



## Тема 1. 2.

# Исследование и моделирование нестационарной составляющей временного ряда



# Структура курса

## **Тема 1. Введения в анализ одномерных временных рядов**

1.1. Временной ряд: основные понятия, определения, характеристики. Простейшие примеры стационарных и нестационарных временных рядов (белый шум, временной ряд с линейным трендом, случайное блуждание, случайным блужданием с дрейфом) и их характеристики. Основные составляющие временного ряда.

1.2. **Исследование и моделирование неслучайной составляющей временного ряда: основные типы трендов (детерминированный и стохастический); проверка наличия тренда во временных рядах; методы выделения тренда.**



## Генезис наблюдений, образующих ВР

4 типа факторов (Айвазян):

(А) Долговременные

(Б) Сезонные

(В) Циклические

(Г) Случайные

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



## Неслучайная составляющая ВР

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

Задача начального этапа анализа ВР:

1. Выявление неслучайной составляющей
2. Аппроксимация



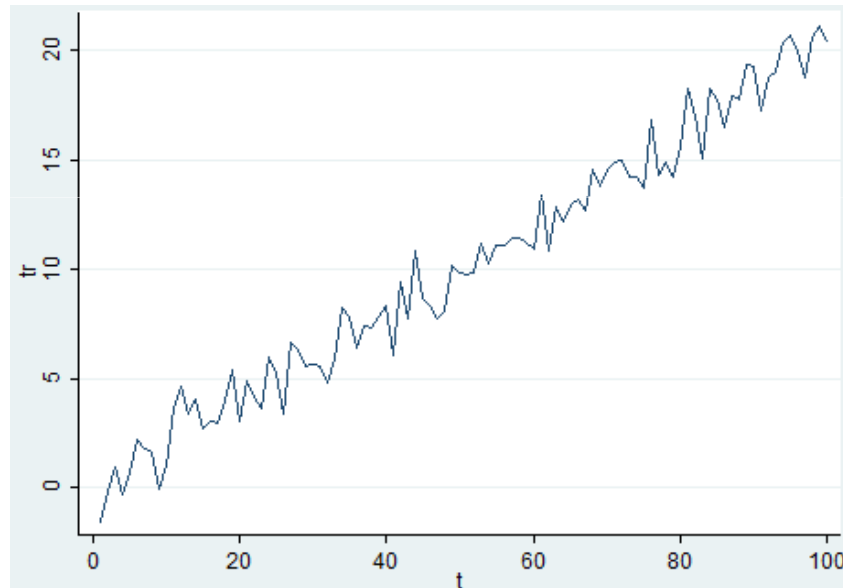
$$Y_t = \chi(A) f_{mp}(t) + \chi(B) \varphi(t) + \chi(B) \psi(t) + \varepsilon_t$$

Под *трендом* понимают изменение, определяющее общее направление развития, основную тенденцию временного ряда. Это систематическая составляющая долговременного действия.

- Детерминированные тренды
- Стохастические тренды

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

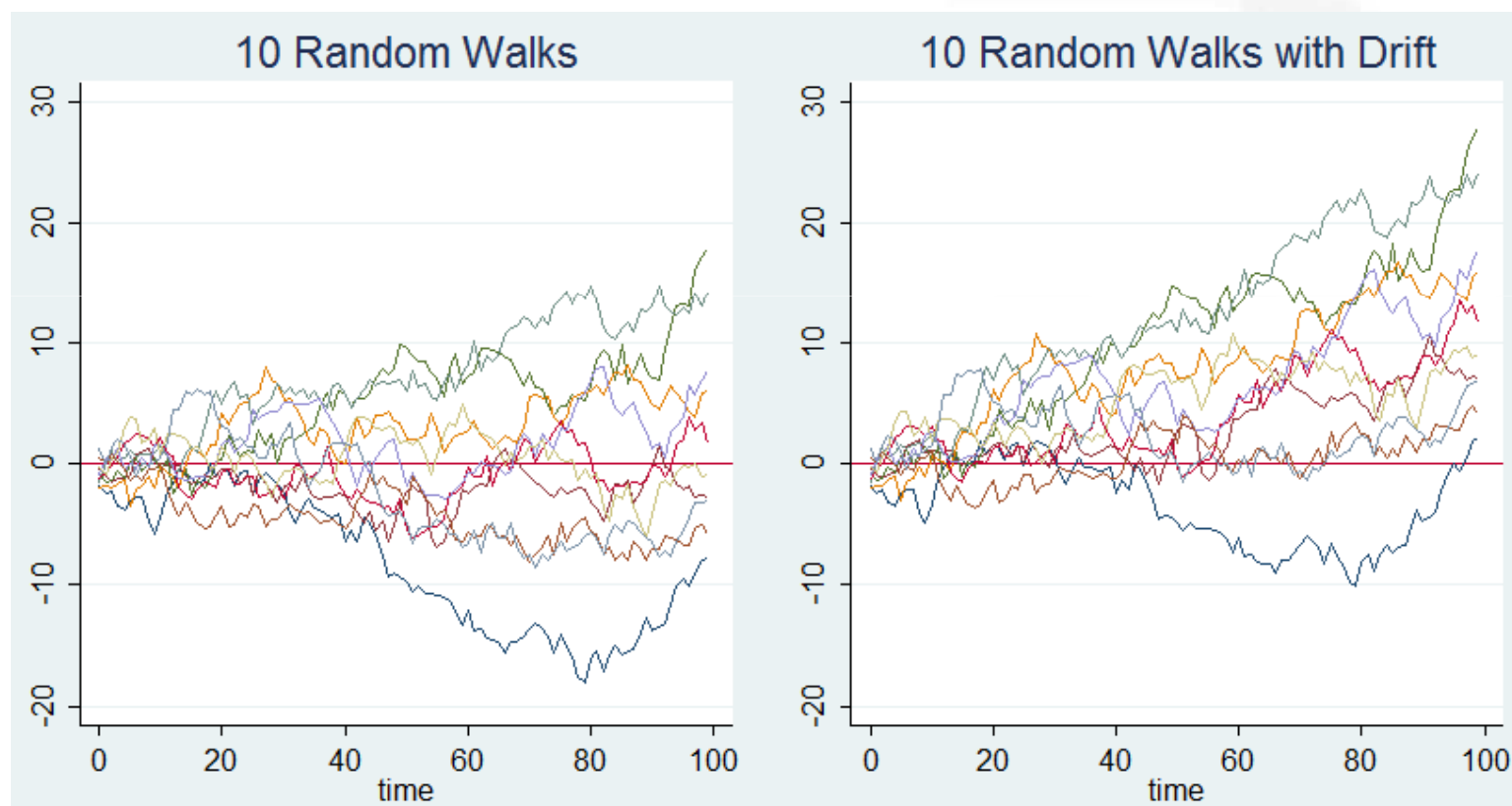
- Детерминированные тренды
- Стохастические тренды



Clive W. J. Granger and P. Newbold Spurious regressions in econometrics // Journal of Econometrics, 1974, vol. 2, issue 2, pages 111-120

Nelson, Charles R. and Charles I. Plosser (1982) Trends and random walks in macroeconomic time series: Some evidence and implications // Journal of Monetary Economics, 10, 139-162.

Clive W. J. Granger and P. Newbold Spurious regressions in econometrics //Journal of Econometrics, 1974, vol. 2, issue 2, pages 111-120



First TS Example.do



# Стохастические тренды

pwcorr Yd1 Yd2 Yd3 Yd4 Yd5 Yd6 Yd7 Yd8 Yd9 Yd10, sig										
	Yd1	Yd2	Yd3	Yd4	Yd5	Yd6	Yd7	Yd8	Yd9	Yd10
Yd1	1.0000									
Yd2	-0.4529 0.0000	1.0000								
Yd3	-0.4337 0.0000	0.7444 0.0000	1.0000							
Yd4	-0.3626 0.0002	0.7302 0.0000	0.8440 0.0000	1.0000						
Yd5	-0.5054 0.0000	0.8417 0.0000	0.9047 0.0000	0.9277 0.0000	1.0000					
Yd6	-0.4244 0.0000	0.8466 0.0000	0.7411 0.0000	0.7127 0.0000	0.8464 0.0000	1.0000				
Yd7	-0.2742 0.0058	0.7187 0.0000	0.7754 0.0000	0.8788 0.0000	0.8643 0.0000	0.8078 0.0000	1.0000			

pwcorr Y1 Y2 Y3 Y4 Y5 Y6 Y7 Y8 Y9 Y10, sig										
	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9	Y10
Y1	1.0000									
Y2	0.1537 0.1267	1.0000								
Y3	-0.6742 0.0000	-0.3150 0.0014	1.0000							
Y4	-0.4395 0.0000	-0.2712 0.0064	0.4944 0.0000	1.0000						
Y5	-0.7970 0.0000	-0.1264 0.2102	0.7509 0.0000	0.7184 0.0000	1.0000					
Y6	-0.1298 0.1981	0.4542 0.0000	0.0118 0.9073	-0.0595 0.5564	0.2520 0.0114	1.0000				
Y7	-0.1540 0.1261	-0.0383 0.7049	0.2637 0.0080	0.5906 0.0000	0.4711 0.0000	0.3905 0.0001	1.0000			





## Spurious regressions

reg Y1 Y2 Y3

Source	SS	df	MS
Model	<b>1395.10078</b>	<b>2</b>	<b>697.550391</b>
Residual	<b>1648.22882</b>	<b>97</b>	<b>16.9920497</b>
Total	<b>3043.3296</b>	<b>99</b>	<b>30.7407031</b>

Number of obs = **100**  
F( 2, 97) = **41.05**  
Prob > F = **0.0000**  
R-squared = **0.4584**  
Adj R-squared = **0.4472**  
Root MSE = **4.1221**

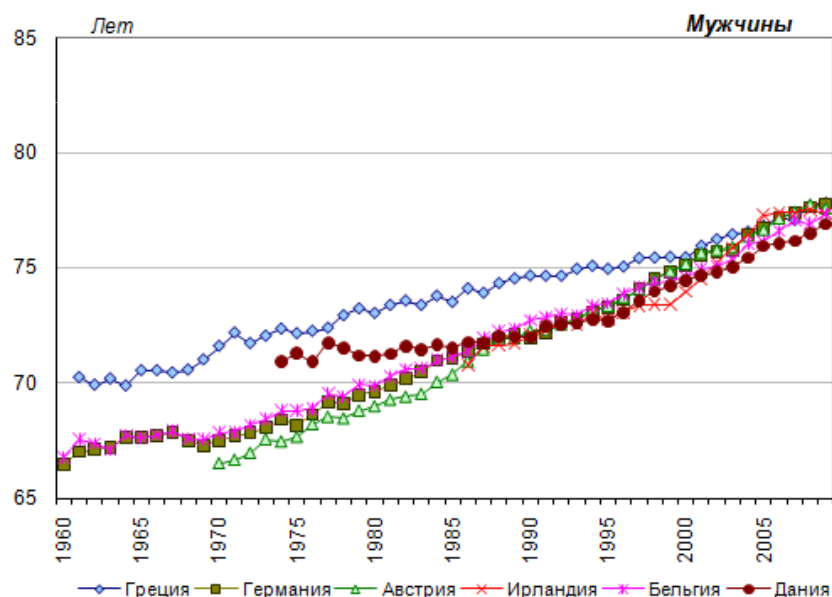
Y1	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
Y2	<b>-.2100813</b>	<b>.2541191</b>	<b>-0.83</b>	<b>0.410</b>	<b>-.7144374</b>	<b>.2942748</b>
Y3	<b>-.8630177</b>	<b>.0977989</b>	<b>-8.82</b>	<b>0.000</b>	<b>-1.057121</b>	<b>-.668914</b>
_cons	<b>-4.159776</b>	<b>.7214576</b>	<b>-5.77</b>	<b>0.000</b>	<b>-5.59167</b>	<b>-2.727883</b>



