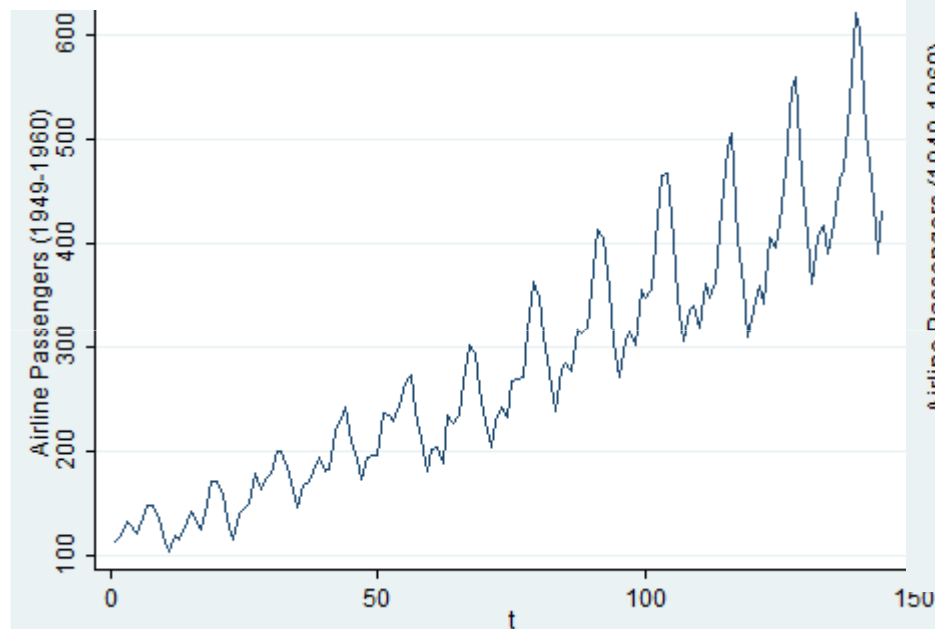
- Тренд: скачок на нулевой частоте



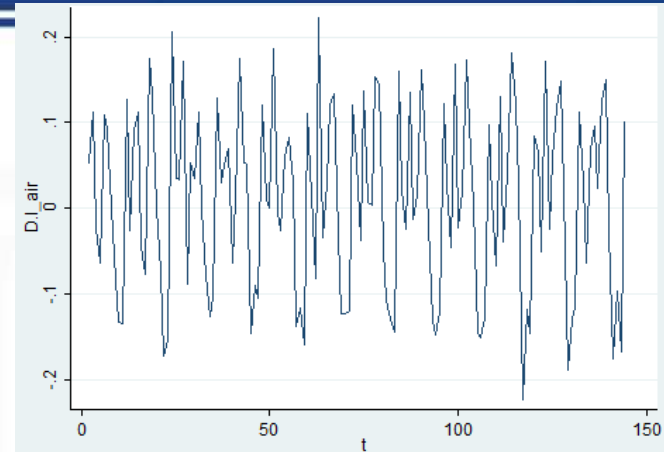
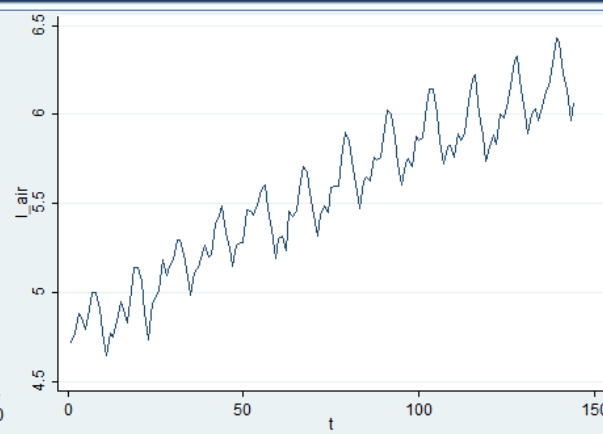
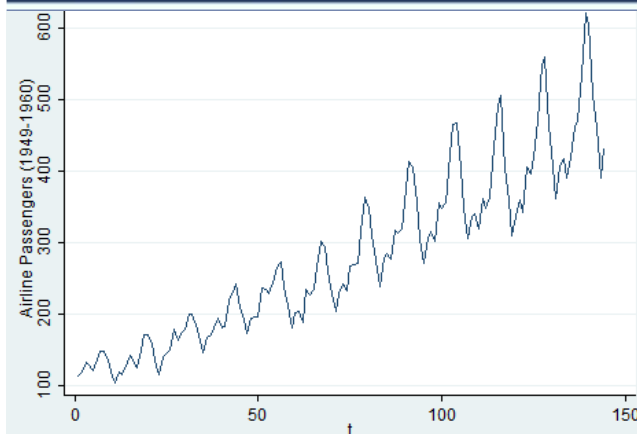


НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ «ВШЭ»

Анализ периодограммы



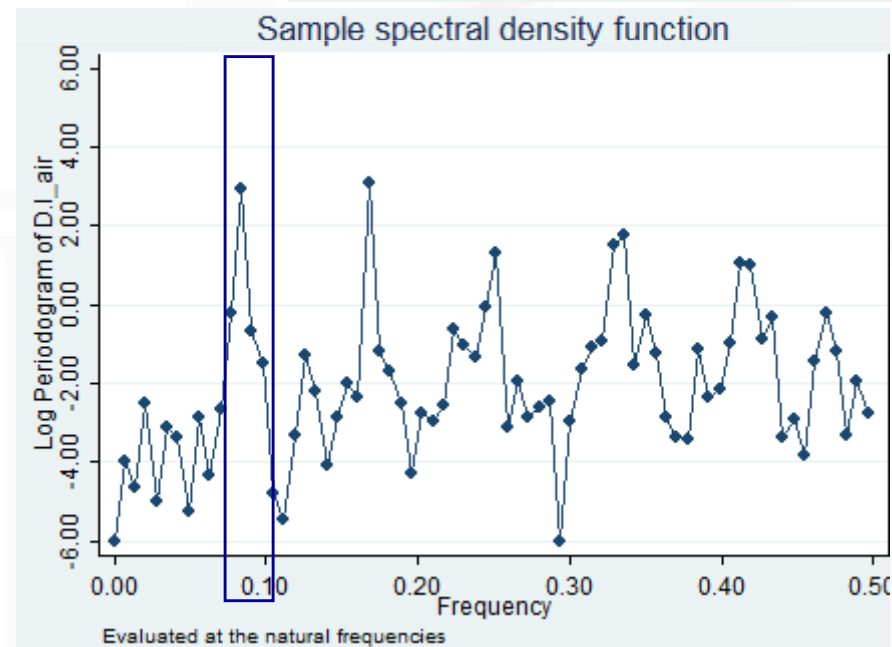
Анализ периодограммы



• период и частота обратны друг другу

Период 12 месяцев соответствует частоте $1 / 12$ (или 0,083). Годовая периодичность подразумевает пик в периодограмме на 0,083 (пик чуть ниже частоты 0.1).

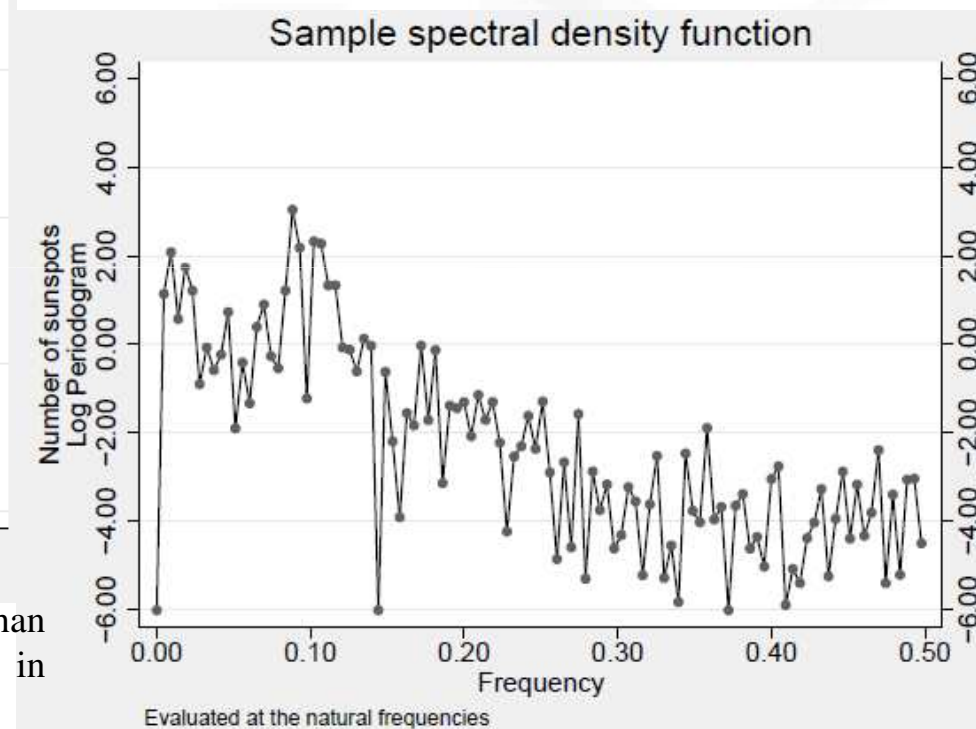
$$f = \frac{1}{p}$$



Анализ периодограммы



The periodogram peaks at a frequency of slightly less than 0.10 cycles per year, indicating a 10- to 12-year cycle in sunspot activity





