

Séries Temporais

Mestrado em Econometria Aplicada e Previsão

Jorge Caiado

CEMAPRE/ISEG, Universidade de Lisboa

Email: jcaiado@iseg.utl.pt

Web: <http://pascal.iseg.utl.pt/~jcaiado/>



LISBOA
SCHOOL OF
ECONOMICS &
MANAGEMENT

Introdução aos métodos de previsão

Conceitos e objectivos

Uma **série temporal** (*time series*) consiste num conjunto de observações de uma variável, feitas em períodos sucessivos de tempo, durante um determinado intervalo e representa-se por Y_t , $t = 1, 2, \dots, n$. São exemplos, as cotações diárias das acções, as vendas semanais de um dado produto financeiro, o número mensal de dormidas na hotelaria, as despesas públicas trimestrais do país, os lucros anuais de uma empresa, as temperaturas mínimas, médias e máximas diárias.

A representação gráfica de uma série temporal designa-se por **cronograma** e constitui o ponto de partida para a sua análise.

Introdução aos métodos de previsão

Conceitos e objetivos

Na análise de séries temporais devem ter-se em conta os seguintes **objetivos** fundamentais:

Descrição. Construção do cronograma da série e caracterização do seu andamento geral, procurando identificar os pontos de viragem (mudança de estrutura) e eventuais observações anómalas (*outliers*).

Explicação ou Modelação. Construção de modelos que permitam explicar o comportamento da série no período observado.

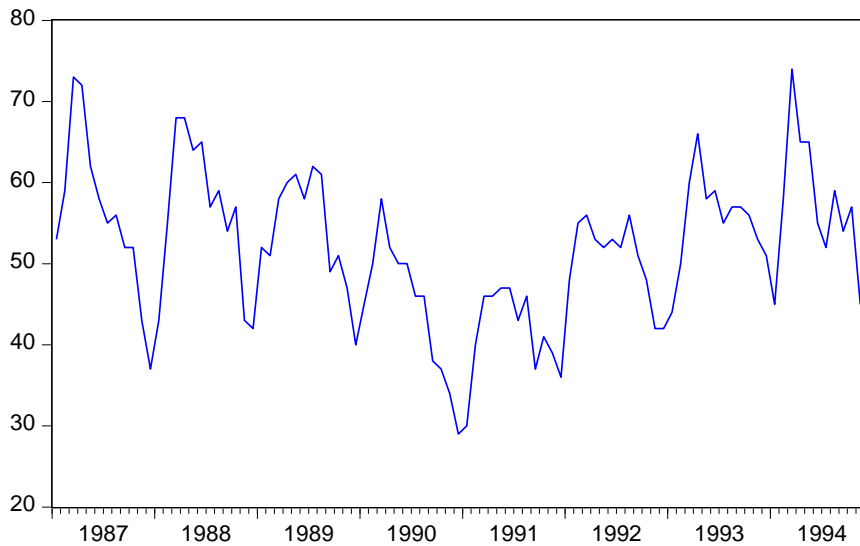
Previsão. Tentar prever a evolução futura da série com base exclusivamente no seu passado (modelos univariados ou não-causais) ou com base no comportamento passado de outras variáveis (modelos multivariados).

Controlo. Procurar modificar o comportamento futuro do processo através do ajustamento de variáveis controláveis. Por exemplo: numa linha de fabrico e montagem de automóveis, é possível prever o número de viaturas produzidas com base nas matérias-primas e mão-de-obra utilizadas na produção.