# Детерминированные тренды

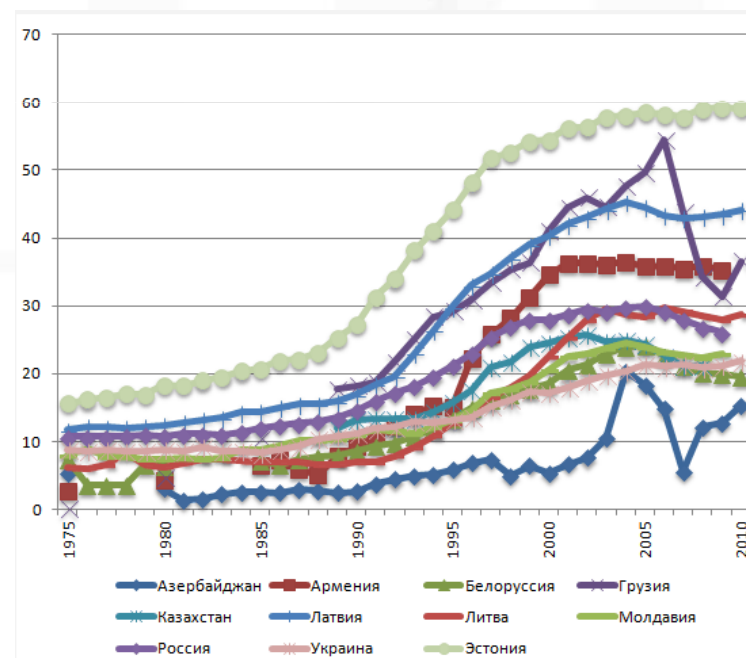
Детерминированные тренды могут быть описаны с помощью функции времени. Все кривые роста делят на 3 класса:

1. Монотонный характер, отсутствие пределов роста (класс полиномов)
2. Имеет предел роста (модифицированная экспонента)
3. S-образные кривые (логистическая кривая)



Ожидаемая продолжительность жизни мужчин при рождении

<http://demoscope.ru/weekly/2011/0463/barom02.php>



Внебрачная рождаемость в некоторых постсоветских странах

# Проверка гипотезы случайности выборки (randomness test)



**Присутствует ли тренд?  
Является ли выборка  
случайной?**

**$H_0: E(y_t) = \mu = \text{const}$**  (о случайности ряда, об отсутствии тренда)

**$H_1: E(y_t) \neq \text{const}$**

Непараметрические критерии:

- критерий серий, основанный на медиане выборки;
- метод Фостера-Стюарта;
- критерий «восходящих» и «нисходящих» серий и др

(Носко)



# Критерий серий (Runs test)

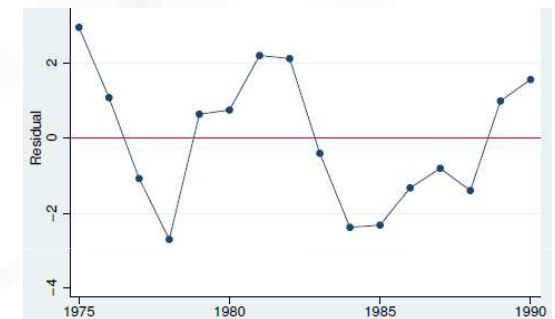
Swed and Eisenhart (1943)

**Идея:** Пусть  $M$  – медиана распределения  $F$ , тогда зн-я  $y_1, \dots, y_n$  не должны «слишком долго» задерживаться по одну сторону от  $M$ .

1. Ранжируем исходный ряд  $y_t$  по возрастанию.

Находим медиану:

$$Me = med(y_1, \dots, y_n) = \begin{cases} y_{n+1/2} & , n - \text{четно} \\ \frac{1}{2} \left( y_{n/2} + y_{n/2+1} \right) & , n - \text{нечетно} \end{cases}$$



2. Определяем **серии** - последовательность «+» и «-»

по правилу:

$$n_i = \begin{cases} +, & y_t > Me, t = \overline{1, n} \\ -, & y_t < Me, t = \overline{1, n} \end{cases}$$

**Серия** – это последовательность подряд идущих плюсов и подряд идущих минусов.



## Критерий серий (Runs test):

### Критерий серий, основанный на медиане

3.  $r(n)$  -число серий в совокупности  $n_i$ .

$\tau(n)$ - протяженность самой длинной серии.

4. Приближенное правило:

(Айвазян)

$$r(n) > \left[ \frac{1}{2} \left( n + 2 - 1,96\sqrt{n-1} \right) \right]$$

$$\tau(n) < [1,43 \ln(n+1)]$$

**Пример: Изменения курса акций промышленной компании в течение месяца**

t	$y_t$	$y_t^ $		t	$y_t$	$y_t^ $		t	$y_t$	$y_t^ $	
1	509	507		7	520	512		15	514	519	
2	507	508	-	8	519	514	+	16	510	520	-
3	508	509	-	9	512	515	-	17	516	521	-
4	509	509	+	10	511	516	-	18	518	524	+
5	518	510	+	11	517	517	+	19	524	524	+
6	515	511	-	12	524	518	+	20	521	526	+
				13	526	518	+				
				14	519	519	+				

Дуброва



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
УНИВЕРСИТЕТ

# Критерий серий (Runs test)

Swed and Eisenhart (1943)

$$\mu_r = \frac{2n_0n_1}{N} + 1, \quad N = n_0 + n_1$$

$$\sigma_r^2 = \frac{2n_0n_1(2n_0n_1 - N)}{N^2(N-1)}$$

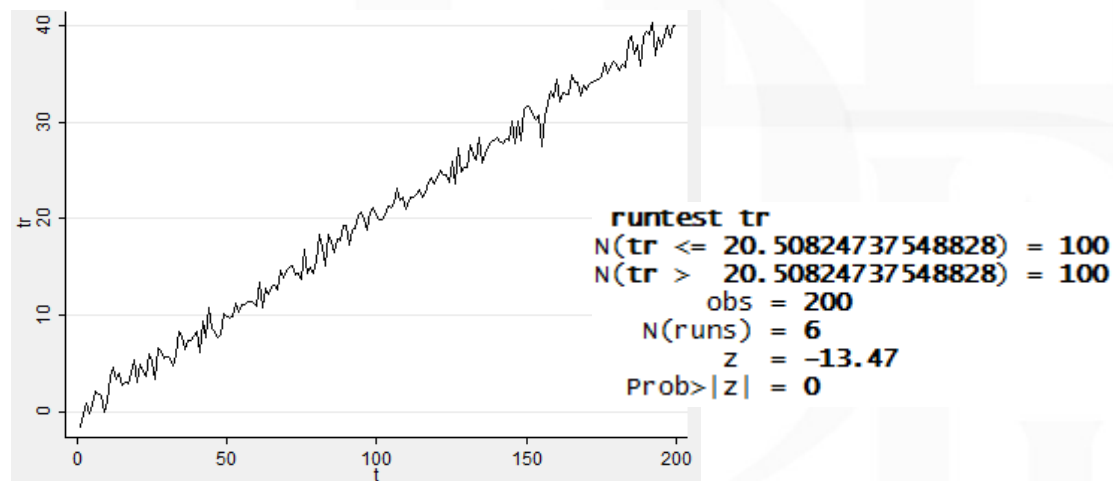
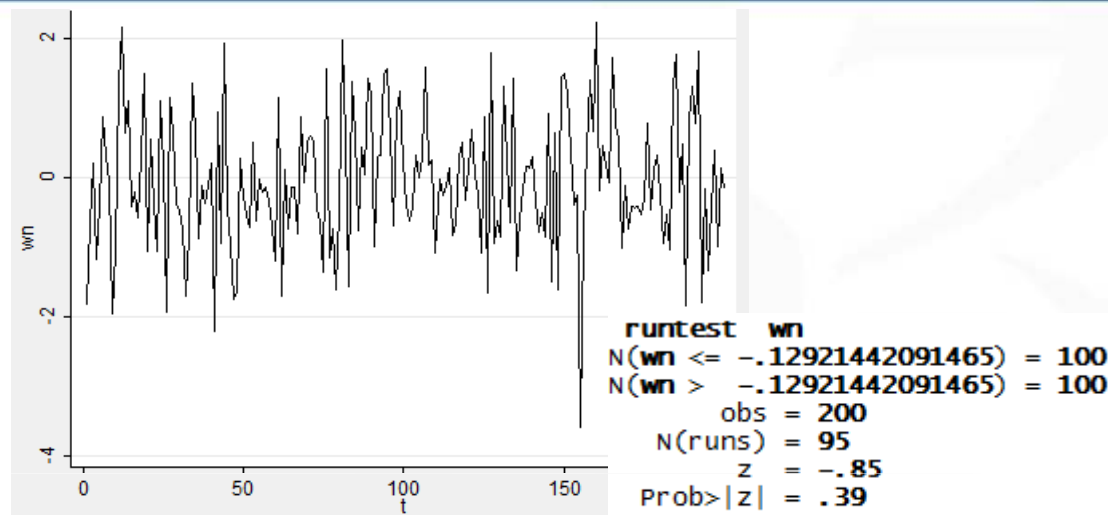
$$\hat{Z} = \frac{r - \mu_r}{\sigma_r} \sim N(0,1)$$

$r$  – число серий

$n_0$  – число наблюдений,  
меньших медианы,

$n_1$  – число наблюдений,  
больших медианы,

$$N = n_0 + n_1$$



# Наличие детерминированного тренда: пример



Индекс МосБиржи

Наличие тренда

```
. runtest IMOEX
N(IMOEX <= 2521.1) = 127
N(IMOEX > 2521.1) = 126
      obs = 253
      N(runs) = 8
      z = -15.06
      Prob>|z| = 0
```

Результаты теста?

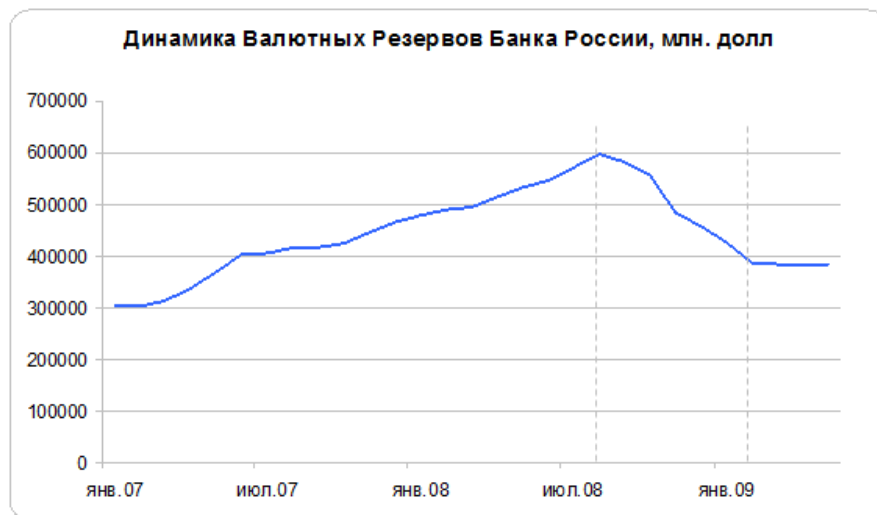


## Анализ структурных сдвигов

**Структурный сдвиг** - неожиданное изменение во временном ряду. Может приводить к изменениям в параметрах модели регрессии, ошибкам прогнозирования и недостоверности модели.