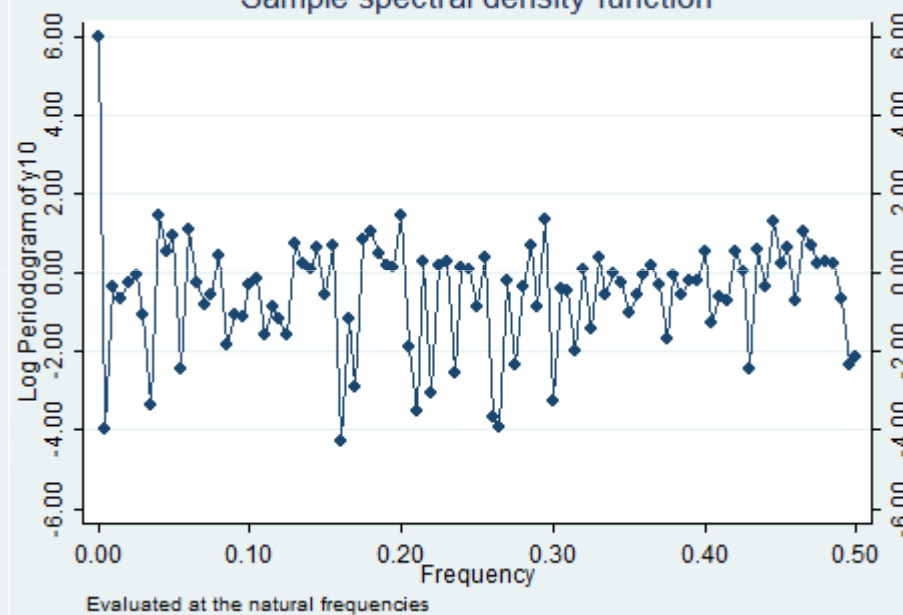
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

Анализ периодограммы

Белый шум



Sample spectral density function





Спектральная плотность

Выборочный спектр ВР – функция, описывающая распределение амплитуд по различным частотам.

$$I(f) = \frac{2}{N} (\alpha_f^2 + \beta_f^2), \quad 0 \leq f \leq 0.5$$

$$I(f) = 2(c_0 + 2 \sum_{k=1}^{N-1} c_k \cos 2\pi f_k)$$

c_k – выборочная автоковариационная функция

Замечание. На практике выборочный спектр сильно флуктуирует.

Спектр мощности (спектральная плотность) (spectral density function)

$$p(f) = \lim_{N \rightarrow \infty} E(I(f)) = 2(\gamma_0 + 2 \sum_{k=1}^{N-1} \gamma_k \cos 2\pi f_k), \quad 0 \leq f \leq 0.5, \quad \lim_{N \rightarrow \infty} E(c_k) = \gamma_k$$

γ_k – теоретическая автоковариационная функция

$$g(f) = p(f) / \sigma_y^2 = 2(1 + 2 \sum_{k=1}^{N-1} \rho_k \cos 2\pi f_k)$$

- Нормированный спектр

Айвазян



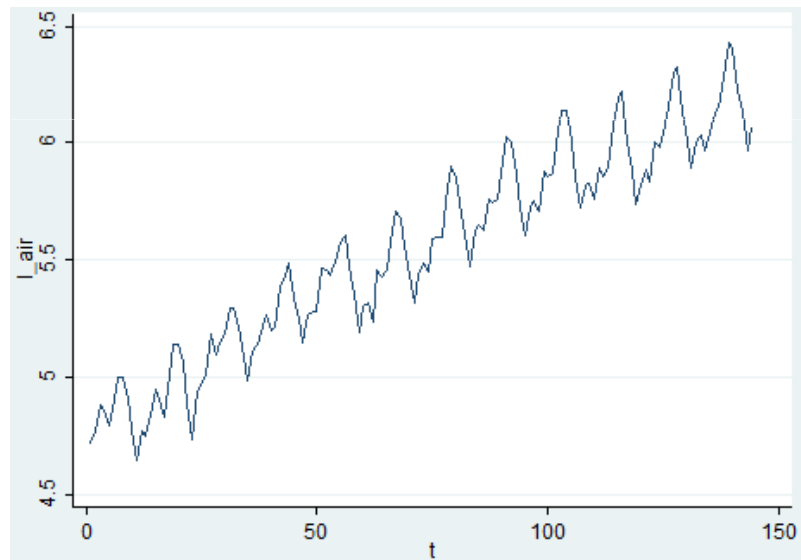
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

4.2. Моделирование сезонных колебаний

- фиктивные переменные
- гармонический анализ
- тренд-сезонные модели
- SARIMA
- адаптивные модели с сезонностью

Аддитивная модель

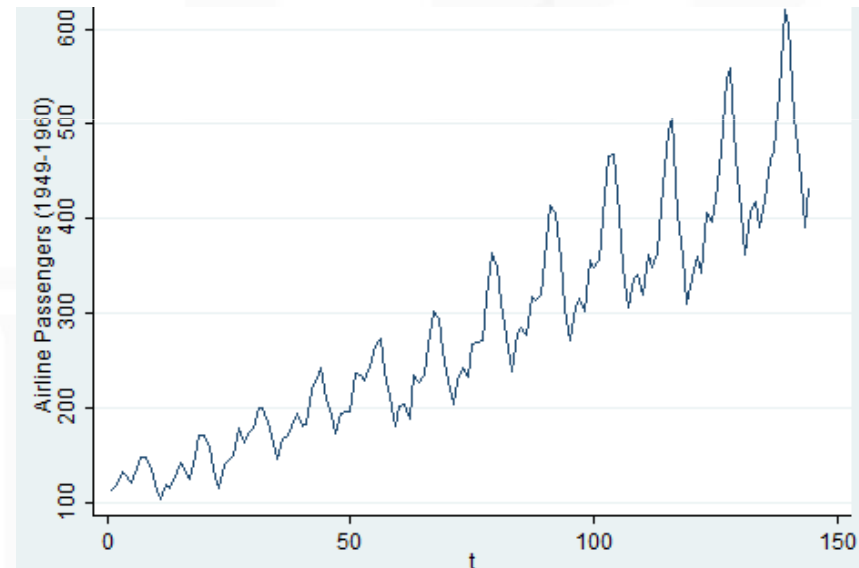
$$Y_t = f_{mp}(t) + \varphi(t) + \varepsilon_t$$



(После логарифмирования)

Мультипликативная модель

$$Y_t = f_{mp}(t) \cdot \varphi(t) \cdot \varepsilon_t$$



4.2.1 Моделирование сезонных колебаний с помощью фиктивных переменных

$$y_t = a + bt + c_2 d_2 + c_3 d_3 + c_4 d_4 + \varepsilon_t,$$

$$d_i = \begin{cases} 1, & \text{если наблюдение } i - \text{му кварталу, } i = 2, 3, 4 \\ 0, & \text{в остальных случаях.} \end{cases}$$

$$\hat{y}_t = \hat{a} + \hat{b}t + \hat{c}_2 d_2 + \hat{c}_3 d_3 + \hat{c}_4 d_4,$$

1-й квартал $\hat{y}_t = \hat{a} + \hat{b}t,$

2-й квартал $\hat{y}_t = \hat{a} + \hat{b}t + \hat{c}_2,$

3-й квартал $\hat{y}_t = \hat{a} + \hat{b}t + \hat{c}_3,$

4-й квартал $\hat{y}_t = \hat{a} + \hat{b}t + \hat{c}_4.$

- оценка сезонных отклонений

$$\hat{y}_t = \hat{a}_{\text{средн}} + \hat{b}t$$



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ УПРАВЛЕНИЯ

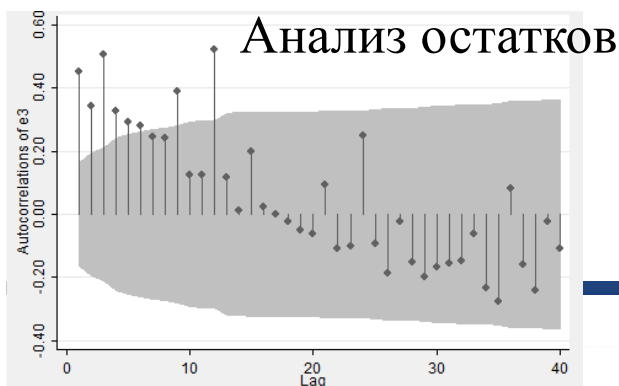
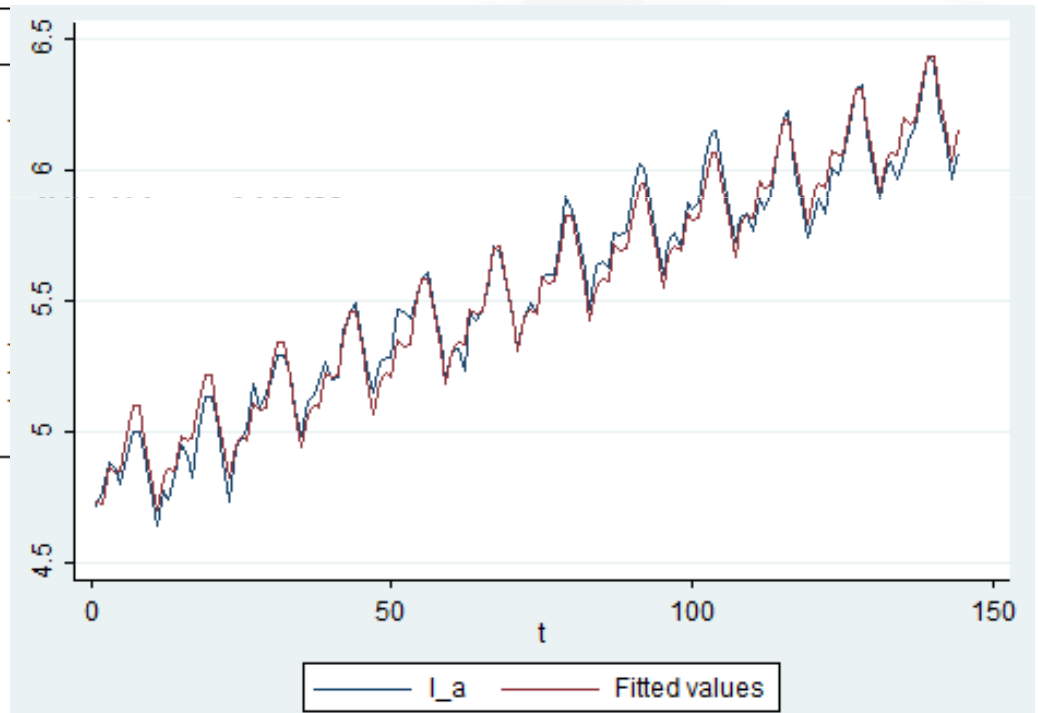
Моделирование сезонных колебаний с помощью фиктивных переменных

reg l_a t mon_2 mon_3 mon_4 mon_5 mon_6 mon_7 mon_8 mon_9 mon_10 mon_11 mon_12

Source	SS	df	MS
Model	27.407665	12	2.28397208
Residual	.460715454	131	.003516912
Total	27.8683804	143	.194883779