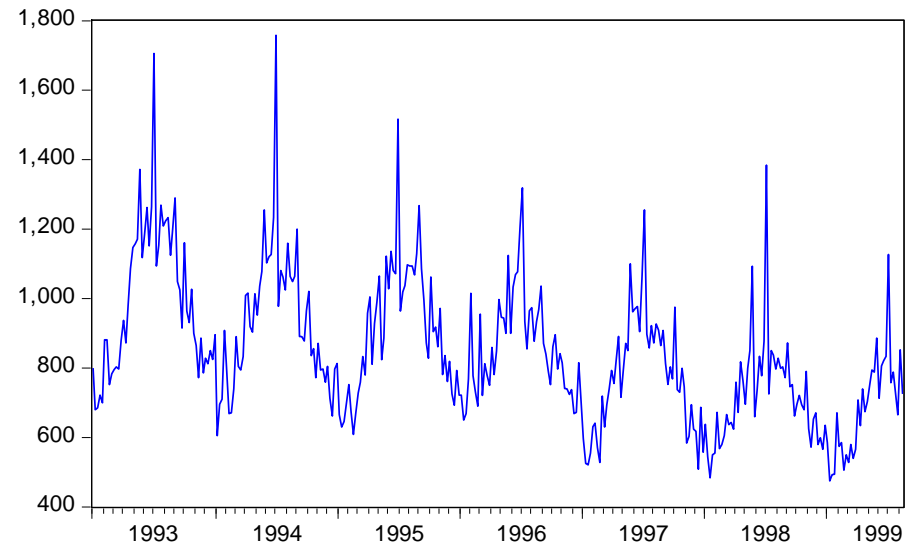
Introdução aos métodos de previsão

Conceitos e objetivos

HOUSE



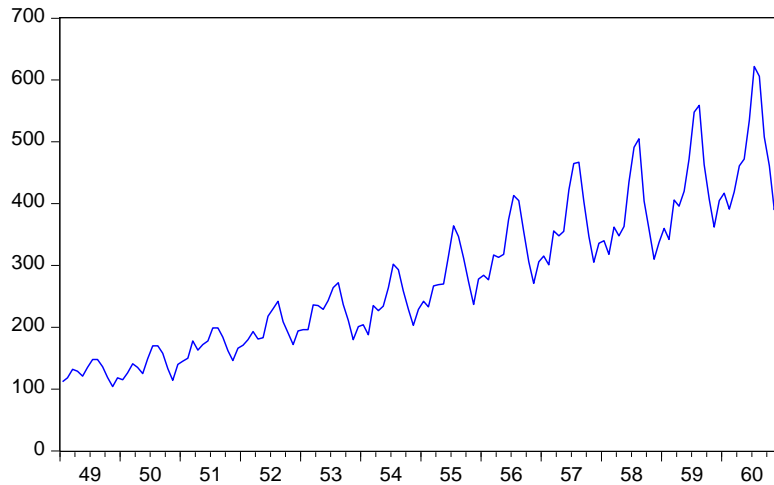
BEER



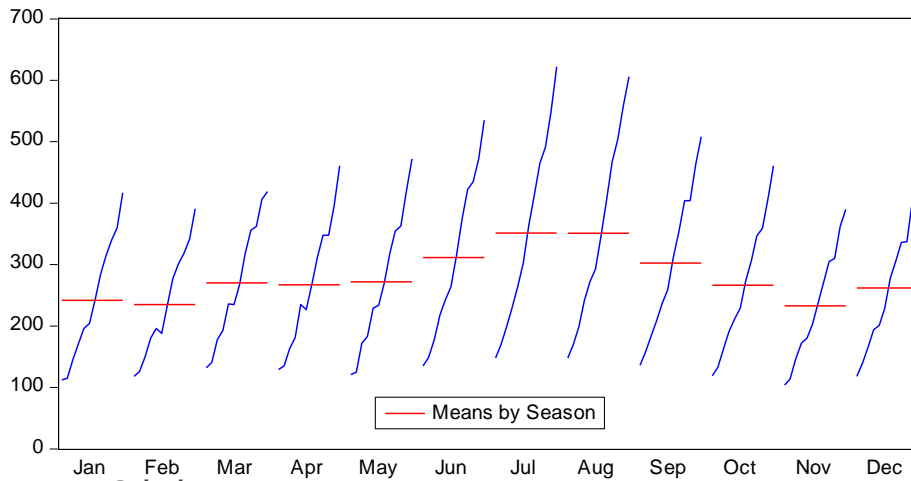
Introdução aos métodos de previsão

Conceitos e objetivos

PASSAG



PASSAG by Season

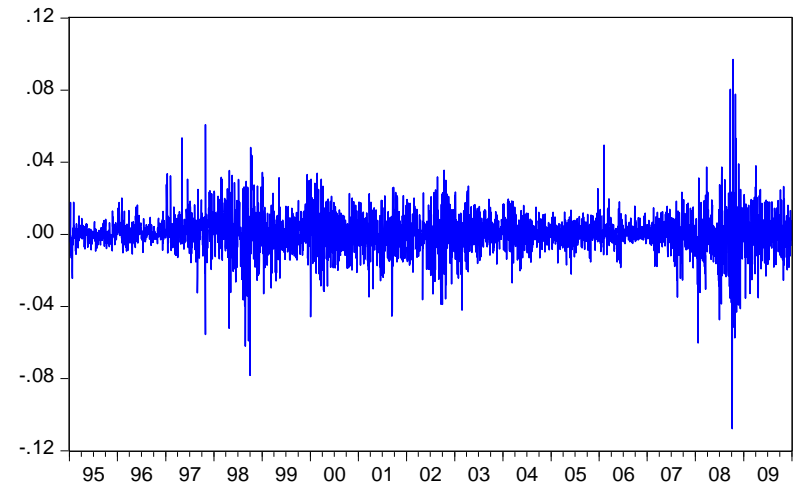


Jorge Caiado

POR



Log Differenced POR



Introdução aos métodos de previsão

Decomposição de séries temporais

Os métodos tradicionais de análise de séries temporais baseiam-se na decomposição de um conjunto de efeitos ou forças componentes, a saber: tendência, movimentos oscilatórios ou cíclicos, sazonalidade e movimentos irregulares ou aleatórios.