- Графический анализ

### Влияние экономического кризиса

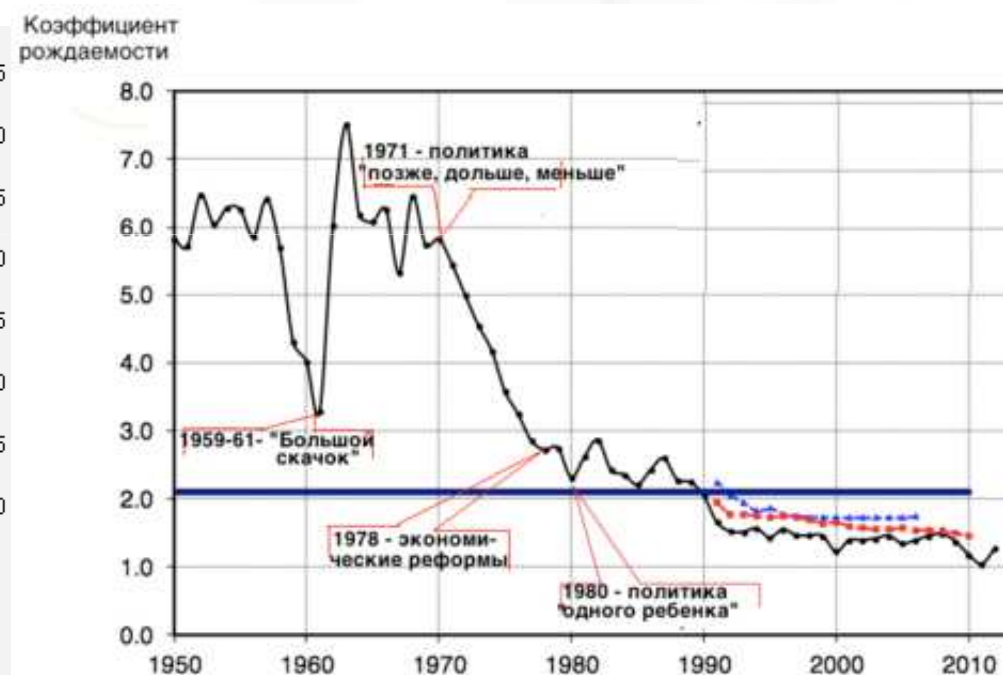
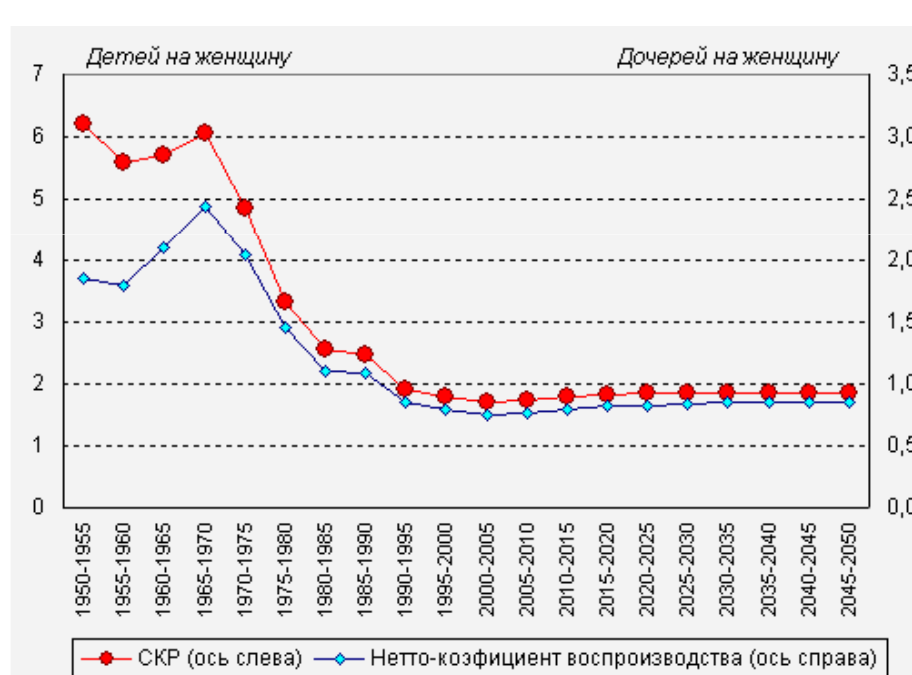


**Кредиты населению в валюте, млн.руб.**



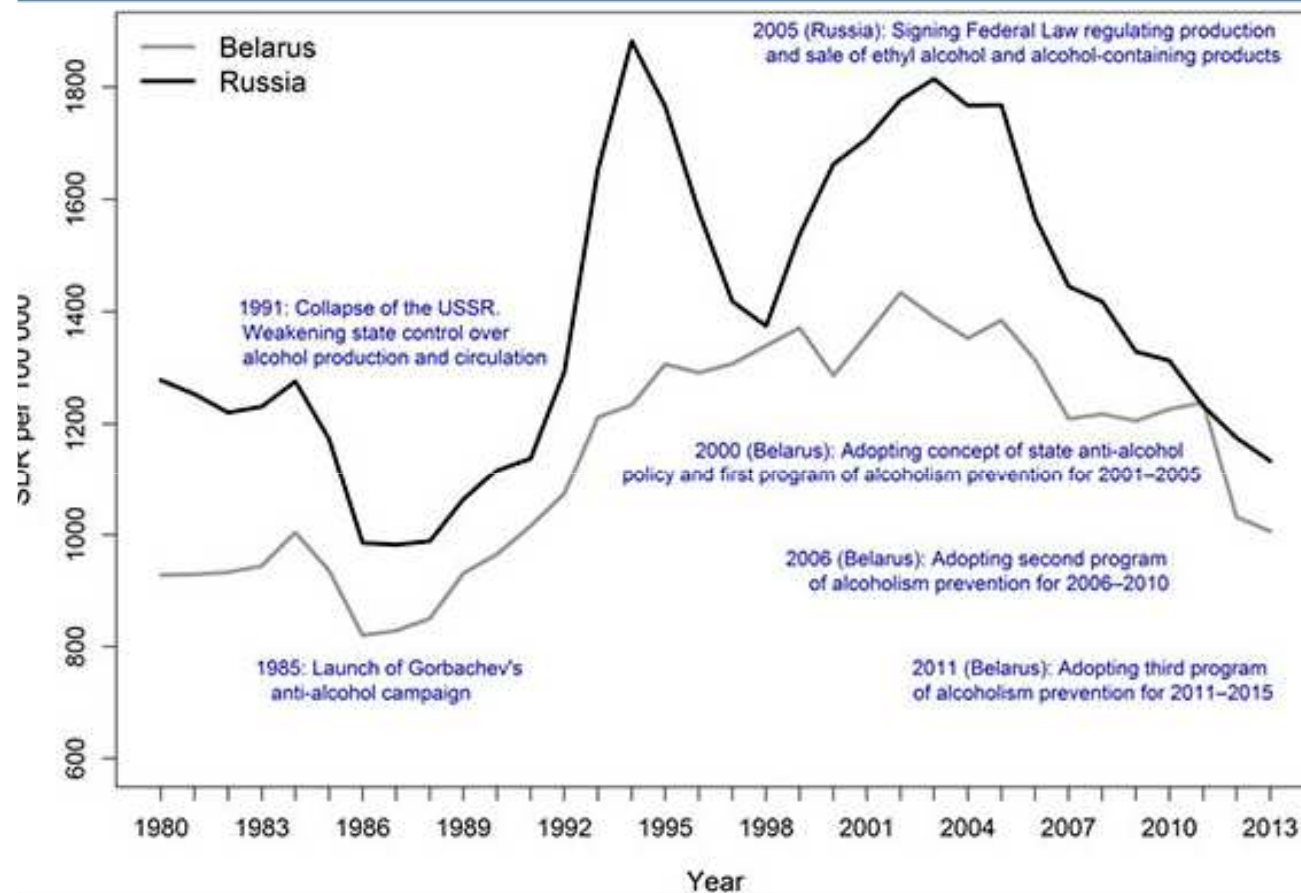
# Анализ структурных сдвигов

## Влияние демографической политики государства



Коэффициент суммарной рождаемости и нетто-коэффициент воспроизводства населения в Китае

# Анализ структурных сдвигов: влияние антиалкогольной кампании



1985 - начало Горбачевской антиалкогольной кампании;

1991 - распад СССР и ослабление контроля за производством и распространением алкоголя;

2000 (Белоруссия) - принята первая программа борьбы с алкоголизмом, рассчитанная на 2001–2005 гг.

2005 (Россия) - принят федеральный закон, регламентирующий производство и продажу алкогольной продукции;

2006 (Белоруссия) - принята вторая программа борьбы с алкоголизмом, рассчитанная на 2006–2010 гг.

2011 (Белоруссия) - принята третья программа борьбы с алкоголизмом, рассчитанная на 2011–2015 гг.

**Стандартизованный коэффициент смертности от причин, связанных с потреблением  
алкоголя, мужчины в возрасте 20–64 лет**

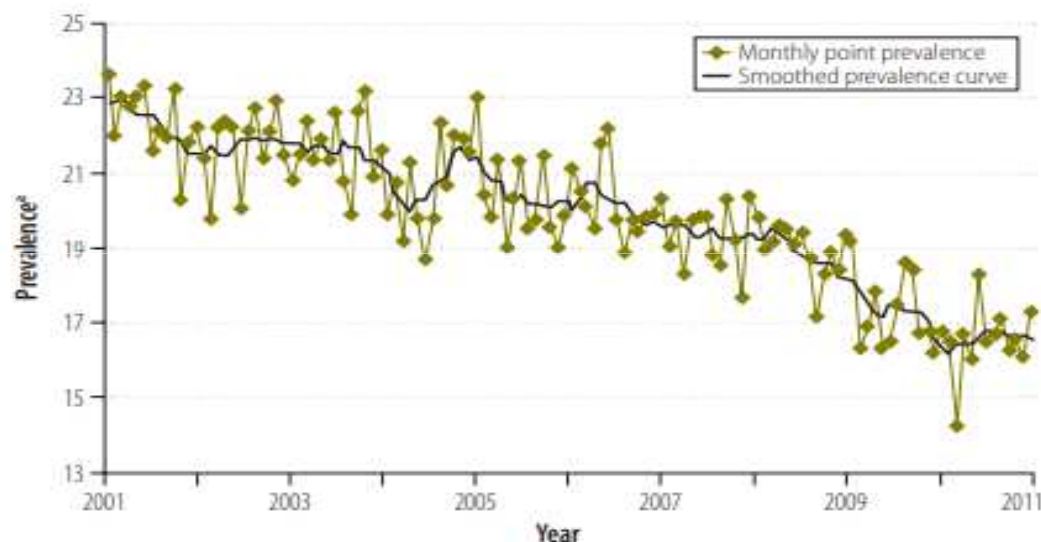
Grigoriev, P., & Andreev, E. M. (2015). The huge reduction in adult male mortality in Belarus and Russia: is it attributable to anti-alcohol measures? PLoS One, 10(9), e0138021

<http://demoscope.ru/weekly/2016/0679/digest01.php>

## Анализ структурных сдвигов: не всегда скачки

### Как меры политики влияют на потребление табака в Австралии?

Fig. 1. Monthly smoking prevalence among Australian adults, January 2001 to June 2011



\* Smokers per 100 respondents.