Number of obs = 144
F(12, 131) = 649.43
Prob > F = 0.0000
R-squared = 0.9835
Adj R-squared = 0.9820
Root MSE = .0593

l_a	Coef.	Std. Err.	t	P> t
t	.0100688	.0001193	84.40	0.000
mon_2	-.0220548	.0242109	-0.91	0.364
mon_3	.1081723	.0242118	4.47	0.000
mon_4	.0769034	.0242132	3.18	0.002
mon_5	.0745308	.0242153	3.08	0.003
mon_6	.196677	.0242179	8.12	0.000
mon_7	.3006193	.0242212	12.41	0.000
mon_8	.2913245	.024225	12.03	0.000
mon_9	.1466899	.0242294	6.05	0.000
mon_10	.0085316	.0242344	0.35	0.725
mon_11	-.1351861	.02424	-5.58	0.000
mon_12	-.0213211	.0242461	-0.88	0.381
_cons	4.72678	.0188935	250.18	0.000





4.2.2. Моделирование сезонных колебаний с помощью гармонического анализа

$$y_t = a_0 + \sum_{i=1}^q a_i \cos \omega_i t + \sum_{i=1}^q b_i \sin \omega_i t + \varepsilon_t,$$

$$y_t = a_0 + a_1 \cos \omega_1 t + b_1 \sin \omega_1 t + a_2 \cos \omega_2 t + b_2 \sin \omega_2 t + \varepsilon_t.$$

$$\omega_i = 2\pi f_i, \quad f_i = i/N.$$



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ЦЕНТР ИНФОРМАТИКИ

Моделирование сезонных колебаний с помощью гармонического анализа

Source	SS	df	MS	Number of obs =	144
Model	26.7467746	3	8.91559154	F(3, 140) =	1112.85
Residual	1.12160577	140	.00801147	Prob > F	= 0.0000
				R-squared	= 0.9598
				Adj R-squared	= 0.9589
Total	27.8683804	143	.194883779	Root MSE	= .08951

l_a	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
t	.010036	.0001798	55.81	0.000	.0096804	.0103915
cos11	-.1417735	.01055	-13.44	0.000	-.1626314	-.1209155
sin11	-.0494811	.0105698	-4.68	0.000	-.0703782	-.028584
_cons	4.814568	.0150203	320.54	0.000	4.784872	4.844264

Source	SS	df	MS	Number of obs =	144
Model	27.2297386	5	5.44594771	F(5, 138) =	1176.78
Residual	.638641843	138	.004627839	Prob > F	= 0.0000
				R-squared	= 0.9771
				Adj R-squared	= 0.9763
Total	27.8683804	143	.194883779	Root MSE	= .06803

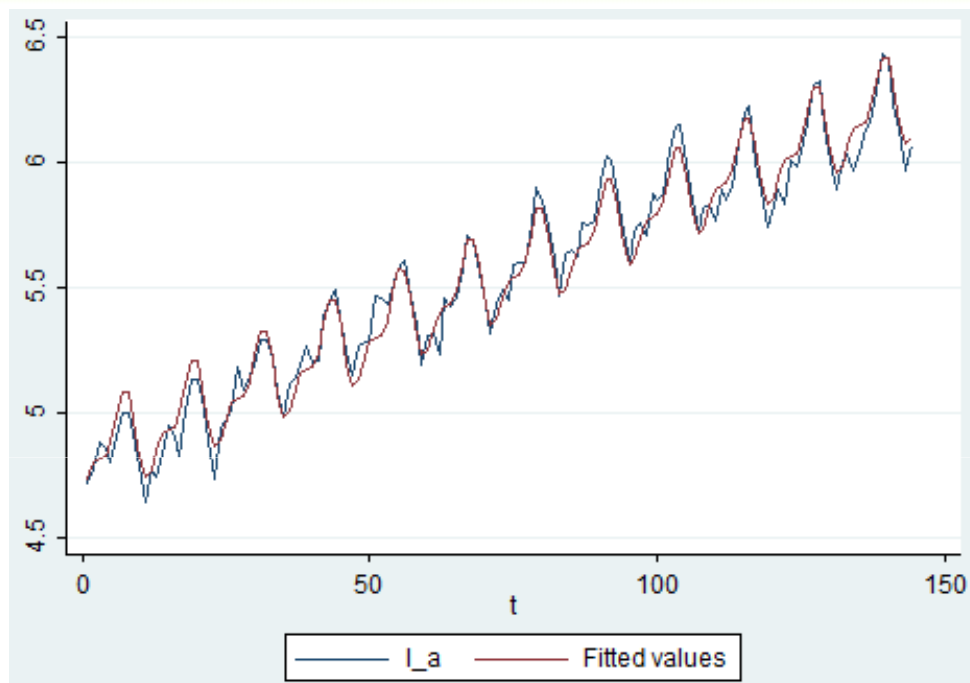
l_a	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
t	.0100822	.0001368	73.73	0.000	.0098118	.0103526
cos11	-.1418197	.0080184	-17.69	0.000	-.1576745	-.125965
sin11	-.0493085	.0080334	-6.14	0.000	-.065193	-.033424
cos2	-.0228262	.0080184	-2.85	0.005	-.0386809	-.0069714
sin2	.078703	.0080207	9.81	0.000	.0628436	.0945623
_cons	4.811215	.0114209	421.26	0.000	4.788632	4.833798

	cos11	sin11	cos2	sin2
1	.8660254	.5	.5	.8660254
2	.5	.8660254	-.5	.8660254
3	6.12e-17	1	-1	1.22e-16
4	-.5	.8660254	-.5	-.8660254
5	-.8660254	.5	.5	-.8660254
6	-1	1.22e-16	1	-2.45e-16
7	-.8660254	-.5	.5	.8660254
8	-.5	-.8660254	-.5	.8660254
9	-1.84e-16	-1	-1	3.67e-16
10	.5	-.8660254	-.5	-.8660254
11	.8660254	-.5	.5	-.8660254
12	1	-2.45e-16	1	-4.90e-16
13	.8660254	.5	.5	.8660254
14	.5	.8660254	-.5	.8660254
15	1.19e-15	1	-1	2.39e-15
16	-.5	.8660254	-.5	-.8660254
17	-.8660254	.5	.5	-.8660254
18	-1	3.67e-16	1	-7.35e-16
19	-.8660254	-.5	.5	.8660254

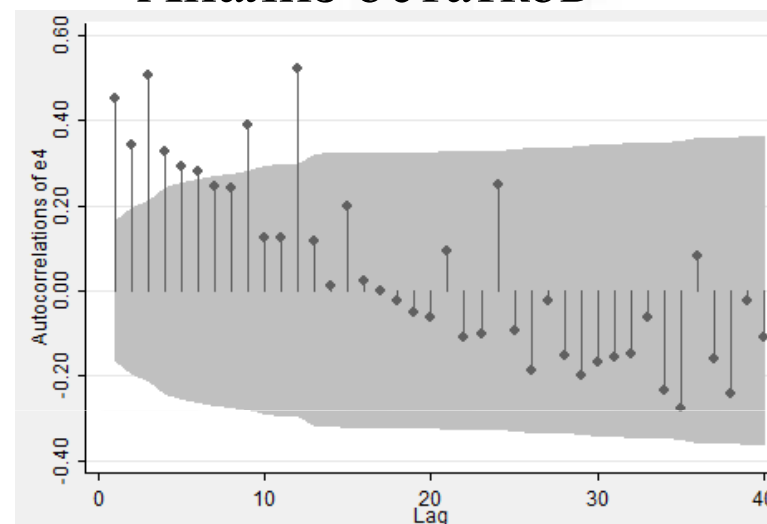


НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ УПРАВЛЕНИЯ

Моделирование сезонных колебаний с помощью гармонического анализа



Анализ остатков



```
. wrtestq e4
```

Portmanteau test for white noise

Portmanteau (Q) statistic =	320.6684
Prob > chi2(40) =	0.0000

4.2.3. Моделирование сезонных колебаний: тренд-сезонные модели (самостоятельно)

Процедура построения тренд-сезонных моделей:

1. Сглаживание ВР с помощью скользящей средней
2. Оценивание сезонной составляющей с учетом характера сезонности (аддитивной или мультипликативной).
3. Сезонная корректировка (десезонализация) исходных данных.
4. Расчет параметров трендовой составляющей.
5. Моделирование динамики исходного ряда с учетом трендовой и сезонной составляющих.
6. Оценка точности и адекватности полученной модели.
7. Использование построенной модели для прогнозирования.