A **tendência** caracteriza o andamento mais notório da série durante um longo período de tempo. Os **movimentos oscilatórios ou cíclicos** estão associados às fases de expansão e recessão dos sistemas económicos. Em ciclos longos, as componentes de tendência e cíclica são difíceis de separar, pelo que se pode tomar estas como uma única componente (tendência-cíclica).

Introdução aos métodos de previsão

Decomposição de séries temporais

A **sazonalidade** refere-se às oscilações periódicas que ocorrem semanalmente, mensalmente ou trimestralmente no decurso do ano. Estas oscilações no comportamento da série podem estar associadas com as estações do ano (temperatura do ar, consumo de água, consumo de eletricidade, turismo), com medidas administrativas (início e fim do ano escolar), com tradições e costumes sociais ou culturais (por exemplo, o aumento das vendas no período natalício) ou com as variações do calendário (número de dias úteis do mês ou semana, número de sábados no mês).

Por último, os **movimentos irregulares ou aleatórios** são os movimentos da série que não são explicados pelas componentes anteriormente referidas.

Introdução aos métodos de previsão

Decomposição de séries temporais

O modelo de decomposição pode traduzir-se pela seguinte expressão:

$$Y_t = f(T_t, C_t, S_t, E_t),$$

onde a série Y_t é função da componente de tendência (*trend*) T_t , da componente cíclica (*cycle*) C_t , da componente sazonal (*seasonal*) S_t e da componente irregular (*irregular*) E_t . Como se referiu anteriormente, alguns autores não separam a componente de tendência da componente cíclica por estas representarem, em conjunto, o comportamento de longo prazo da série, dando lugar à expressão, $Y_t = f(TC_t, S_t, E_t)$, onde TC_t representa a componente de tendência-cíclica (*trend-cycle*).

As formas funcionais mais correntemente utilizadas são o modelo multiplicativo e o modelo aditivo. O **modelo multiplicativo** descreve a série como o produto das componentes tendência, sazonalidade, cíclica e irregular:

$$Y_t = T_t \times C_t \times S_t \times E_t.$$

Introdução aos métodos de previsão

Decomposição de séries temporais

O modelo multiplicativo é adequado a situações em que a amplitude da componente sazonal aumenta ou diminui com a tendência da série, o que acontece na maioria das séries económicas. Contudo, este método não pode ser implementado para séries temporais de valores negativos ou nulos.

Quando as oscilações de carácter sazonal não variam com o nível da série, podemos utilizar o **modelo aditivo**. Segundo o modelo aditivo, a série observada resulta da soma das componentes tendência, sazonalidade, cíclica e irregular, através da expressão:

$$Y_t = T_t + C_t + S_t + E_t.$$

Introdução aos métodos de previsão

Avaliação dos erros de previsão

Erro quadrático médio. Traduz o valor médio dos desvios ao quadrado entre os valores observados e as previsões para os instantes 1, 2, ..., m :

$$EQM = \frac{1}{m} \sum_{t=1}^m (Y_t - P_t)^2$$

Erro absoluto médio. Traduz o valor absoluto médio dos desvios entre os valores observados e as previsões para os instantes 1, 2, ..., m :

$$EAM = \frac{1}{m} \sum_{t=1}^m |Y_t - P_t|$$

Erro percentual absoluto médio. Traduz o valor percentual absoluto médio dos desvios entre os valores observados e as previsões para os instantes 1, 2, ..., m :

$$EPAM = \frac{1}{m} \sum_{t=1}^m \left| \frac{(Y_t - P_t)}{Y_t} \right| \times 100$$

Métodos de Alisamento Exponencial

Alisamento exponencial simples

Considere-se uma série de dados observados até ao instante $t-1$ e que se pretende obter uma previsão pontual para o instante t . Seja Y_t o valor da série no instante t e P_t a previsão obtida no mesmo instante. Quando for conhecida a observação Y_t , o erro de previsão vem dado por $Y_t - P_t$. Para obter as estimativas dos valores futuros da série, o método de alisamento exponencial simples utiliza a previsão do último instante ajustada pelo respectivo erro de previsão, através da expressão:

$$P_{t+1} = P_t + \alpha(Y_t - P_t),$$

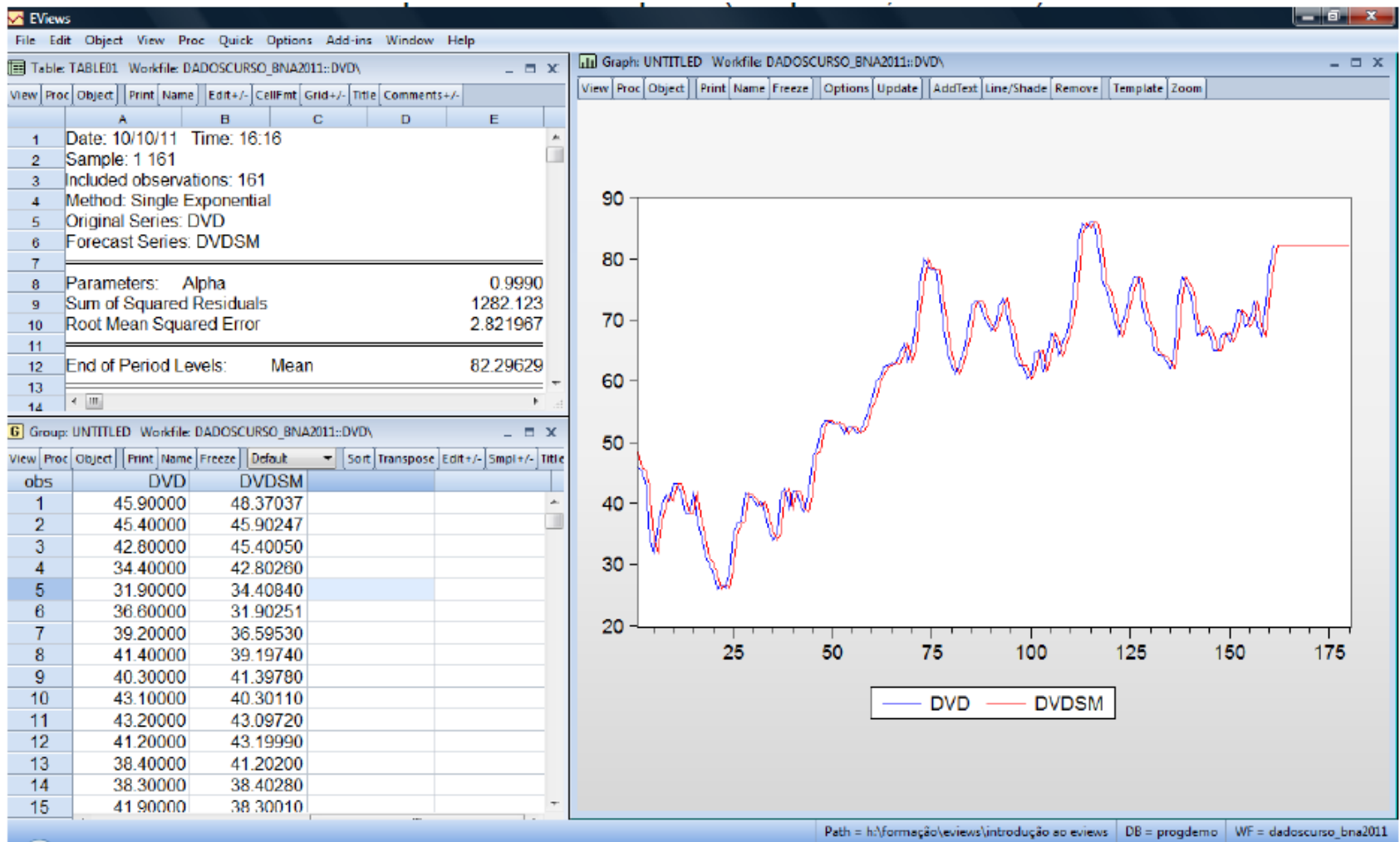
onde α é uma constante de amortecimento ou alisamento que varia entre 0 e 1.

Dado que o método de alisamento simples utiliza apenas o último valor observado e a previsão para esse instante, as previsões para horizontes maiores são constantes para todo o horizonte temporal, isto é:

$$P_{t+h} = P_t, \quad h = 1, 2, \dots$$

Métodos de Alisamento Exponencial

Alisamento exponencial simples



Métodos de Alisamento Exponencial

Alisamento exponencial duplo

O modelo de alisamento exponencial duplo é apropriado para séries temporais que apresentam tendência linear e consiste na aplicação do método de alisamento exponencial simples duas vezes, utilizando a mesma constante de alisamento. Para proceder ao alisamento duplo de uma série utilizam-se as seguintes equações de actualização:

$$M_t = \alpha Y_t + (1 - \alpha)M_{t-1}$$

$$D_t = \alpha M_t + (1 - \alpha)D_{t-1},$$

As previsões com horizonte temporal de h passos à frente são obtidas através da expressão:

$$P_{t+h} = a(t) + b(t) \times h, \quad h = 1, 2, \dots$$

onde $a(t) = 2M_t - D_t$ e $b(t) = (M_t - D_t) \left[\frac{\alpha}{(1 - \alpha)} \right]$