<https://www.who.int/bulletin/volumes/92/6/13-118448.pdf?ua=1>

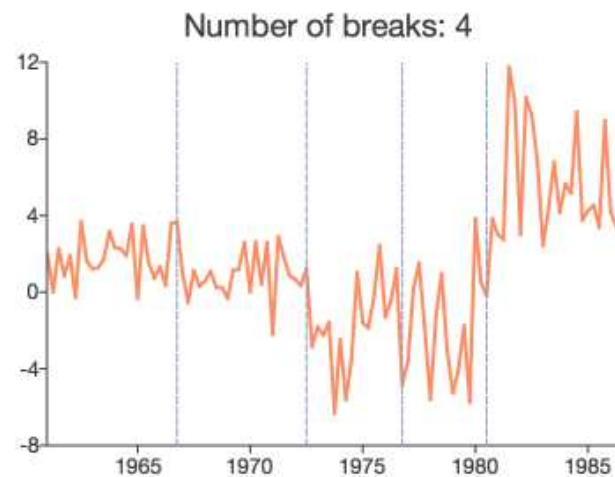
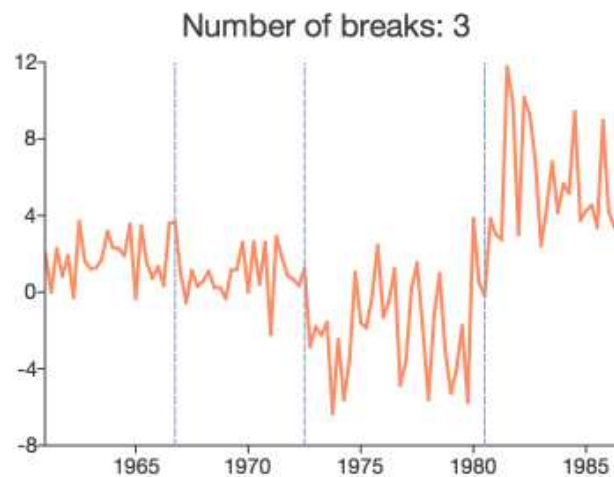
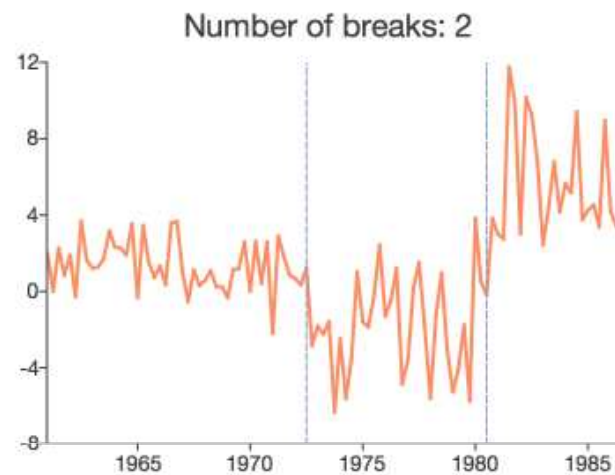
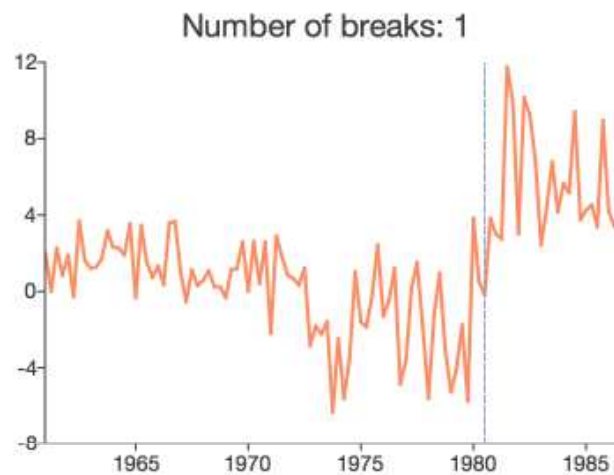
Средняя распространенность курения снизилась с 23,6% (в январе 2001 г.) до 17,3% (в июне 2011 г.).

- увеличение налогов на табачные изделия,
- усиление законов о запрете курения,
- повышенная интенсивность ежемесячных трансляций по телевидению кампаний в СМИ по борьбе против табака и рекламы фармацевтической компании, предлагающей никотинзаместительную терапию (НЗТ),
- обязательное размещение на пачках сигарет графических предупреждений о вреде курения.



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
УНИВЕРСИТЕТ «ВШЭ»  
(VSES)

# Анализ структурных сдвигов: несколько структурных сдвигов





## Анализ структурных сдвигов

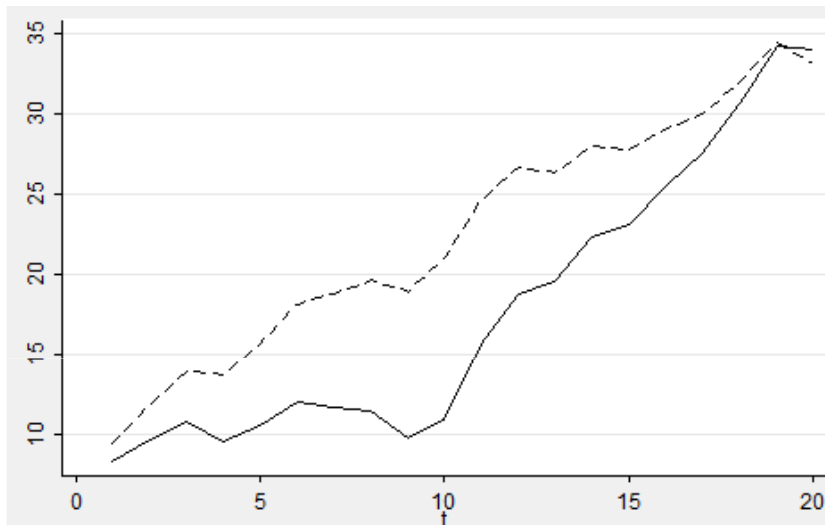
- Графический анализ
- **Тест на структурное изменение**
- Тест Чоу (Chow breakpoint test, Chow test)
- тест CUSUM (Харвея-Коллера (Harvey-Coller))

### **Анализ стационарности с учетом структурных сдвигов.**

- Тест Эндрюса-Зивота (Zivot, Andrews, 1992).
- Тест на несколько структурных сдвигов (Clemente, Montanes, Reyes, 1998)
- Тест Kapetanios (2005) до 5 структурных сдвигов

Kapetanios, G. (2005). Unit-root testing against the alternative hypothesis of up to m structural breaks. Journal of Time Series Analysis, 26(1), 123-133.

## Тест Чоу на структурное изменение (Chow breakpoint test)



$$y_t = x_t^T b + \varepsilon_t$$

$$y_t = x_t^T b_1 + \varepsilon_t, t \in S_1$$

$$y_t = x_t^T b_2 + \varepsilon_t, t \in S_2$$

$$d_t = \begin{cases} 1, t \in S_1 \\ 0, t \in S_2 \end{cases} \rightarrow y_t = x_t^T (d_t b_1 + (1 - d_t) b_2) + \varepsilon_t$$

$$H_0: b_1 = b_2$$

**Дата сдвига  
известна**

$$F = \frac{(ESS_R - ESS_1 - ESS_2) / k}{(ESS_1 + ESS_2) / (T - 2k)} \sim F(k, T - 2k)$$

ESS-сумма квадратов остатков,

k – количество оцениваемых параметров модели,

T – длина ВР

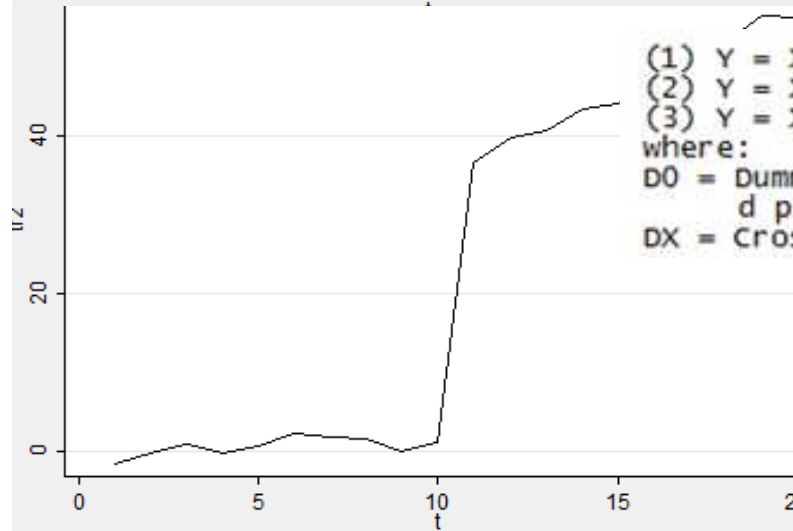
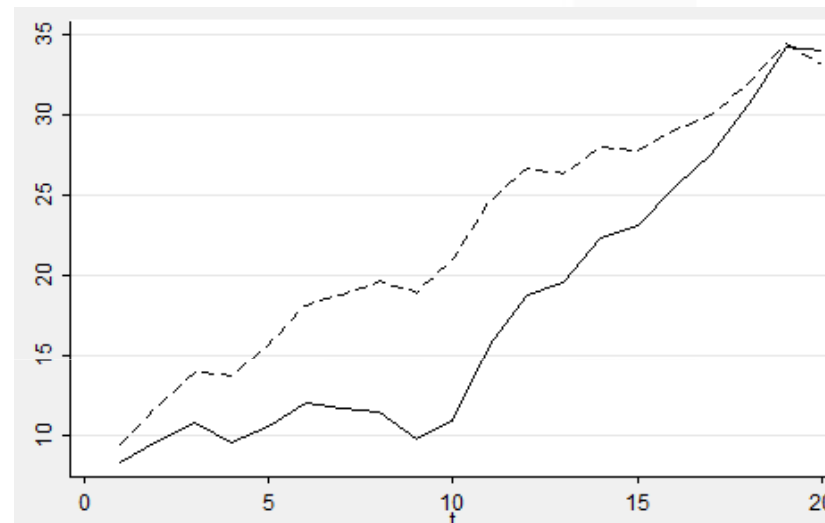
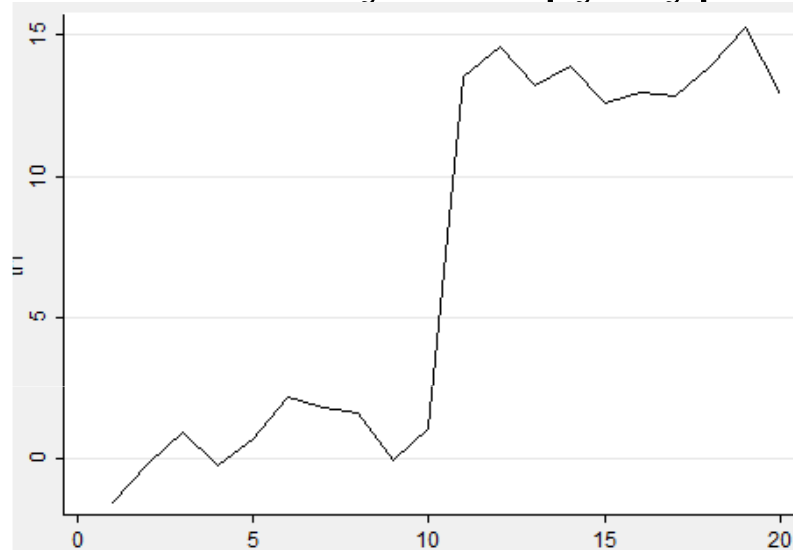




НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
УНИВЕРСИТЕТ

## Типы структурных сдвигов : Stata

### Тест Чоу на структурное изменение (Chow breakpoint test)



- (1)  $Y = X + D0$
- (2)  $Y = X + DX$
- (3)  $Y = X + D0 + DX$

where:

$D0$  = Dummy variable (0,1), takes (0) in first period, and (1) in second period.