См. Дуброва Т.А. Статистические методы прогнозирования 2003

<http://www.twirpx.com/file/18580/>

4.2.4. Сезонные модели ARIMA

1. Тренд-стационарные ВР (с трендом определенного вида)

$$y_t = \varphi(t) f(t) + \alpha(L) y_t + \theta(L) \varepsilon_t + c_2 d_2 + c_3 d_3 + c_4 d_4$$

Для квартальных данных

2. Сезонно-интегрированные ВР – приводятся к стационарному виду взятием сезонной разности

$$y_t = y_{t-4} + \varepsilon_t$$

Сезонный разностный оператор:

$$\Delta^s y_t = (1 - L^s) y_t = y_t - y_{t-s}$$

Сезонная разность

$$\Delta^s y_t = y_t - y_{t-s}$$

$$\Delta^4 y_t = y_t - y_{t-4}$$

$$\Delta^{12} y_t = y_t - y_{t-12}$$

3. SARIMA(p, d, q)(Ps, Ds, Qs)_s

Модели SARIMA: общий вид

SARIMA(p, d, q)(P_s, D_s, Q_s)_s,

P_s - сезонный параметр авторегрессии;

Q_s - сезонный параметр скользящего среднего;

D_s - порядок сезонной разности.

$$\alpha_p(L)\alpha_{P_s}(L)\Delta^d\Delta_s^D y_t = \theta_q(L)\theta_{Q_s}(L)\varepsilon_t$$

$$\alpha_{P_s}(L) = 1 - \alpha_{s1}L^s - \alpha_{s2}L^{2s} \dots - \alpha_{sP}L^{P_s},$$

$$\theta_{Q_s}(L) = 1 + \theta_{s1}L^s + \theta_{s2}L^{2s} \dots + \theta_{sQ}L^{Q_s}$$

мультипликативная модель



Модели SARIMA:

аддитивная и мультипликативная модель

$$SARIMA(0,1,1)(0,1,1)_4$$

$$\Delta\Delta_4 y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_4 \varepsilon_{t-4}$$

$$= (1 + \theta_1 L + \theta_4 L^4) \varepsilon_t$$

$$\Delta\Delta_4 y_t = (1 + \theta_1 L)(1 + \theta_4 L^4) \varepsilon_t$$

$$SARIMA(1,1,0)(1,1,0)_4$$

$$(1 + \theta_1 L + \theta_4 L^4) \Delta\Delta_4 y_t = \varepsilon_t$$

$$(1 + \theta_1 L)(1 + \theta_4 L^4) \Delta\Delta_4 y_t = \varepsilon_t$$

Задание. Записать (самостоятельно) аддитивный и мультипликативный вид моделей:

$$SARIMA(2,1,1)(1,1,2)_{12}$$

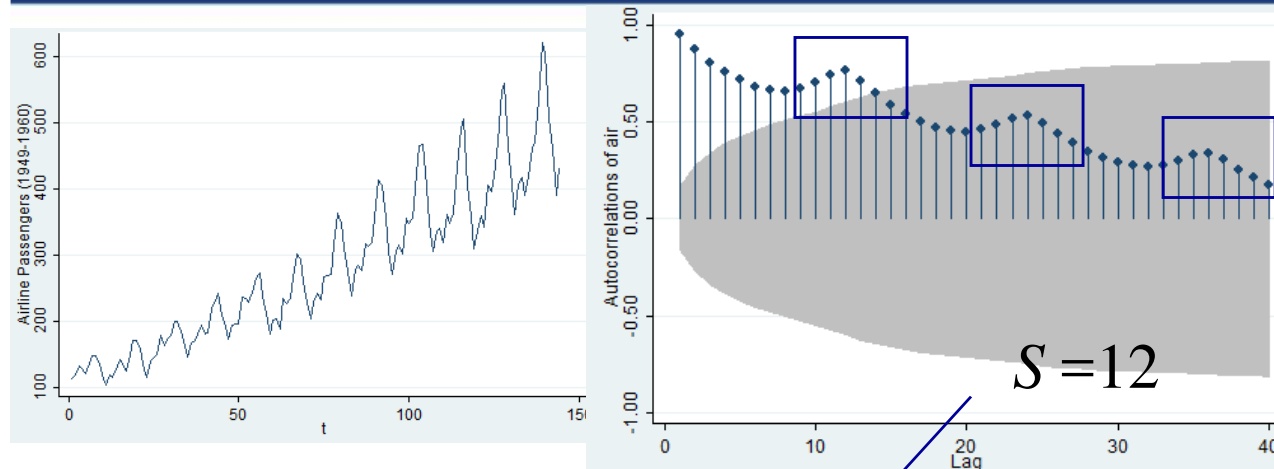
$$SARIMA(2,1,2)(2,1,2)_6$$



Модели SARIMA: практика

- *Не следует* строить смешанных моделей, одновременно включать SAR- и SMA-составляющих.
- **Идентификация** полной сезонной модели
$$\text{SARIMA}(p, d, q)(P_s, D_s, Q_s),$$
 - 1) логарифмирование исходного ряда (снижение дисперсии процесса),
 - 2) взятие несезонной и сезонной разности.
 - 3) исследование поведения ACF и PACF для полученного ряда.

Модели SARIMA



$S=12$

arima lnair, arima(0,1,1) sarima(0,1,1,12) noconstant

arima - ARIMA, ARMAX, and other dynamic regression

Model Model 2 Model 3 by/ff/in Weights SE/Robust

Seasonal ARIMA specification

☒ SARIMA (P,D,Q,S) specification

Autoregressive order (P)

Integrated (difference) order (D)

Moving-average order (Q)

Seasonal lag (S)

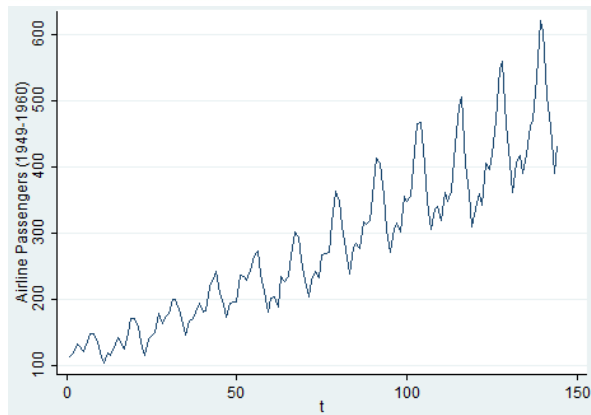
Sample: 14 - 144 Number of obs = 131
 wald chi2(2) = 84.53
 Log likelihood = 244.6965 Prob > chi2 = 0.0000

		Coef.	OPG Std. Err.	z	P> z
ARMA	DS12.lnair				
	ma L1.	-.4018324	.0730307	-5.50	0.000
ARMA12	ma L1.	-.5569342	.0963129	-5.78	0.000
	/sigma	.0367167	.0020132	18.24	0.000



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ ЭКОНОМИКИ И СТАТИСТИКИ

Модели SARIMA



arima lnair, arima(0,1,1) sarima(0,1,1,12) noconstant

Sample: 14 - 144 Number of obs = 131
 Wald chi2(2) = 84.53
 Log likelihood = 244.6965 Prob > chi2 = 0.0000

DS12.lnair	Coef.	OPG Std. Err.	z	P> z
ARMA				
ma L1.	-.4018324	.0730307	-5.50	0.000
ARMA12				
ma L1.	-.5569342	.0963129	-5.78	0.000
/sigma	.0367167	.0020132	18.24	0.000

$$\Delta\Delta_{12}\ln\text{air}_t = -0.402\epsilon_{t-1} - 0.557\epsilon_{t-12} + 0.224\epsilon_{t-13} + \epsilon_t$$

$$\hat{\sigma} = 0.037$$



Модели SARIMA: стационарность

Анализ стационарности SARIMA-моделей аналогичен анализу ARIMA-моделей.

Показать стационарность процессов.

$$y_t = y_{t-4} + \varepsilon_t$$

$$y_t = \frac{2}{3} y_{t-1} - \frac{1}{48} y_{t-4} + \varepsilon_t + \frac{1}{5} \varepsilon_{t-1}$$

$$y_t = 0.4 y_{t-1} + \varepsilon_t + 0.3 \varepsilon_{t-1} + 0.8 \varepsilon_{t-4}$$



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ
УПРАВЛЕНИЯ

Модели SARIMA: прогнозирование

Прогнозирование по SARIMA-моделям аналогично прогнозированию в классе ARIMA-моделей.

Построить прогноз

$$y_t = 0.4y_{t-1} + 0.2y_{t-4} + \varepsilon_t + 0.3\varepsilon_{t-1}$$



4.3. Сезонные единичные корни



Сезонные единичные корни

SARIMA(p, d, q)(P_s, D_s, Q_s)_s

Нестационарность процесса:

- Обычный единичный корень
- Сезонный единичный корень
- Обычный и сезонный единичный корень

Вопрос: брать обычную разность, сезонную или обе разности?

выводы о порядке интегрированности некоторых рядов

1. процедура сезонного сглаживания рядов и исследование порядка интегрированности тестами на единичные корни (тест Дики – Фуллера, тест Филипса – Перрона).

2. тестирование временных рядов на сезонные единичные корни:

- **DHF-test** – расширение ADF-test на случай сезонных ед. корней

Dickey D. A.; Hasza D. P.; Fuller W. A. (1984) Testing for Unit Root in Seasonal Time Series // Journal of the American Statistic Association, Vol. 79, pp 355-367.

Недостаток: не рассматривает возможность наличия обычных ед.корней

- **HEGY-test** – тестирование на случай обычных и сезонных ед.корней

Hylleberg S. Engle R. F. Granger C. W. J. Yoo B. S.(1990) Seasonal Integration and Cointegration // Journal of Econometrics 44 pp. 215-238.