Métodos de Alisamento Exponencial

Alisamento exponencial duplo - Inicialização

1) **Método das médias.** Estimar o declive da série através da expressão,

$$b(1) = \frac{\bar{Y}_2 - \bar{Y}_1}{k},$$

onde \bar{Y}_1 consiste na média das primeiras k observações da série e \bar{Y}_2 a média das k observações seguintes, e utilizar este declive para obter a estimativa do nível da série usando a expressão,

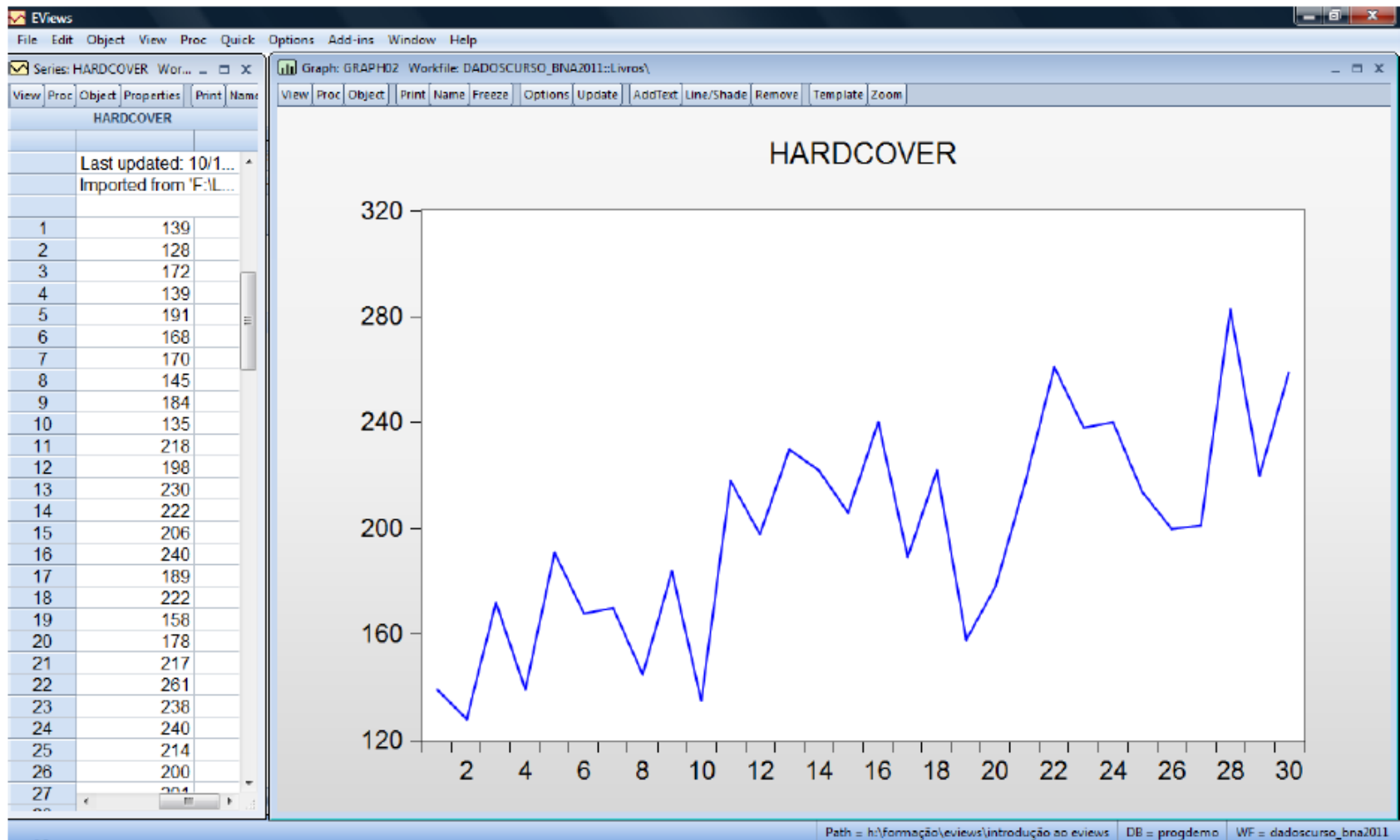
$$a(1) = \bar{Y}_1 - b(1) \times \frac{k+1}{2}$$