$DX$  = Cross product of each  $x_i$  times in  $D0$

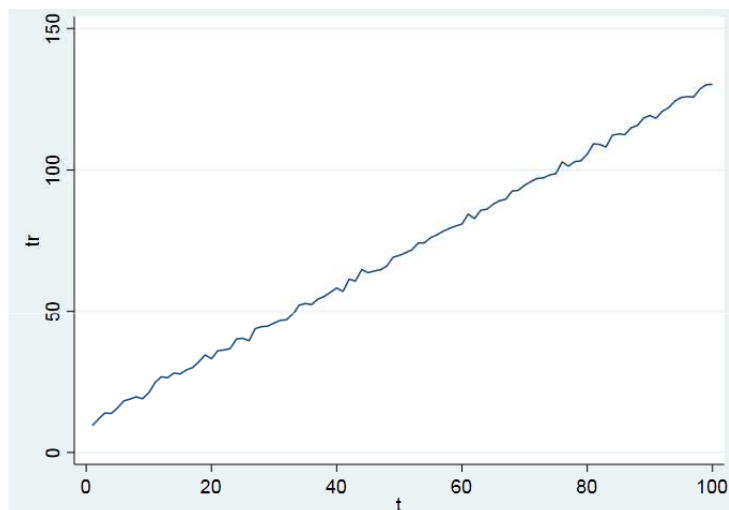
Пример моделирования и анализа сдвигов в Stata

<https://www.stata.com/support/faqs/statistics/computing-chow-statistic/>



# Тест Чой: Stata

chowreg tr t, dum(10) type(1)



Source	SS	df	MS	Number of obs	=	100
Model	121039.538	2	60519.7692	F(2, 97)	=	61359.69
Residual	95.6722167	97	.986311513	Prob > F	=	0.0000
				R-squared	=	0.9992
				Adj R-squared	=	0.9992
Total	121135.211	99	1223.58799	Root MSE	=	.99313

tr	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
t	1.203102	.0040269	298.77	0.000	1.19511	1.211094
D0	.3961573	.3874649	1.02	0.309	-.3728534	1.165168
_cons	9.509317	.3148359	30.20	0.000	8.884455	10.13418

( 1) D0 = 0

F( 1, 97) = 1.05  
Prob > F = 0.3091

Ho: no Structural Change

- N1: 1st Period Obs = 10  
- N2: 2nd Period Obs = 90  
- Chow Test = 1.0454 P-Value > F(1 , 97) 0.3091

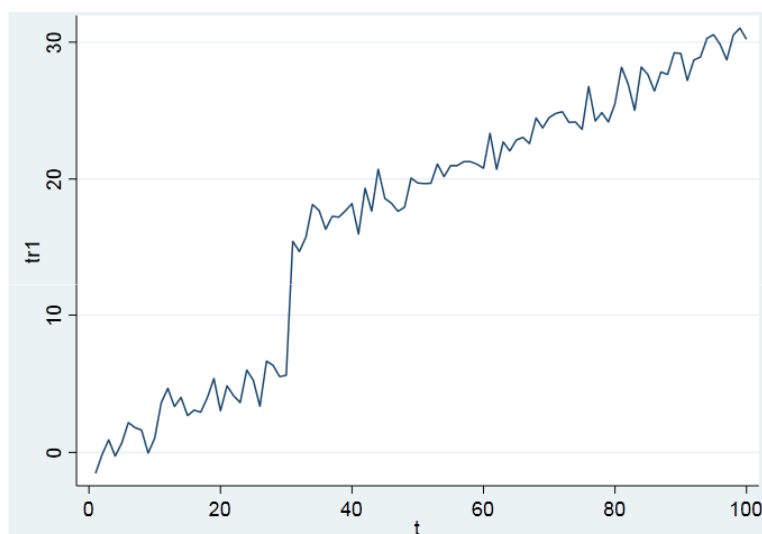




НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
УНИВЕРСИТЕТ УПРАВЛЕНИЯ

## Тест Чой: Stata

chowreg tr t, dum(30) type(1)



Source	SS	df	MS	Number of obs	=	100
Model	9716.7776	2	4858.3888	F(2, 97)	=	5138.98
Residual	91.7037498	97	.945399482	Prob > F	=	0.0000
Total	9808.48135	99	99.0755691	R-squared	=	0.9907
				Adj R-squared	=	0.9905
				Root MSE	=	.97232

tr1	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
t	.2153505	.005538	38.89	0.000	.2043591	.226342
D0	9.097784	.3488463	26.08	0.000	8.405421	9.790148
_cons	-.1911343	.1971846	-0.97	0.335	-.5824911	.2002226

Ho: no Structural Change

- N1: 1st Period Obs = 30
- N2: 2nd Period Obs = 70
- Chow Test = 680.1473 P-Value > F(1 , 97) 0.0000



### Тесты на структурное изменение

- Test for structural breaks with known break dates
- Test for a structural break with an unknown break date
- Wald and likelihood-ratio tests
- Robust to heteroskedasticity
- Cumulative sum (CUSUM) test for multiple breaks

<https://www.stata.com/features/overview/structural-breaks/>



## Моделирование структурных сдвигов

### -Фиктивные переменные

Пример: Айвазян С.А., Бродский Б.Е. и др. Макроэконометрическое моделирование экономик России и Армении // ПЭ. 2013. №3

### - Передаточные функции (transfer function) (Enders)

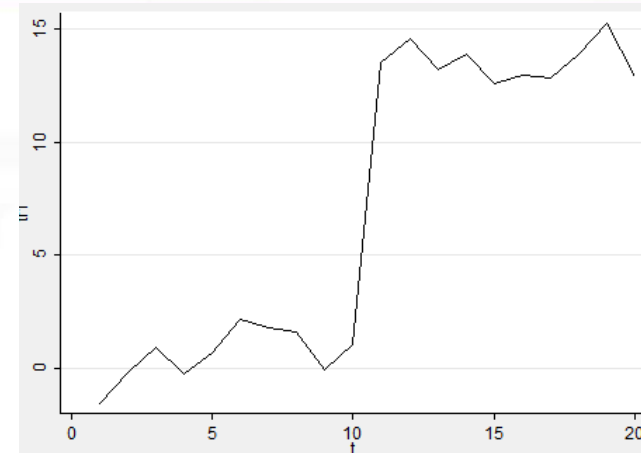
Пример: Clarke, H. D., W. Mishler and P. Whiteley. 1990. "Recapturing the Falklands: Models of Conservative Popularity, 1979-83." *British Journal of Political Science* 20:63-82.

# Моделирование структурных сдвигов: фиктивные переменные

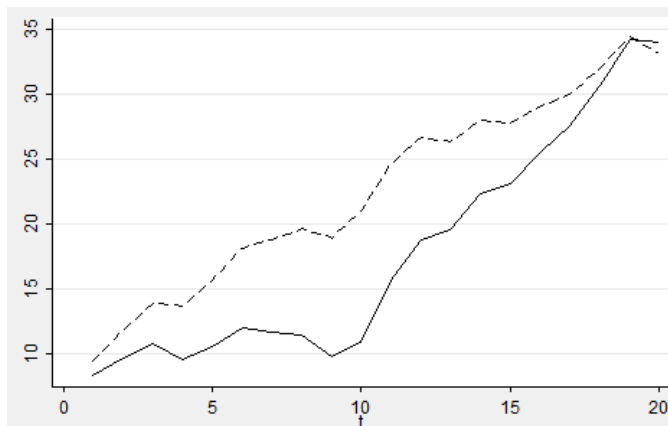
- Фиктивная переменная

$$d = \begin{cases} 0, & t < t^* \\ 1, & t \geq t^* \end{cases} \quad \hat{y}_t = b_0 + b_1 t + \boxed{\gamma d}$$

*Если  $d=1$ , меняется константа*



- Перекрестная дамми-переменная:  $d \cdot t$



$$\hat{y}_t = b_0 + b_1 t + \boxed{\gamma d t}$$

*Если  $d=1$ , меняется коэффициент при  $t$*

# Моделирование структурных сдвигов: пример



Индекс МосБиржи

$$d = \begin{cases} 0, & t < \text{май 2019} \\ 1, & t \geq \text{май 2019} \end{cases}$$

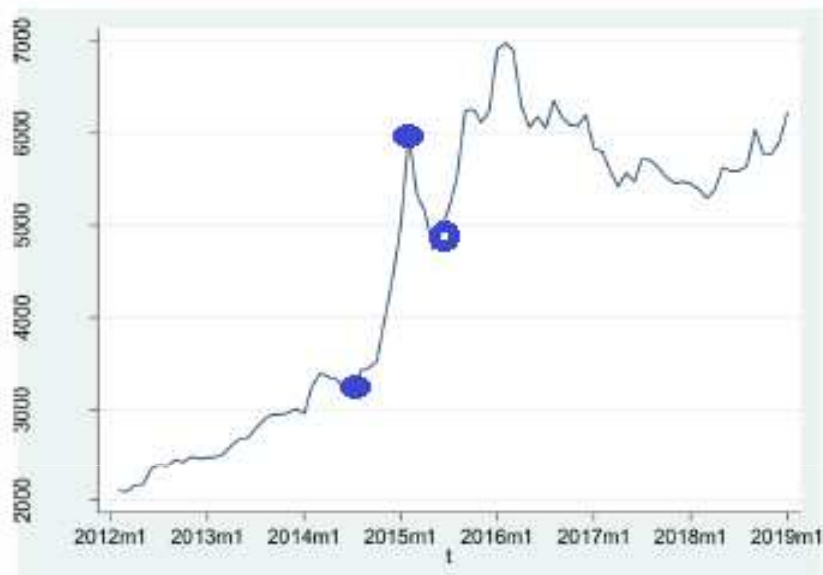
$$\hat{y}_t = b_0 + b_1 t + \gamma_1 d + \gamma_2 dt$$

Source	SS	df	MS	Number of obs =	253
Model	4922879.12	3	1640959.71	F( 3, 249) =	518.54
Residual	787980.273	249	3164.57941	Prob > F =	0.0000
				R-squared =	0.8620
				Adj R-squared =	0.8604
Total	5710859.39	252	22662.1405	Root MSE =	56.255

IMOEX	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
t	1.269665	.0962891	13.19	0.000	1.08002	1.45931
D0	172.1014	46.22868	3.72	0.000	81.05232	263.1505
Dx_t	-.2996196	.2376733	-1.26	0.209	-.767726	.1684868
_cons	2350.404	8.93649	263.01	0.000	2332.804	2368.005

Зачем нужна перекрестная дамми-переменная?

# Анализ и моделирование структурных сдвигов: пример



банковские вклады юридических и физических лиц в иностранной валюте и драгоценных металлах (млн.руб.)

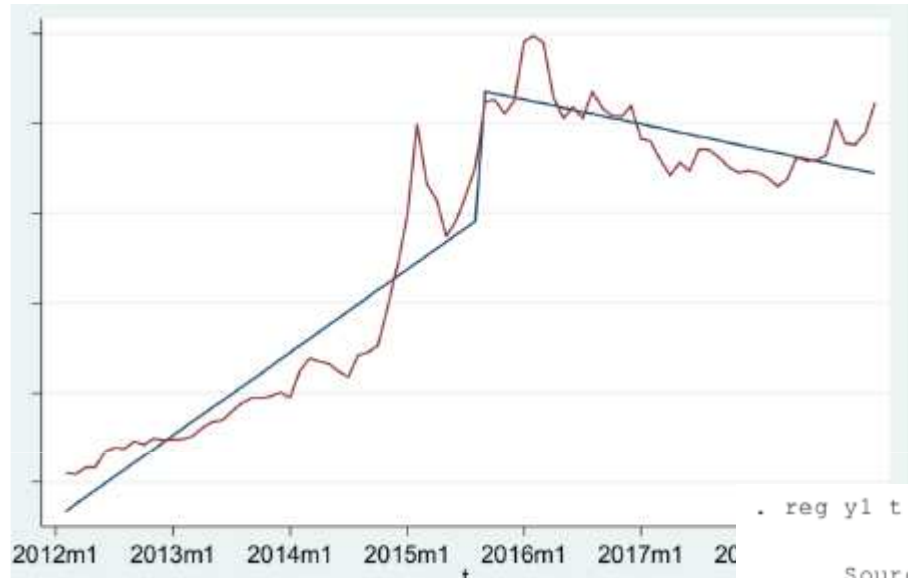
Точки структурного сдвига:  
2014m10,  
2015m1,  
2015m8

Тест Чоу для каждой точки:

Тест Чоу	Точка структурного сдвига	Результат	Знач.	Вывод
(1)	2015 m1	$R^2 = 0.88$ Есть структурный сдвиг типа 1	0.000	Тест Чоу для всех моделей показал наличие структурных сдвигов соответствующих типов в точке 2015 m1. Наилучшие результаты дала модель со скачком и изменением угла наклона, что можно было предположить на основании графика зависимой переменной.
(2)	2015 m1	$R^2 = 0.76$ Есть структурный сдвиг типа 2 Незначима переменная t1	0.001	
(3)	2015 m1	$R^2 = 0.932$ Есть структурный сдвиг типа 3	0.000	

Наилучшие результаты показала модель, в которой предполагалось, что структурный сдвиг произошел в момент времени 2015m8 – ее объясняющая способность составила 93,4 %.