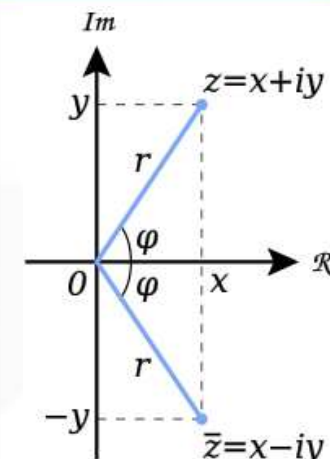
HEGY-test в случае квартальных данных ($s=4$)

$(1-L^4)=0 \rightarrow (1-z^4)=0$ имеет 4 единичных корня:

$z_1 = 1$, обычный ед.корень

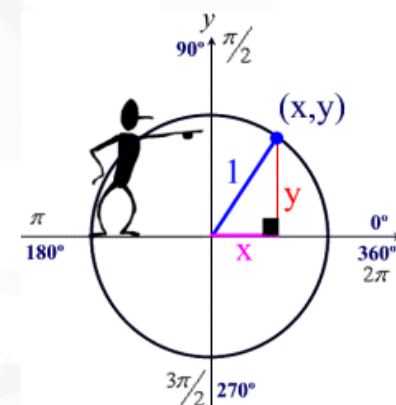
$z_2 = -1$, сезонный полугодовой ед.корень

$z_{3,4} = \pm i$, $i^2 = -1$ сезонный квартальный ед.корень



Спецификация теста:

$$B(L)y_t = \left\{ \alpha_0 + \sum_{k=1}^{s-1=3} \alpha_k d_t^{(k)} + \delta t \right\} + \varepsilon_t, \varepsilon_t \sim WN \quad (4.1)$$



$B(L)$ необходимо представить в виде суммы

Hylleberg S. Engle R. F. Granger C. W. J. Yoo B. S. (1990) Seasonal Integration and Cointegration // Journal of Econometrics 44 pp. 215-238.

HEGY-test в случае квартальных данных (s=4)

$$B(L)y_t = \left\{ \alpha_0 + \sum_{k=1}^{s-1=3} \alpha_k d_t^{(k)} + \delta t \right\} + \varepsilon_t, \varepsilon_t \sim WN$$

Теорема Лагранжа (теория приближенных вычислений)

1. Пусть z_k – корни уравнения $1-z^s=0$.

$$B(z) = \sum_{k=1}^s \lambda_k B_k(z) + B^*(z)(1-z^s) \quad (4.2)$$

$$B_k(z) = \frac{z}{z_k} \prod_{j \neq k} \left(1 - \frac{z}{z_j} \right), \quad k = \overline{1, s}$$

2. Разложение (4.2) обладает следующим свойством:

$$\lambda_k = 0 \Leftrightarrow z_k \text{ – корень уравнения } B(z) = 0.$$

(4.2) к (4.1):

$$B^*(L)\Delta_s y_t = \left\{ \alpha_0 + \sum_{k=1}^3 \alpha_k d_t^{(k)} + \delta t \right\} - \sum_{k=1}^4 \lambda_k B_k(L)y_t + \varepsilon_t \quad (4.3)$$



HEGY-test в случае квартальных данных ($s=4$)

Схема теста

$$B(z) = \sum_{k=1}^s \lambda_k B_k(z) + B^*(z)(1 - z^s) \quad (4.2)$$

$$B_k(z) = \frac{z}{z_k} \prod_{j \neq k} \left(1 - \frac{z}{z_j}\right), \quad k = \overline{1, s}$$

$\lambda_k = 0 \Leftrightarrow z_k$ – корень уравнения $B(z) = 0$.

$$B^*(L)\Delta_s y_t = \left\{ \alpha_0 + \sum_{k=1}^3 \alpha_k d_t^{(k)} + \delta t \right\} - \sum_{k=1}^4 \lambda_k B_k(L) y_t + \varepsilon_t \quad (4.3)$$

$k = 1, 2$ $H_0 : \lambda_k = 0 (\pi_k = 0)$ наличие несезон. ($k=1$) и сезон ($k=2$) ед. корня

$H_1 : \lambda_k > 0 (\pi_k < 0)$ стационарность процесса

t-статистика

$k = 3, 4$ $H_0 : \lambda_3 = 0, \lambda_4 = 0 (\pi_3 = \pi_4 = 0)$

F-статистика

Реализация в статистических пакетах

Инструменты - Пакеты функций - На сервере

GNU Regression, Econometric and
Time-series Library

gretl



gretl 2016c

MS Windows (x86_64)

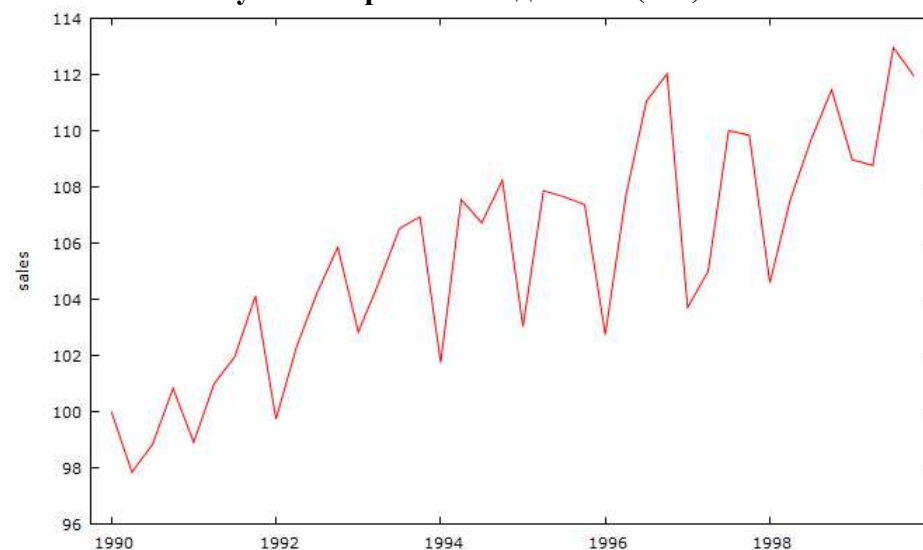
Дата сборки 2016-07-06

"От эконометристов, для эконометристов."

<http://gretl.sourceforge.net/>

Пакет	Версия	Описание
buys_ballot	1.6	Plots for seasonal time series
criteria	1.0	Calculate model selection criteria
DHF_test	1.041	Dickey-Hasza-Fuller Seasonal Unit Root Test
genr_dates	0.2	generate yyyy, mm, day, weekday
GHegy	1.1	General seasonal unit roots tests (HEGY) and P-values
gig	2.2	An assortment of univariate GARCH models
HIP	0.41	Heteroskedastic IV Probit

Случай квартальных данных ($s=4$)



HEGY test of seasonal unit roots for series sales:

AR order = 0 (determined by BIC with max.order=8)

Deterministic component: constant + (s-1) trigonometric terms

Dof (T-k) = 28

Statistic	p-value	Ang. Frequency	Period
t1= -1,46	1,00000	zero	infinity
F1= 6,76	0,99924	+pi/2	4
t2= -1,64	1,00000	pi	2
Fs= 7,37	0,18778	All the seasonal cycles	
Ft= 6,26	0,00924	Delta_s (all the seas. + zero freq.)	



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

Сезонные единичные корни в случае $s=12$

Тест HEGY для ежемесячных данных

$(1-L^{12})=0 \rightarrow 12$ единичных корней:

$$\pm 1; \pm i; -\frac{1 \pm i\sqrt{3}}{2}; \frac{1 \pm i\sqrt{3}}{2}; -\frac{\sqrt{3} \pm i}{2} \text{ и } \frac{\sqrt{3} \pm i}{2}.$$

Единичные корни и частоты в случае месячных данных

Сезонные единичные корни	Частота	
	В единицах π	Циклов/год
-1	π	6
$\pm i$	$\pm \frac{\pi}{2}$	3, 9
$-\frac{1}{2}(1 \pm \sqrt{3}i)$	$\mp \frac{2\pi}{3}$	8, 4
$\frac{1}{2}(1 \pm \sqrt{3}i)$	$\pm \frac{\pi}{3}$	2, 10
$-\frac{1}{2}(\sqrt{3} \pm i)$	$\mp \frac{5\pi}{6}$	7, 5
$\frac{1}{2}(\sqrt{3} \pm i)$	$\pm \frac{\pi}{6}$	1, 11

Критические значения для теста HEGY в случае месячных данных

Частота	Односторонняя t-статистика		F-статистика				
	1	π	$\pm \frac{\pi}{2}$	$\mp \frac{2\pi}{3}$	$\pm \frac{\pi}{3}$	$\mp \frac{5\pi}{6}$	$\pm \frac{\pi}{6}$
Критические значения	-3.19	-2.65	5.77	5.77	5.77	5.84	5.82

Beaulieu J.J., Miron J.A. (1993). Seasonal unit roots in aggregate U.S. data // Journal of Econometrics. № 55. P. 305–328.