2) **Método da regressão.** Efectuar a regressão linear das primeiras k observações da série, utilizando como variável independente o tempo (t), o que equivale a ajustar a equação da recta

$$\hat{Y}_t = a(1) + b(1)t,$$

Métodos de Alisamento Exponencial

Alisamento exponencial duplo - Inicialização

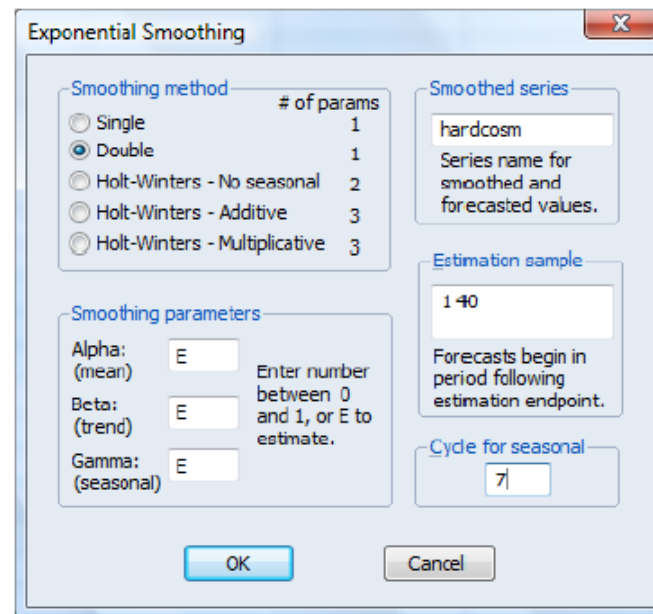


Métodos de Alisamento Exponencial

Alisamento exponencial duplo - Inicialização

A aplicação do método de alisamento exponencial duplo com o EViews permite a escolha manual e a pesquisa óptima da constante de alisamento. Ora veja:

- 1) Seleccionar a série “DVD”
- 2) Seleccionar *Proc/Exponential Smoothing/Double*