# Анализ и моделирование структурных сдвигов: пример



$$d = \begin{cases} 0, & t < 43 \\ 1, & t \geq 43 \end{cases}$$

$$\hat{y}_t = b_0 + b_1 t + \gamma_1 d + \gamma_2 dt$$

банковские вклады юридических и физических лиц в иностранной валюте и драгоценных металлах (млн.руб.)

```
. reg y1 t t43 tt43
```

Source	SS	df	MS
Model	184826268	3	61608756.1
Residual	12508595.5	80	156357.444
Total	197334864	83	2377528.48

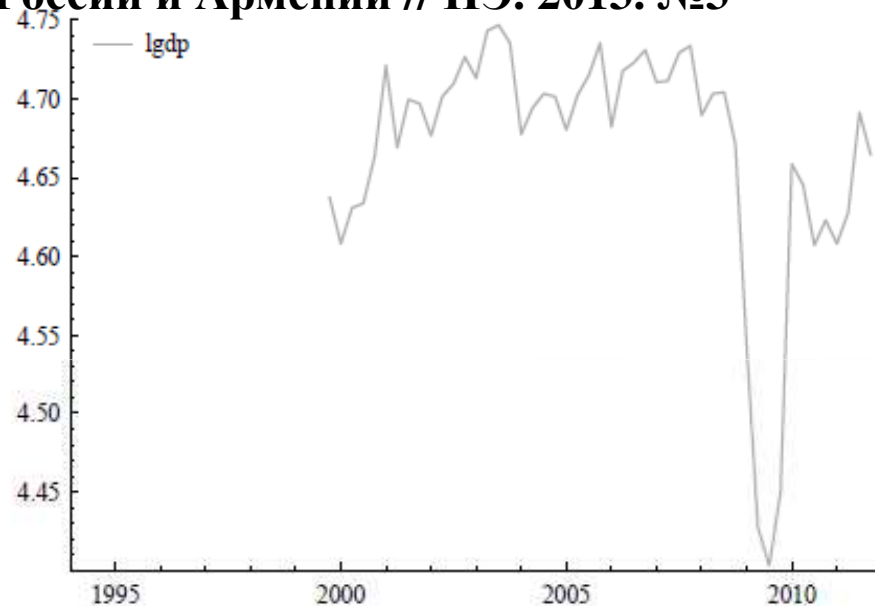
Number of obs = 84  
F( 3, 80) = 394.03  
Prob > F = 0.0000  
R-squared = 0.9366  
Adj R-squared = 0.9342  
Root MSE = 395.42

y1	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
t	-22.82867	5.219193	-4.37	0.000	-33.21519 -12.44214
t43	-5762.289	361.1826	-15.95	0.000	-6481.065 -5043.513
tt43	99.9003	7.131044	14.01	0.000	85.70907 114.0915
_cons	21608.74	3591.336	6.02	0.000	14461.76 28755.73

Анализ остатков?

# Моделирование структурных сдвигов: фиктивные переменные

**Пример: Айвазян С.А., Бродский Б.Е. и др. Макроэконометрическое моделирование экономик России и Армении // ПЭ. 2013. №3**



**Рис. 3. Квартальная динамика ВВП Армении  
в период 1-й квартал 2000 г. — 4-й квартал 2011 г.**

$$D\ln(GDP_t) = -0.011_{(-2.31)} - 0.312_{(-3.78)} R\ln(GDP_{t-1}) + 0.174_{(6.39)} m_t - 0.111_{(-3.98)} i2009p2_t,$$

где  $m_t = M2_t / M2_{t-1} - 1$  — темп роста денежного агрегата M2;  $i2009p2$  — фиктивная переменная (индикатор 2-го квартала 2009 года), отражающая влияние мирового финансового кризиса на динамику ВВП РА.





# Методы выделения неслучайной составляющей: Как выделить трендовую составляющую?

## аналитические

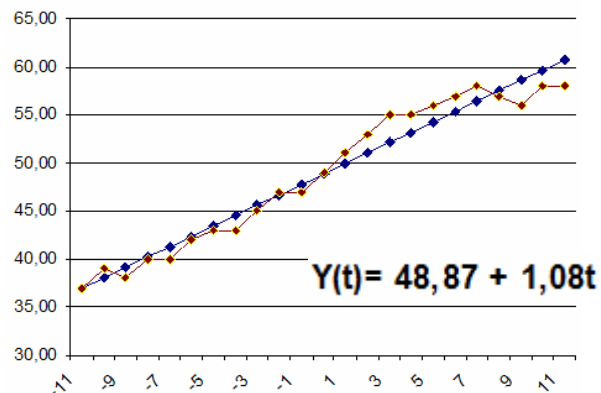
**Предположение:** известен общий вид неслучайной составляющей

$$y_t = f(t) + \varepsilon_t,$$

$$f(t) = \theta_0 + \theta_1 t$$

**Необходимо** найти  $\hat{\theta}_1, \hat{\theta}_0$   
для параметров  $\theta_1, \theta_0$

**Реализация:** модели регрессии

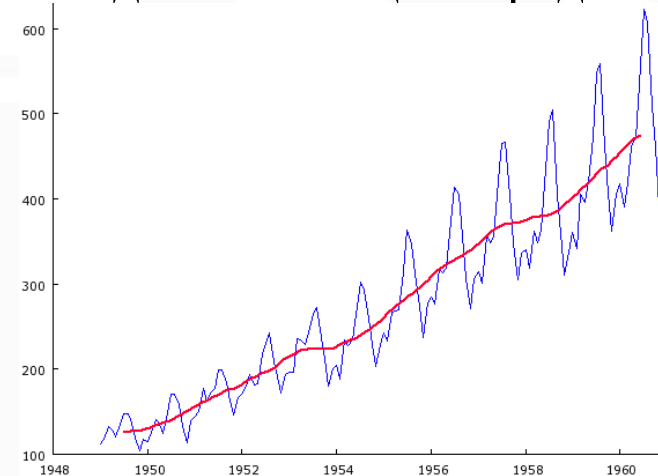


## алгоритмические

Общий аналитический вид функции неизвестен, на «выходе» - алгоритм расчета оценки  $\hat{f}(t)$  для искомой функции  $f(t)$  в точке  $t$ .

**Реализация:**

методы скользящего среднего

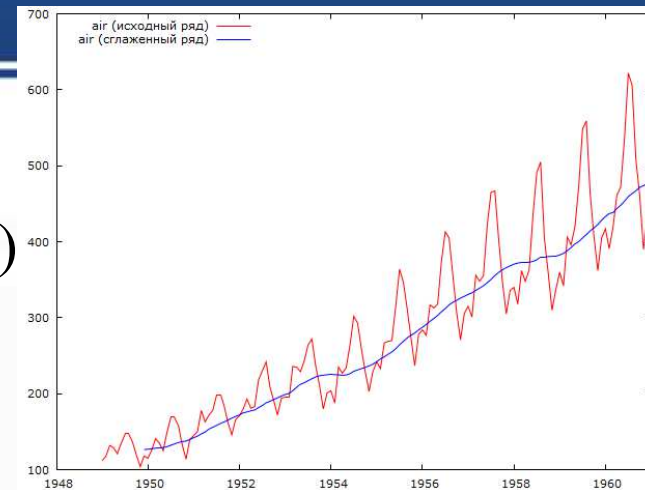


# Выделение трендовой составляющей: основные подходы

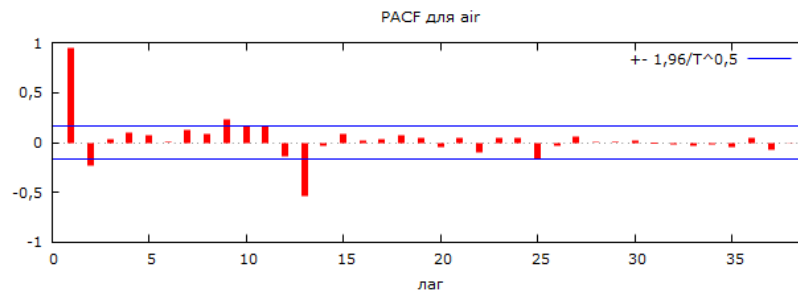
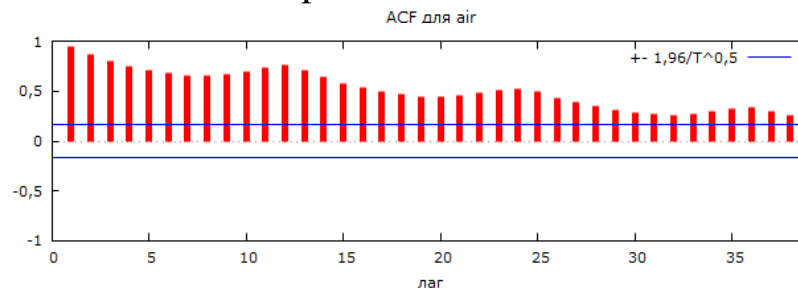
## 1. Скользящие средние (разных типов: простые, взвешенные и т.д)

См. Дуброва Т.А. Статистические методы прогнозирования 2003

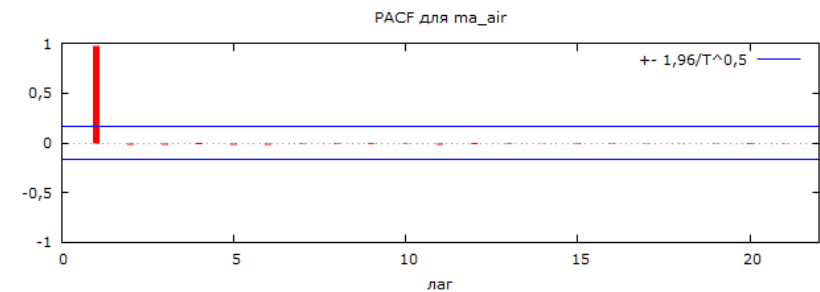
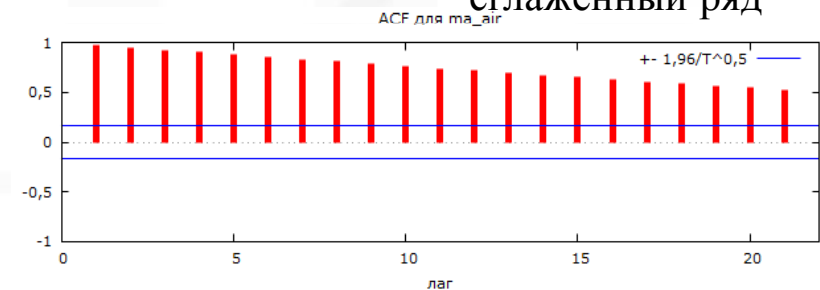
<http://www.twirpx.com/file/18580/>



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.

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

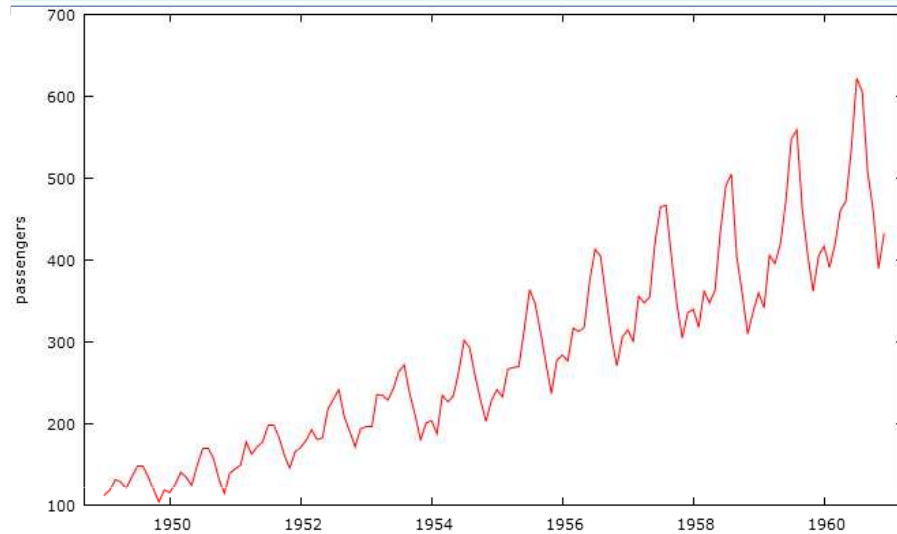
$$\hat{y}_t = \frac{\frac{1}{2}y_{t-6} + y_{t-5} + \dots + y_t + \dots + y_{t+5} + \frac{1}{2}y_{t+6}}{12}$$