Сезонные единичные корни в случае $s=12$

Тест HEGY для ежемесячных данных

Тестовое уравнение

$$\begin{aligned} \psi(L) \cdot (1 - L^{12}) \cdot y_t = & \pi_1 \cdot y_{1,t-1} + \pi_2 \cdot y_{2,t-1} + \pi_3 \cdot y_{3,t-1} + \pi_4 \cdot y_{4,t-1} + \pi_5 \cdot y_{5,t-1} + \pi_6 \cdot y_{6,t-1} \\ & + \pi_7 \cdot y_{7,t-1} + \pi_8 \cdot y_{8,t-1} + \pi_9 \cdot y_{9,t-1} + \pi_{10} \cdot y_{10,t-1} + \pi_{11} \cdot y_{11,t-1} + \pi_{12} \cdot y_{12,t-1} + \mu_t + \varepsilon_t \end{aligned}$$

$$y_{1,t} = (1 + L) \cdot (1 + L^2) \cdot (1 + L^4 + L^8) \cdot y_t$$

$$y_{7,t} = -(1 - L^4) \cdot (1 + \sqrt{3} \cdot L + L^2) \cdot (1 + L^2 + L^4) \cdot y_t$$

$$y_{2,t} = -(1 - L) \cdot (1 + L^2) \cdot (1 + L^4 + L^8) \cdot y_t$$

$$y_{8,t} = y_{7,t-1}$$

$$y_{3,t} = -(1 - L^2) \cdot (1 + L^4 + L^8) \cdot y_t$$

$$y_{9,t} = -(1 - L^4) \cdot (1 - L^2 + L^4) \cdot (1 - L + L^2) \cdot y_t$$

$$y_{4,t} = y_{3,t-1}$$

$$y_{10,t} = y_{9,t-1}$$

$$y_{5,t} = -(1 - L^4) \cdot (1 - \sqrt{3} \cdot L + L^2) \cdot (1 + L^2 + L^4) \cdot y_t$$

$$y_{11,t} = -(1 - L^4) \cdot (1 - L^2 + L^4) \cdot (1 + L + L^2) \cdot y_t$$

$$y_{6,t} = y_{5,t-1}$$

$$y_{12,t} = y_{11,t-1}$$



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

Реализация в статистических пакетах

GNU Regression, Econometric and
Time-series Library

gretl



gretl 2016c

MS Windows (x86_64)

Дата сборки 2016-07-06

"От эконометристов, для эконометристов."

<http://gretl.sourceforge.net/>

Случай ежемесячных данных ($s=12$)

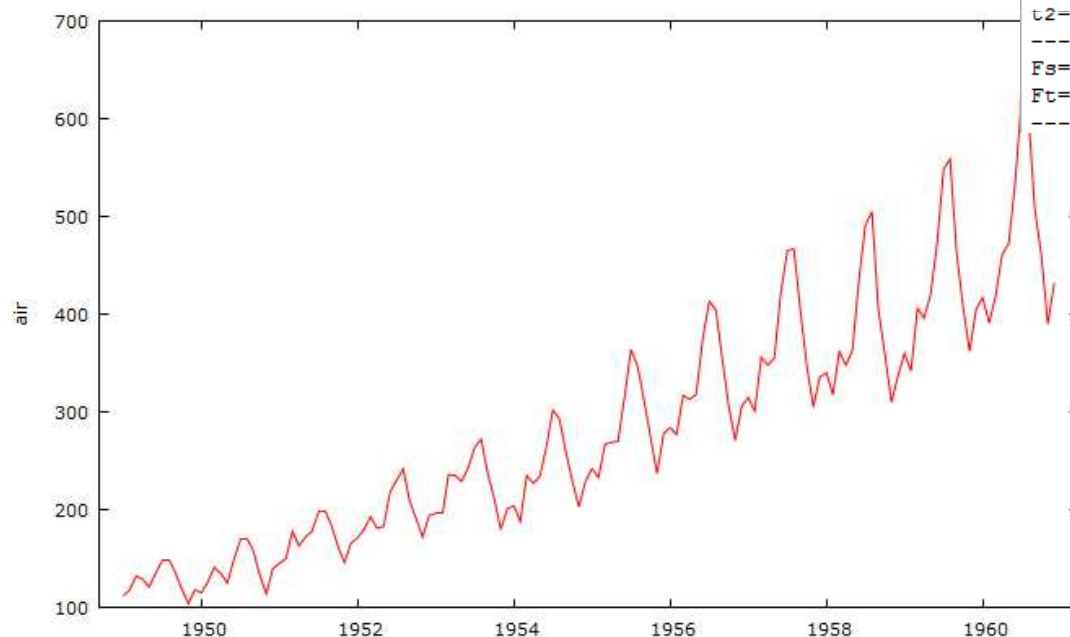
HEGY test of seasonal unit roots for series air:

AR order = 2 (determined by BIC with max.order=24)

Deterministic component: constant + (s-1) trigonometric terms

Dof (T-k) = 104

Statistic	p-value	Ang. Frequency	Period
t1= 0,86	0,99887	zero	infinity
F1= 0,08	0,99711	+pi/6	12
F2= 0,28	0,98899	+pi/3	6
F3= 2,27	0,72881	+pi/2	4
F4= 0,95	0,94293	+2*pi/3	3
F5= 4,87	0,24939	+5*pi/6	2,40
t2= -2,89	0,07047	pi	2
Fs= 2,33	0,73584	All the seasonal cycles	
Ft= 2,20	0,91422	Delta_s (all the seas. + zero freq.)	



Сезонные единичные корни: пример

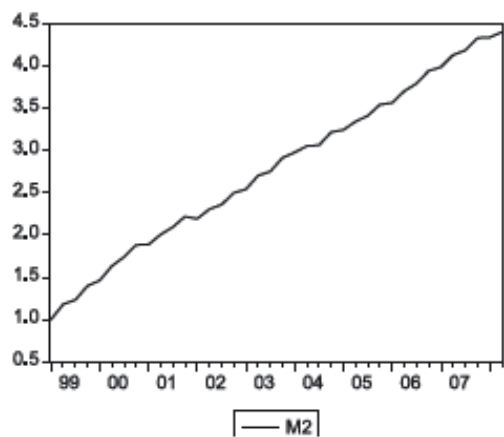


Рис. 7. Динамика логарифмов денежного агрегата М2 (I квартал 1999 г. – II квартал 2008 г.)

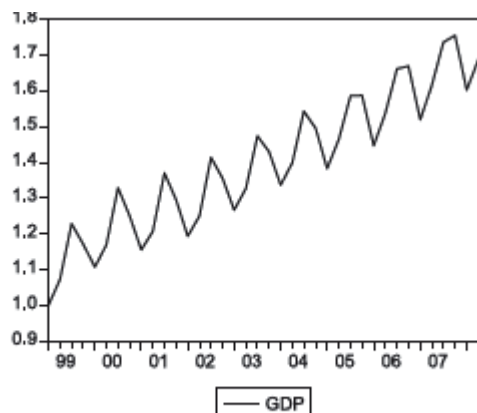


Рис. 10. Динамика логарифмов ВВП (I квартал 1999 г. – II квартал 2008 г.)

Моделирование спроса на деньги в российской экономике в 1999–2008 гг./ С. Дробышевский [и др.];– М.: ИЭПП, 2010.

http://www.iep.ru/files/text/working_papers/136.pdf

Ряд/ тест	HEGY	С константой	С константой и дамми	С константой и трендом	С констан- той, дамми, трендом
<i>m2</i>	Единичный Полугодовой	Единичный Полугодовой	Единичный	Единичный Полугодовой	Квартальный
<i>gdp</i>	Единичный Полугодовой Квартальный	Единичный Полугодовой Квартальный	Единичный Полугодовой	Единичный Полугодовой Квартальный	Единичный Полугодовой

устойчивость результатов теста к смене модификации

Ряд	Модифика- ция теста	Значение статистик	Критические значения при уровне значимости 0,05	Вывод
<i>m2</i>	С константой	$\gamma_1 = -1.30$ $\gamma_2 = -1.39$ $\gamma_3 = -2.97$ $\gamma_4 = -3.51$	-2.96 -1.95 -1.90 -2.06 или 2.04	Единичный и полугодовой единичный корни
<i>gdp</i>	С константой и трендом	$\gamma_1 = -0.40$ $\gamma_2 = -0.59$ $\gamma_3 = -0.47$ $\gamma_4 = -0.79$	-3.56 -1.91 -1.92 -2.05 или 1.96	Единичный, полугодовой единичный и квартальный корни

Dickey D. A.; Hasza D. P.; Fuller W. A. (1984) Testing for Unit Root in Seasonal Time Series. // Journal of the American Statistic Association, Vol. 79, pp 355-367.

Franses P. H. and Taylor R. M. A. (2000) Determining the Order of Differencing in Seasonal Time Series Processes // Econometric Journal Vol.3, pp. 250-264.

Franses P. H. and Paap R. (2004) Periodic Time Series Models // Oxford University Press 2004

Ghysels E., Lee H. S., and Noh J. (1994) Testing for Unit Root in Seasonal Time Series // Journal of Econometrics Vol. 62 pp. 415-442.

Hasza D. P. and Fuller W. A. (1982) Testing for Nonstationary Parameter Specifications in Seasonal Time Series Models // The Annals of Statistics Vol. 10. No. 4, pp.1209-1216.

Hylleberg S. Engle R. F. Granger C. W. J. Yoo B. S. (1990) Seasonal Integration and Cointegration // Journal of Econometrics 44 pp. 215-238.

Моделирование спроса на деньги в российской экономике в 1999–2008 гг./ С. Дробышевский [и др.]. – М.: ИЭПП, 2010. – 144 с. http://www.iep.ru/files/text/working_papers/136.pdf

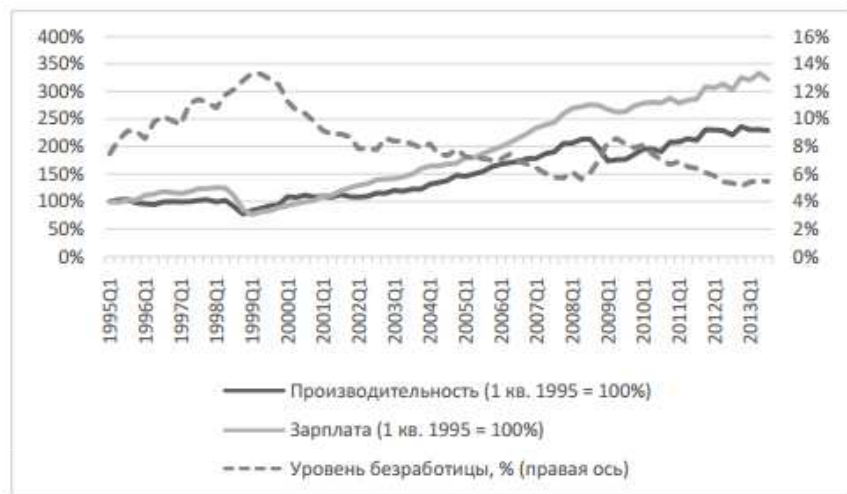
Выдумкин Платон. Сезонные единичные корни и сезонная коинтеграция. Теоретические основы и практическое применение. (обзор)

<https://www.hse.ru/data/2010/10/21/1222645995/статья%20Выдумкин%20Платон.pdf>

A seasonal unit root test with STATA

http://www.stata.com/meeting/italy08/depalo_2008.pdf

Корректировка сезонности



Анализ взаимозависимости и оценивание тенденции развития требует очистки от эффекта сезонности.