Métodos de Alisamento Exponencial

Alisamento exponencial duplo - Inicialização

Date: 10/10/11 Time: 16:37

Sample: 1 30

Included observations: 30

Method: Double Exponential

Original Series: HARDCOVER

Forecast Series: HARDCOSM

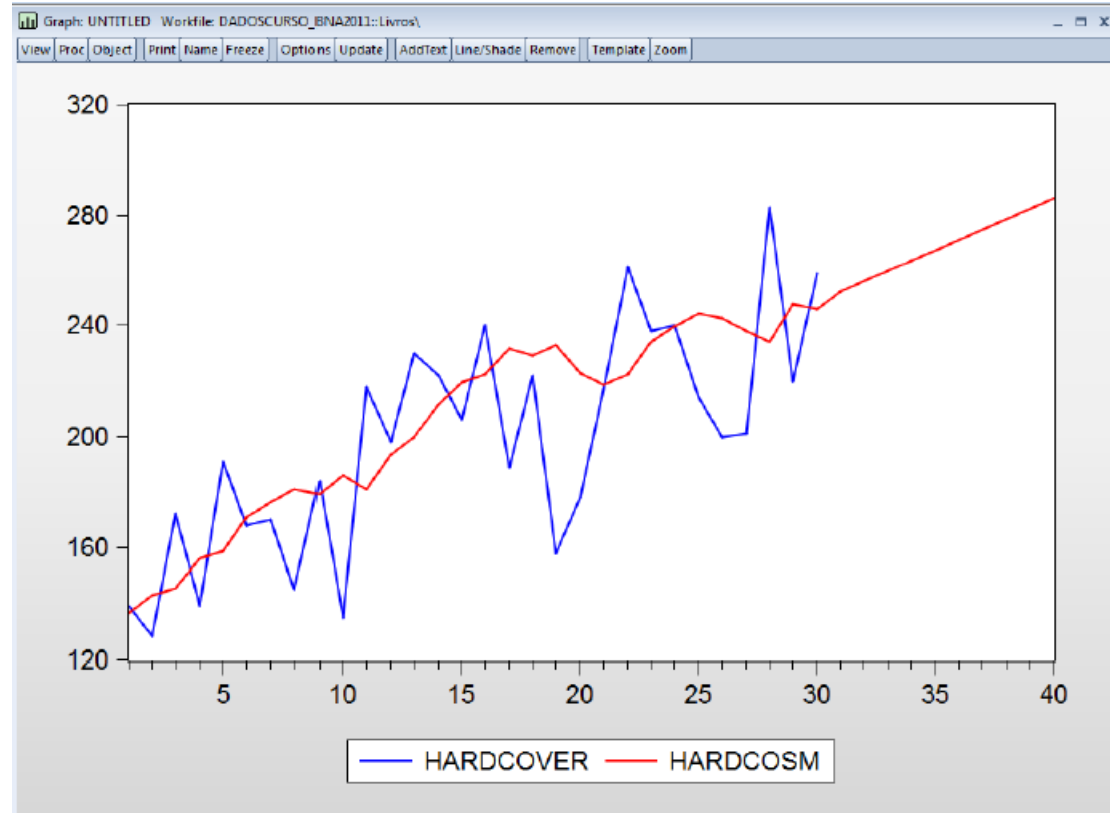
Parameters: Alpha 0.1020

Sum of Squared Residuals 27595.44

Root Mean Squared Error 30.32900

End of Period Levels: Mean 248.5509

Trend 3.776153



Métodos de Alisamento Exponencial

Método de Holt

O método de Holt é adequado a séries com tendência linear e sem movimentos de carácter sazonal e constitui uma alternativa ao método de alisamento exponencial duplo para estimar o nível e o declive de tendência da série. Este modelo utiliza as seguintes equações de actualização:

$$a(t) = \alpha Y_t + (1 - \alpha)[a(t - 1) + b(t - 1)] , 0 < \alpha < 1$$

$$b(t) = \beta[a(t) - a(t - 1)] + (1 - \beta)b(t - 1) , 0 < \beta < 1$$

onde $a(t)$ e $b(t)$ são as equações de actualização do nível e do declive da série, respectivamente, e α e β são as constantes de alisamento. As previsões a h passos à frente são obtidas através da função:

$$P_{t+h} = a(t) + b(t) \times h , h = 1, 2, \dots$$

Métodos de Alisamento Exponencial

Método de Holt

Exponential Smoothing

Smoothing method # of params

- Single 1
- Double 1
- Holt-Winters - No seasonal 2
- Holt-Winters - Additive 3
- Holt-Winters - Multiplicative 3

Smoothing parameters

Alpha: (mean) Enter number between 0 and 1, or E to estimate.

Beta: (trend)

Gamma: (seasonal)

Smoothed series

hardcover_holt

Series name for smoothed and forecasted values.

Estimation sample

140

Forecasts begin in period following estimation endpoint.

Cycle for seasonal

5

OK Cancel

Date: 10/10/11 Time: 19:31

Sample: 1 30

Included observations: 30

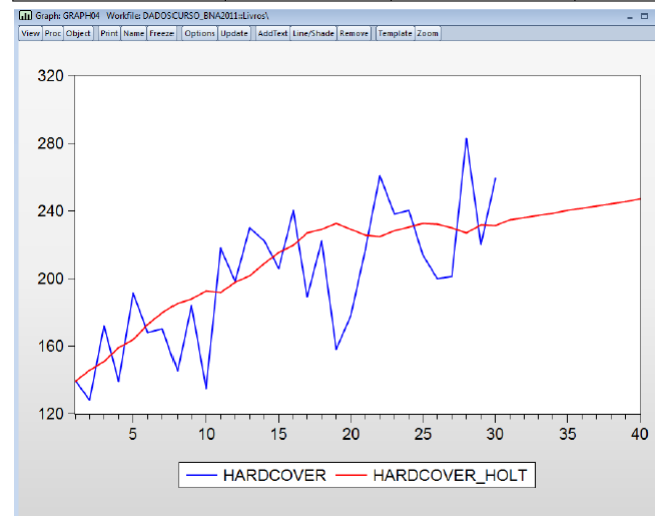
Method: Holt-Winters No Seasonal

Original Series: HARDCOVER

Forecast Series: HARDCOVER_HOLT

Parameters:	Alpha	0.0700
	Beta	0.4900
	Sum of Squared Residuals	26706.89
	Root Mean Squared Error	29.83672

End of Period Levels:	Mean	233.0897
	Trend	1.386631



Métodos de Alisamento Exponencial

Método de Holt-Winters

O método de Holt-Winters é apropriado para séries que apresentam tendência linear e movimentos sazonais. Na forma **multiplicativa**, o método de Holt-Winters é definido pelas seguintes equações de actualização:

$$a(t) = \alpha \frac{Y_t}{S_{t-s}} + (1-\alpha)[a(t-1) + b(t-1)] \quad , 0 < \alpha < 1$$

$$b(t) = \beta[a(t) - a(t-1)] + (1-\beta)b(t-1) \quad , 0 < \beta < 1$$

$$S_t = \gamma \frac{Y_t}{a(t)} + (1-\gamma)S_{t-s} \quad , 0 < \gamma < 1,$$

onde $a(t)$, $b(t)$ e S_t representam as expressões do nível, do declive e do índice sazonal, respectivamente; s é o comprimento de sazonalidade, isto é, o número de meses, trimestres ou semestres do ano (12 se mensal, 4 se trimestral, 2 se semestral); e α , β e γ são as constantes de alisamento.