## 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)

# Выделение трендовой составляющей: 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*.

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	***



## Выделение трендовой составляющей: основные подходы

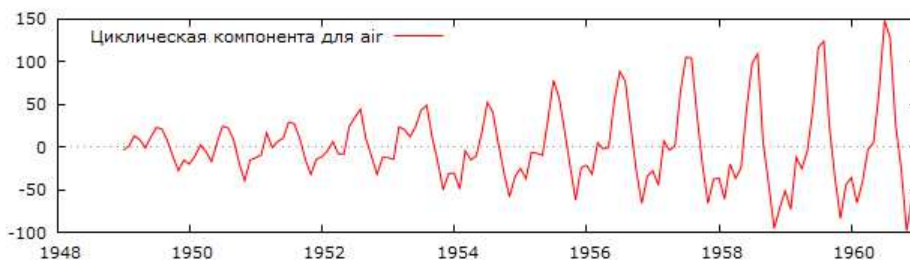
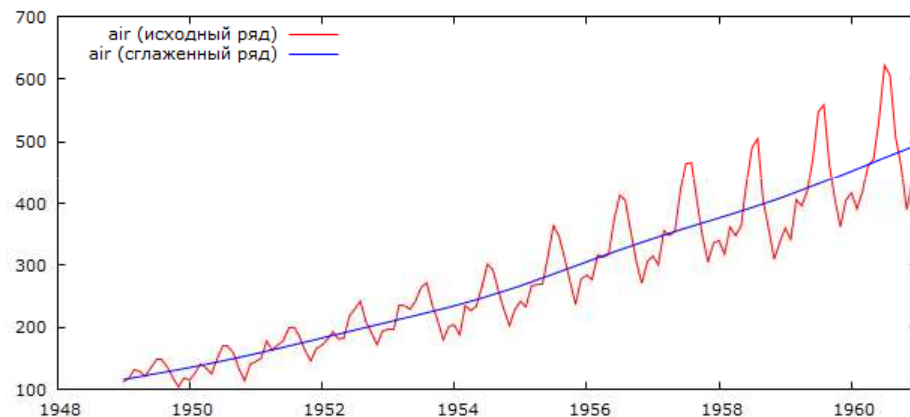
### 3. Использование фильтров – фильтр Ходрика-Прескотта

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

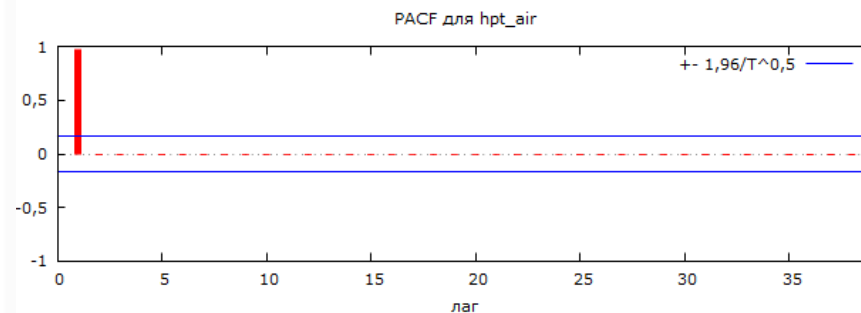
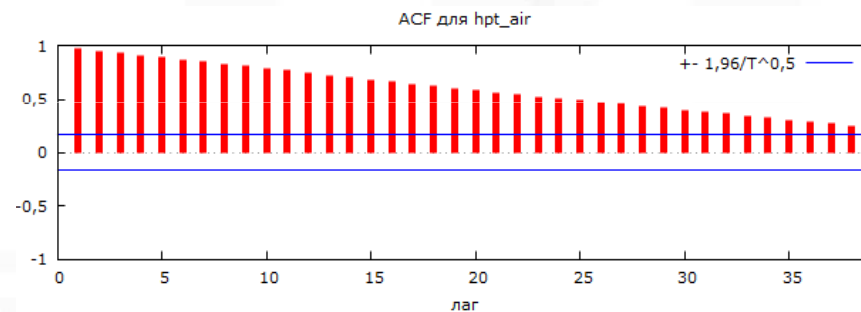
$\lambda=100$  – годовые данные

$\lambda=1600$  – квартальные данные

$\lambda=14400$  – месячные данные



#### сглаженный ряд



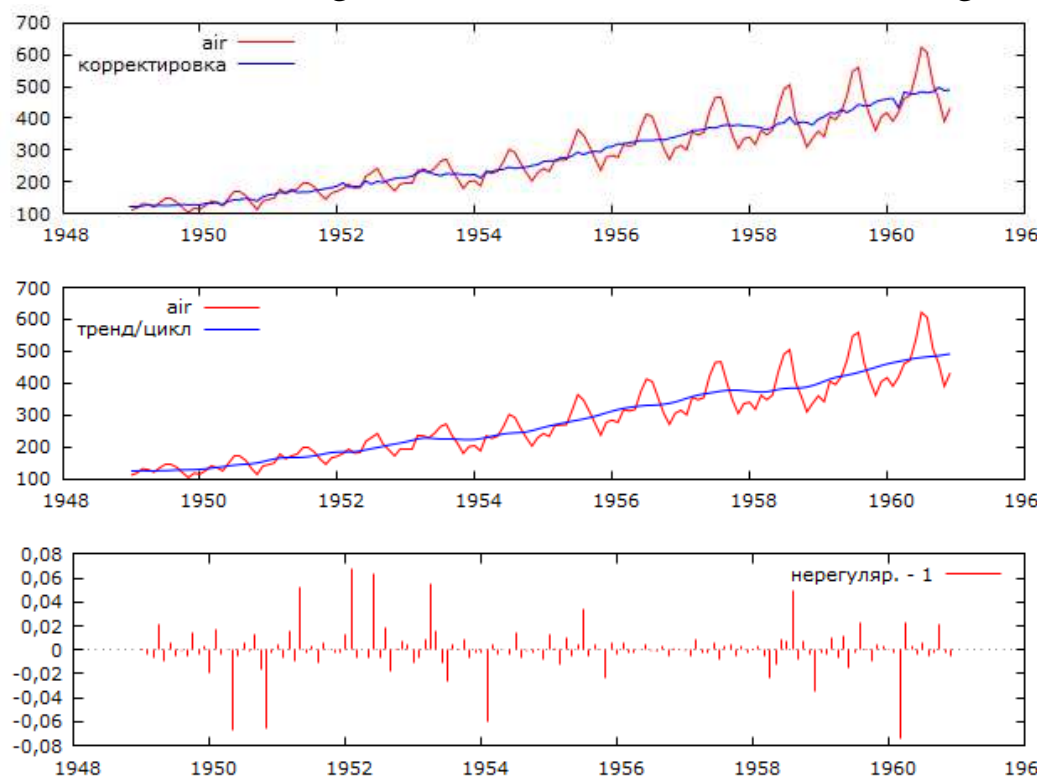
Hodrick, Robert; Prescott, Edward C. (1997). "Postwar U.S. Business Cycles: An Empirical Investigation". Journal of Money, Credit, and Banking. 29 (1): 1–16.

# Выделение трендовой составляющей: основные подходы

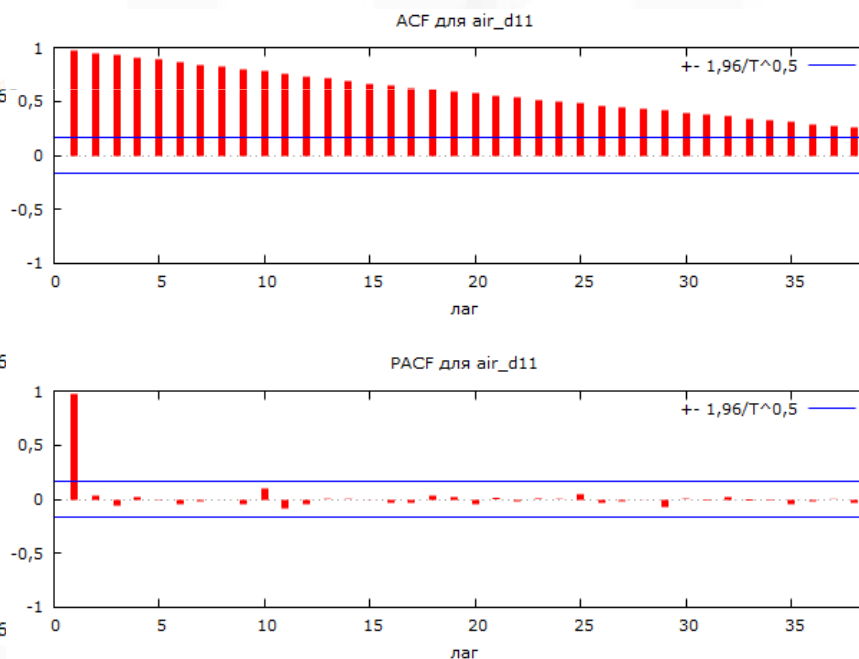
## 4. U.S. Census Bureau's EuroSTAT

## X-12-ARIMA TRAMO

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



сглаженный ряд



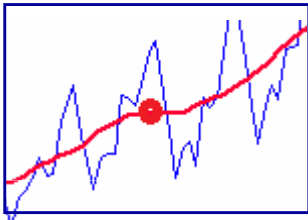
- Для устранения *сезонных* колебаний часто требуется использовать 4- и 12-членные скользящие средние.
- При четном числе уровней принято первое и последнее наблюдение на активном участке брать с половинными весами

### Пример

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

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

- Общая идея подхода: уменьшение дисперсии, сглаживание траектории
- Алгоритм



$$y_t : \mu \rightarrow \sigma^2$$

$$(y_1 + y_2 + \dots + y_k) / k : \mu \rightarrow \sigma^2 / k$$

$$\hat{f}(t) = \sum_{k=-m}^m w_k y_{t+k}$$

- Определение весовых коэффициентов, теорема Вейерштрасса

«любая гладкая функция при самых общих допущениях может быть локально (т.е. в ограниченном интервале изменения ее аргумента  $t$ ) представлена алгебраическим полиномом подходящей степени»



# Сглаживание по простой скользящей средней (Moving Average (MA))

$$1. f(t) = a_0 + a_1 t$$

2. Задаем интервал сглаживания и берем первые  $(2m+1)$  членов ВР:  $t=1,2,\dots,(2m+1)$ ;  $t=m+1$  – средняя точка отрезка

3. Преобразование времени:  $t' = t - (m+1) \rightarrow t' = -m, -(m-1), \dots, 0, \dots, (m-1), m$

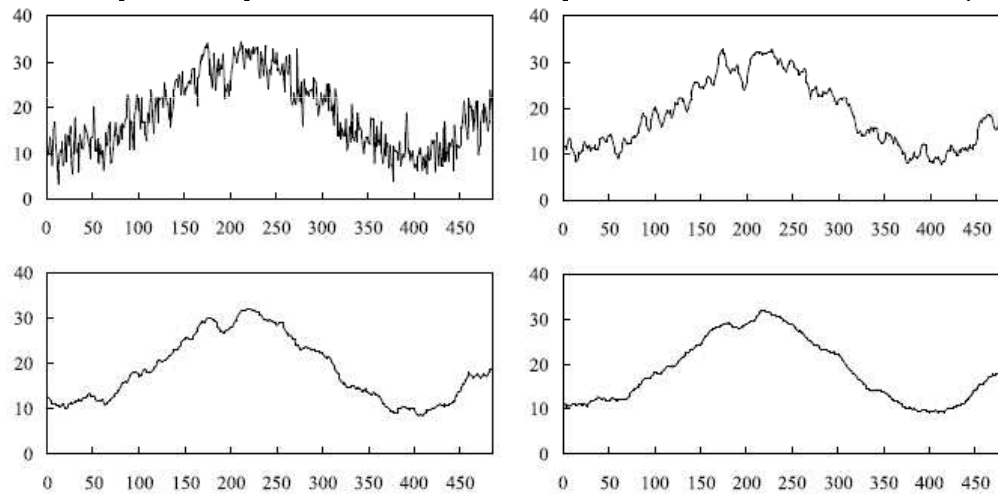


Figure 1.4 Maximum temperature data and its moving average. Top left: original data, top right: moving average with  $k = 5$ , bottom left:  $k = 17$ , bottom right:  $k = 29$ .