Рис. 3. Динамика производительности труда, средней заработной платы и уровня безработицы, сглаженные на сезонность⁸

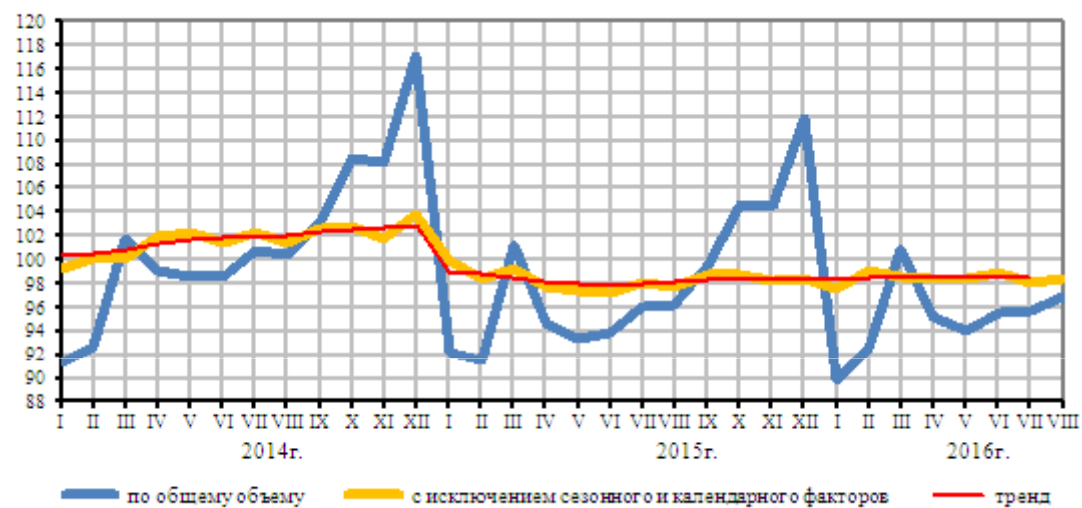
Вакуленко

http://www.hse.ru/data/2014/12/22/1103212064/WP3_2014_08_f.pdf

Пример: 26 апреля 2016. FINMARKET.RU - ВВП России с **исключением сезонного и календарного факторов** в марте 2016 года снизился на 0,4% после роста на 0,3% в феврале, говорится в ежемесячном обзоре Внешэкономбанка (ВЭБ). До февраля ВВП РФ **со снятой сезонностью** снижался в течение четырех месяцев подряд.
<http://www.finmarket.ru/news/4270683>

Корректировка сезонности

Индексы промышленного производства
в % к среднемесячному значению 2013г.



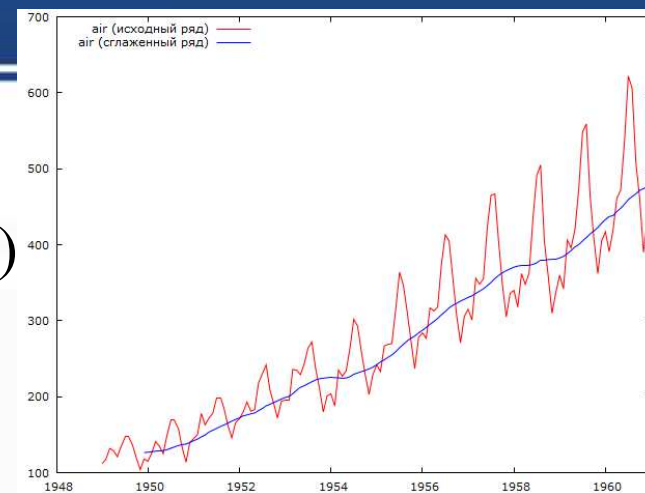
http://www.gks.ru/bgd/free/b04_03/IssWWW.exe/Stg/d01/image1592.gif

Корректировка сезонности: основные подходы

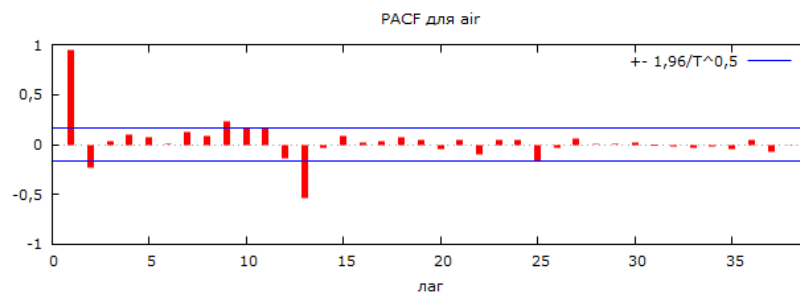
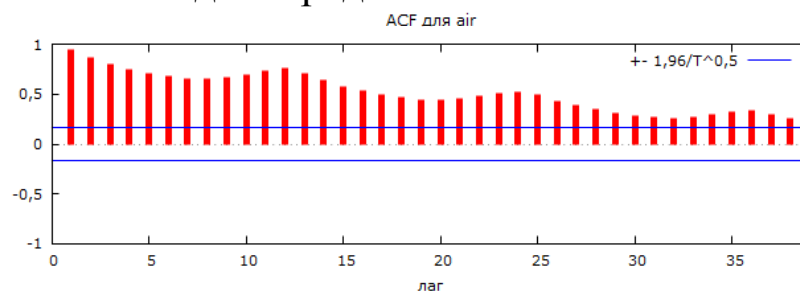
1. Скользящие средние (разных типов: простые, взвешенные и т.д)

См. Дуброва Т.А. Статистические методы прогнозирования 2003

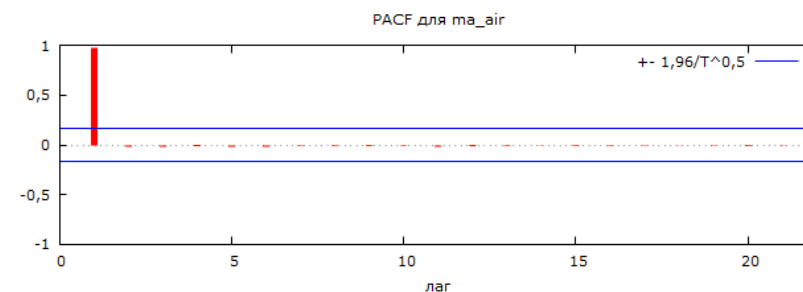
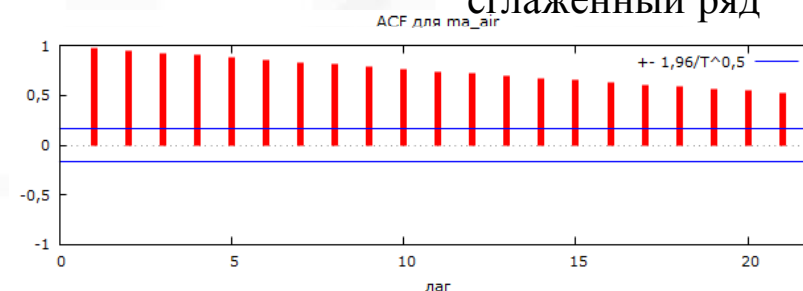
<http://www.twirpx.com/file/18580/>



исходный ряд



сглаженный ряд



Примеры.

1. Простая СС $\hat{y}_t = \frac{1}{3}(y_{t-1} + y_t + y_{t+1})$ (по 3-м точкам)

Для устранения сезонных колебаний часто требуется использовать 4- и 12-членные скользящие средние.

$$\hat{y}_t = \frac{\frac{1}{2}y_{t-2} + y_{t-1} + y_t + y_{t+1} + \frac{1}{2}y_{t+2}}{4}$$

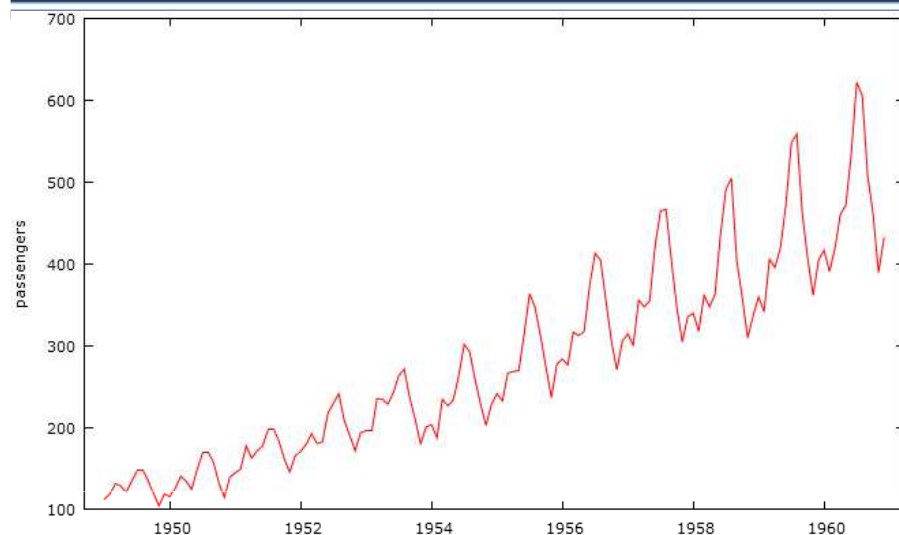
$$\hat{y}_t = \frac{\frac{1}{2}y_{t-6} + y_{t-5} + \dots + y_t + \dots + y_{t+5} + \frac{1}{2}y_{t+6}}{12}$$

2. Взвешенная СС

$$\hat{y}_t = \frac{1}{35}[-3y_{t-2} + 12y_{t-1} + 17y_t + 12y_{t+1} - 3y_{t+2}]$$

(по 5-ти точкам, аппроксимация полиномом 2-го порядка)

Корректировка сезонности: основные подходы



Gretl: Добавить – Фиктивную переменную для периода

2. оценивание регрессии для ряда на сезонные фиктивные переменные:

D1,..., D12 (месячные данные) D1,..., D4 (квартальные).