Genshiro Kitagawa

4. МНК

$$\sum_{t'=-m}^m (y_{t'} - a_0 - a_1 t')^2 \rightarrow \min_{a_0, a_1}$$

$$\hat{f}(t) = \hat{y}_t = \frac{1}{2m+1} \sum_{k=-m}^m y_{t+k}$$



# Сглаживание с помощью взвешенной скользящей средней

1.  $f(t) = a_0 + a_1 t + a_2 t^2$

*Взвешенная* скользящая средняя  
(Weighted Moving Average WMA)

2. Задаем интервал сглаживания ( $k=5$  точек)

3. МНК  $\sum_{t'=-2}^2 (y_{t'} - a_0 - a_1 t' - a_2 (t')^2)^2 \rightarrow \min_{a_0, a_1, a_2}$

$$\hat{f}(t) = \hat{a}_0 + \hat{a}_1 t' + \hat{a}_2 t'^2 \Big|_{t'=0} = \hat{a}_0 = \frac{1}{35} [-3y_{-2} + 12y_{-1} + 17y_0 + 12y_1 - 3y_2]$$

Символическая запись:  $\frac{1}{35} [-3, 12, 17]$

**Свойства весовых коэффициентов:**

- 1) Симметричность относительно центрального уровня.
- 2) Сумма весов с учетом общего множителя, вынесенного за скобки, равна единице.

## Примеры таблиц весовых коэффициентов

### Весовые коэффициенты при сглаживании по полиномам второго и третьего порядка

Длина интервала Сглаживания	Весовые коэффициенты
5	$\frac{1}{35}[-3, +12, +17]$
7	$\frac{1}{21}[-2, +3, +6, +7]$
9	$\frac{1}{231}[-21, +14, +39, +54, +59]$
11	$\frac{1}{429}[-36, +9, +44, +69, +84, +89]$
13	$\frac{1}{143}[-11, 0, +9, +16, +21, +24, +25]$

Сглаживание по параболе 2-го порядка по 7-членной взвешенной скользящей средней:

$$\hat{y}_i = \frac{1}{21} \cdot (-2y_{t-3} + 3y_{t-2} + 6y_{t-1} + 7y_t + 6y_{t+1} + 3y_{t+2} - 2y_{t+3})$$



## Весовые коэффициенты

*Второй и третий порядок*

$$[5] \quad \frac{1}{35} [-3, 12, 17],$$

$$[7] \quad \frac{1}{21} [-2, 3, 6, 7],$$

$$[9] \quad \frac{1}{231} [-21, 14, 39, 54, 59],$$

$$[11] \quad \frac{1}{429} [-36, 9, 44, 69, 84, 89],$$

$$[13] \quad \frac{1}{143} [-11, 0, 9, 16, 21, 24, 25],$$

$$[15] \quad \frac{1}{1105} [-78, -13, 42, 87, 122, 147, 162, 167],$$

$$[17] \quad \frac{1}{323} [-21, -6, 7, 18, 27, 34, 29, 42, 43],$$

$$[19] \quad \frac{1}{2261} [-136, -51, 24, 89, 144, 189, 224, 249, 264, 269],$$

$$[21] \quad \frac{1}{3059} [-171, -76, 9, 84, 149, 204, 249, 284, 309, 324, 329].$$

Кендалл, Стюарт, с.509

## Задание

Пусть длина интервала сглаживания  $k=7$ , локальное поведение сглаженного временного ряда внутри каждого активного участка описывается с помощью *полинома 2-го и 3-го порядка*.

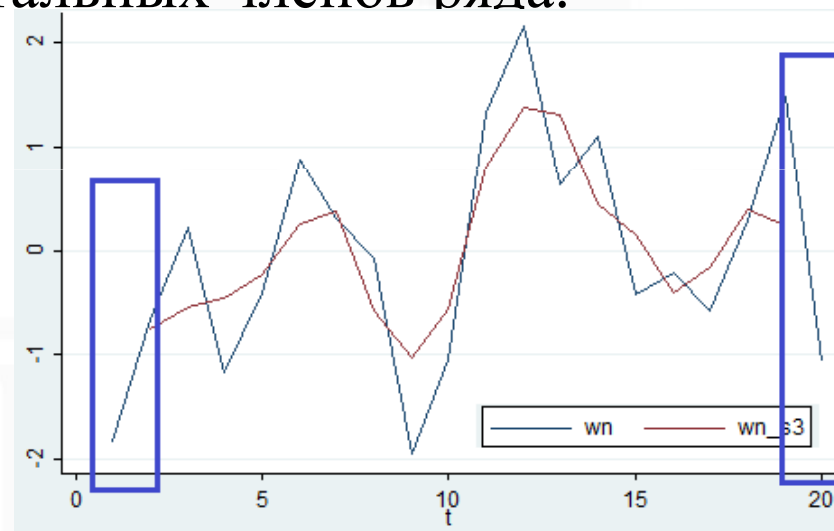
Вывести значения весовых коэффициентов для 7-членной взвешенной скользящей средней

$$\hat{y}_i = \frac{1}{21} \cdot (-2y_{t-3} + 3y_{t-2} + 6y_{t-1} + 7y_t + 6y_{t+1} + 3y_{t+2} - 2y_{t-3})$$

Показать, что значения весовых коэффициентов для 7-членной взвешенной скользящей средней для полиномов 2-го и 3-го порядка не будут отличаться. Почему?

- Для определения  $m$  первых и  $m$  последних потерянных уровней анализируемого ВР используют расчетные значения, полученные с помощью аппроксимирующих полиномов той же степени, что и для сглаживания остальных членов ряда.

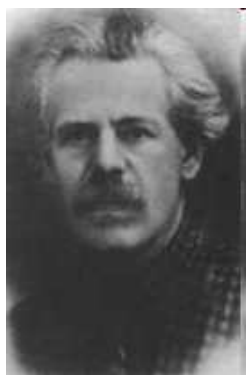
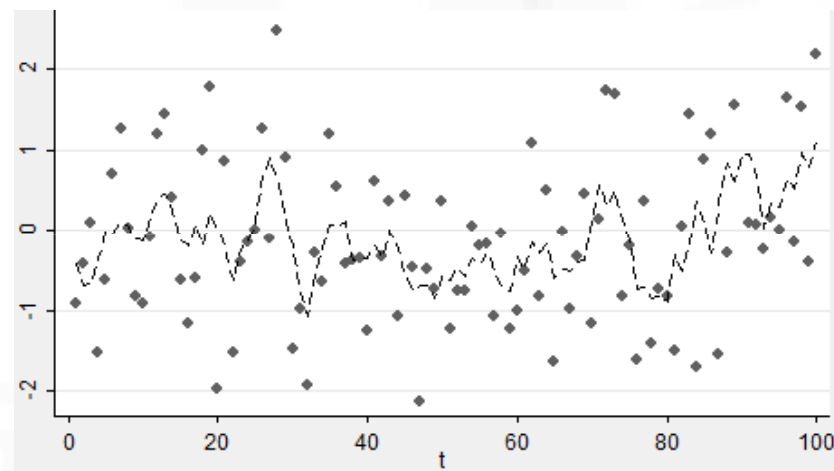
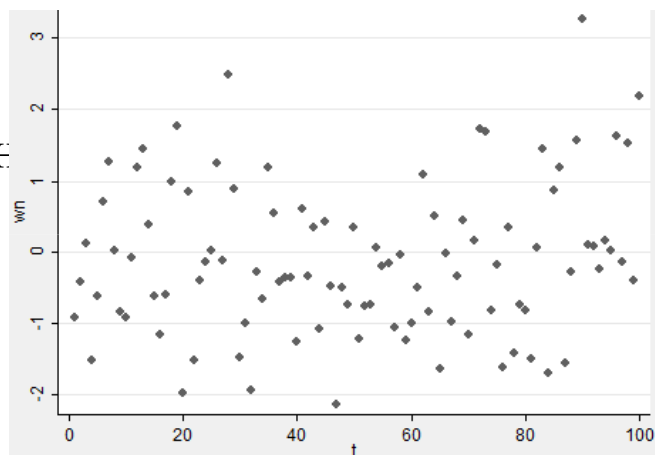
**Пример:**  $\hat{f}(t) = \hat{a}_0 + \hat{a}_1 t' + \hat{a}_2 t'^2 \Big|_{t'=1}$   
 $\hat{f}(t) = \hat{a}_0 + \hat{a}_1 t' + \hat{a}_2 t'^2 \Big|_{t'=-1}$



# Влияние скользящего усреднения на остаточную компоненту

**Эффект Слуцкого-Юла (1927):** Ряд скользящих средних может содержать циклические компоненты, отсутствующие в исходном ряде.

Док-во:  
См. Айвазян



Евгений  
Евгеньевич  
Слуцкий (1880-1948)

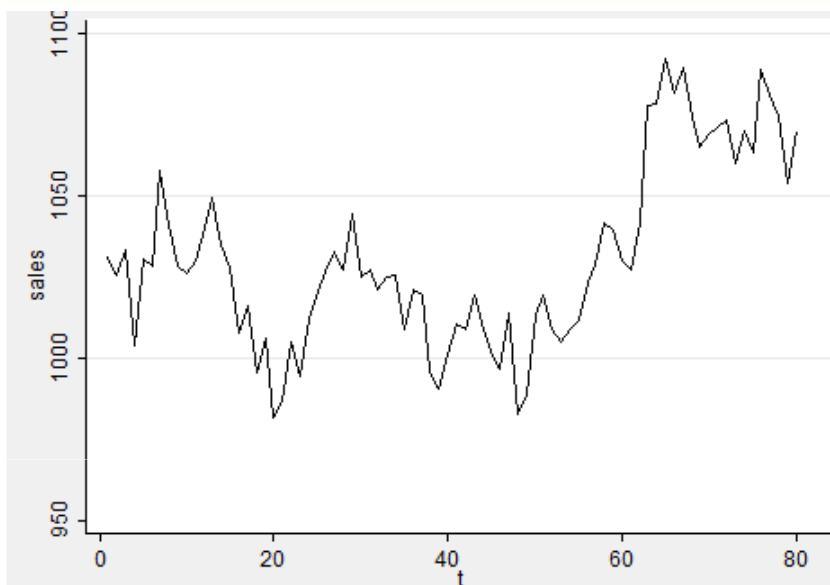
## Эксперимент Слуцкого:

[http://www.minneapolisfed.org/publications\\_papers/pub\\_display.cfm?id=4348](http://www.minneapolisfed.org/publications_papers/pub_display.cfm?id=4348)

Е.Е.Слуцкий. Сложение случайных величин как источник циклических процессов// Вопросы конъюнктуры. 1927

Yule G. On a method of investigating periodicities in disturbed series, with special reference to Wolfer's sunspot numbers. Phil., 1927.

# Восстановление пропущенных наблюдений в середине выборки



use <http://www.stata-press.com/data/r13/sales1>, clear  
tssmooth exponential sm1=sales, parms(.7) forecast(3)  
generate sales2=sales if t!=28  
tssmooth exponential sm3=sales2, parms(.7) forecast(3)

```
exponential coefficient = 0.7000
sum-of-squared residuals = 6842.4
root mean squared error = 11.817
. list t sales2 sm3 if t>25 & t<31
```

	t	sales2	sm3
26.	26	1011.5	1007.5
27.	27	1028.3	1010.3
28.	28	.	1022.9
29.	29	1028.4	1022.9
30.	30	1054.8	1026.75

$$S_{29} = \alpha S_{28} + (1 - \alpha) S_{28} = S_{28}$$



## Приложения скользящего среднего

1. Метод выделения тренда.
2. В техническом анализе, в качестве самостоятельного **технического индикатора** либо в составе других инструментов.
3. В технике, в качестве **фильтра** (используется для обработки сигналов и изображений, системах автоматического управления и для других прикладных целей).

### Стратегия форекс - RSI Profit Taking