-Остатки от оцененной регрессии - *очищенный ряд*, к которому можно применять DF-test. [Dickey, Bell, Miller (1986)]: асимптотическое распределение DF статистики не изменяется при исключении из ряда детерминированных сезонных компонент.

тест с константой включая сезонные фиктивные переменные
модель: $(1-L)y = b_0 + (a-1)*y(-1) + \dots + e$
коэф. автокорреляции 1-го порядка для e: 0,019
оценка для $(a - 1)$: -0,00924108
тестовая статистика: $\tau_c(1) = -0,705617$
асимпт. p-значение 0,8436

Регрессия расширенного теста Дики-Фуллера
МНК, использованы наблюдения 1949:03-1960:12 (T = 142)
Зависимая переменная: d_g

	Коэффициент	Ст. ошибка	t-статистика	P-значение	
const	37,5225	6,70374	5,597	1,27e-07	***
g_1	-0,00924108	0,0130965	-0,7056	0,8436	
d_g_1	0,188766	0,0870947	2,167	0,0321	**
dm1	-33,3151	9,03628	-3,687	0,0003	***
dm2	-44,4442	8,10237	-5,485	2,12e-07	***
dm3	1,09000	7,50030	0,1453	0,8847	
dm4	-44,7475	9,29879	-4,812	4,13e-06	***
dm5	-29,7223	7,60609	-3,908	0,0002	***
dm6	3,92624	7,86946	0,4989	0,6187	
dm7	-2,49487	9,57433	-0,2606	0,7948	
dm8	-42,0135	9,60322	-4,375	2,49e-05	***
dm9	-82,8976	7,79895	-10,63	2,65e-019	***
dm10	-61,3745	7,31233	-8,393	7,56e-014	***
dm11	-62,0449	7,13996	-8,690	1,48e-014	***

Корректировка сезонности: основные подходы

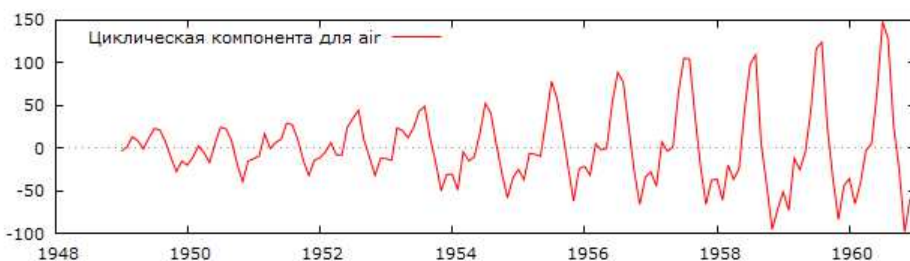
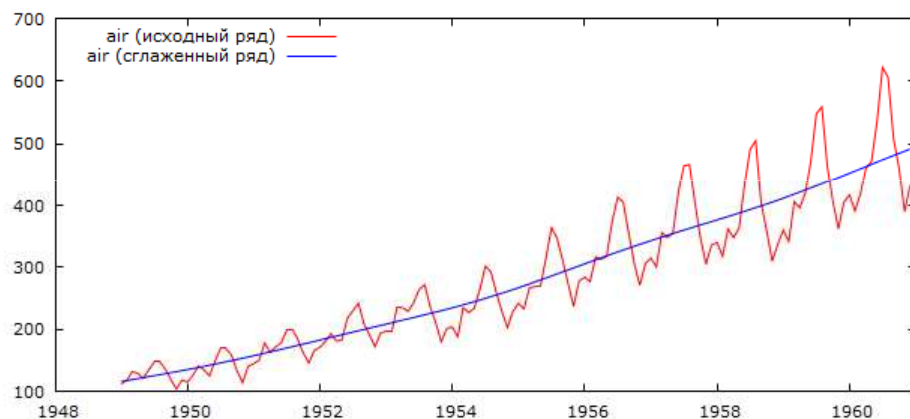
3. Использование фильтров – фильтр Ходрика-Прескотта

$$HP_{\lambda}(L) = \frac{1}{1 + \lambda \Delta^2 (1 - L^{-1})^2}, \lambda > 0$$

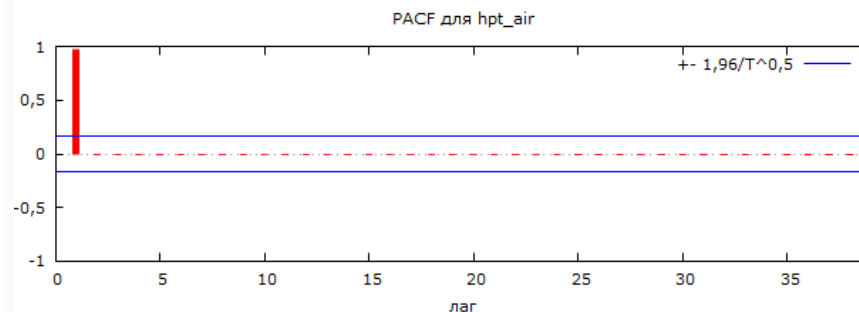
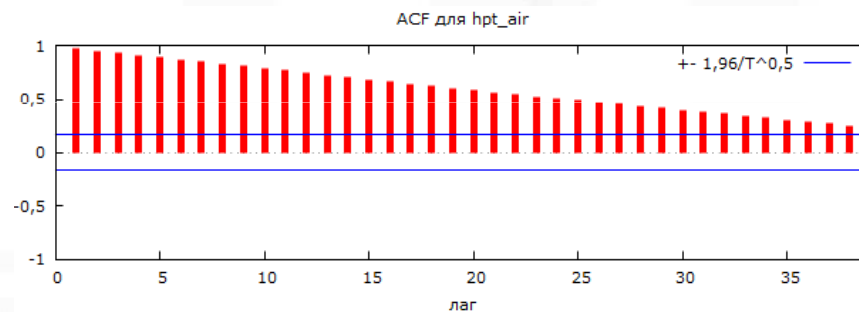
$\lambda=100$ – годовые данные

$\lambda=1600$ – квартальные данные

$\lambda=14400$ – месячные данные



сглаженный ряд



Hodrick, Robert; Prescott, Edward C. (1997). "Postwar U.S. Business Cycles: An Empirical Investigation". Journal of Money, Credit, and Banking. 29 (1): 1–16.



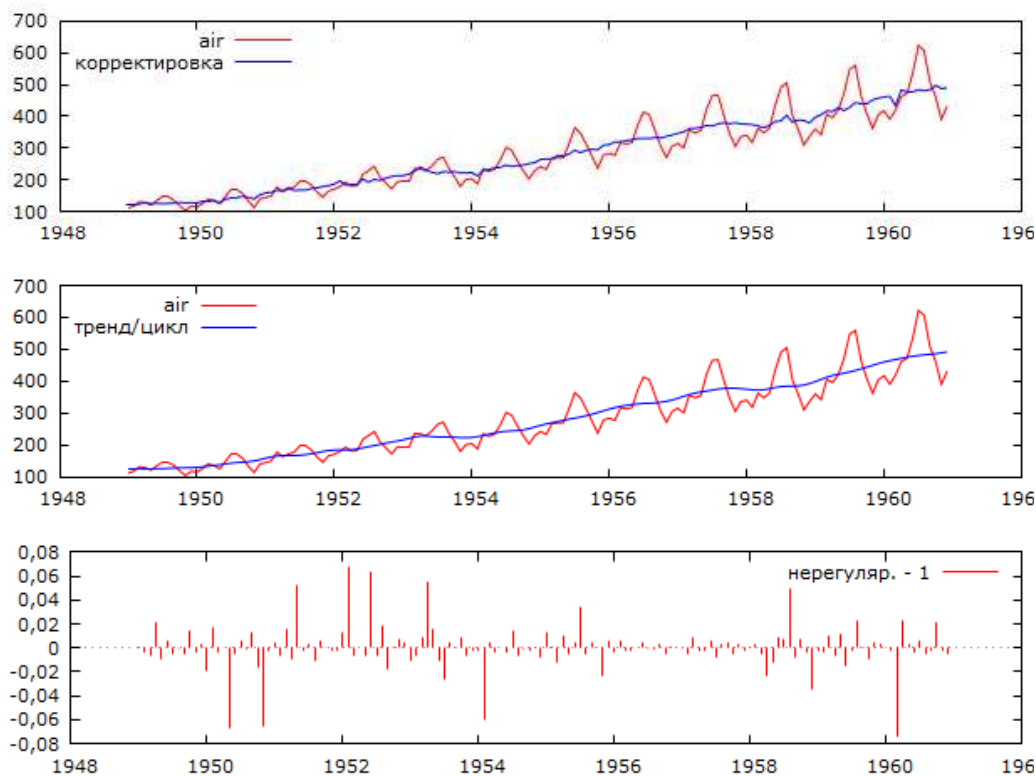
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

Корректировка сезонности: основные подходы

4. U.S. Census Bureau's EuroSTAT

X-12-ARIMA TRAMO

(Time Series Regression with ARIMA Noise, Missing Observations and Outliers)



сглаженный ряд