As previsões a h passos à frente do método de Holt-Winters multiplicativo são obtidas através da função:

$$P_{t+h} = [a(t) + b(t) \times h] \times S_{t+h-s} \quad , h = 1, 2, \dots$$

Métodos de Alisamento Exponencial

Método de Holt-Winters - Inicialização

Tal como os outros métodos de alisamento exponencial, é necessário obter os valores iniciais de $a(t)$, $b(t)$ e também de S_t para arrancar com o algoritmo de Holt-Winters. Assim, para inicializar o nível, calcula-se a média das primeiras s observações,

$$a(s) = (1/s) \sum_{t=1}^s Y_t$$

Para inicializar o valor do declive, utiliza-se a expressão

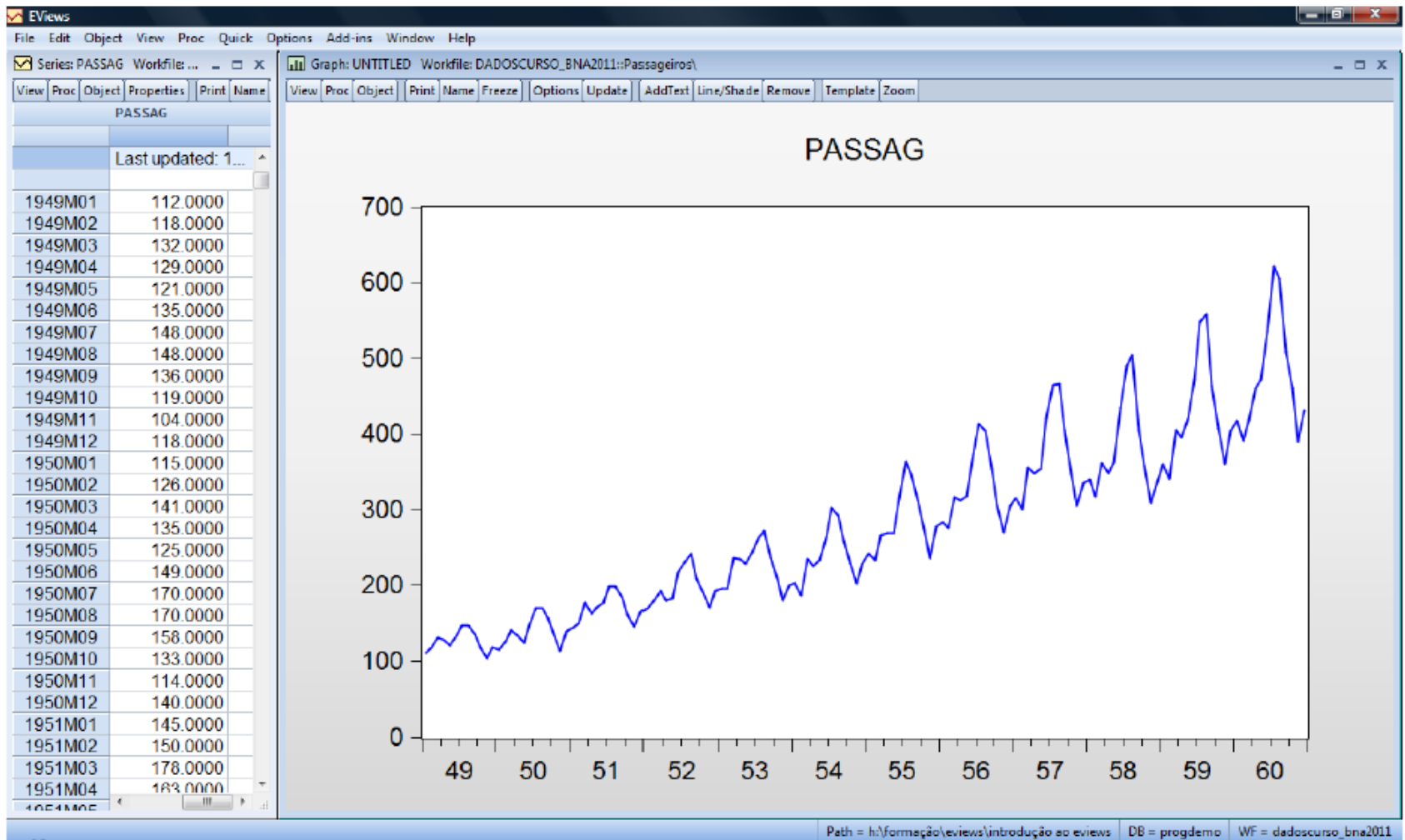
$$b(s) = (1/s^2) \left(\sum_{t=s+1}^{2s} Y_t - \sum_{t=1}^s Y_t \right)$$

E por último, os primeiros índices sazonais são obtidos por

$$S_1 = Y_1/a(s), S_2 = Y_2/a(s), \dots, S_s = Y_s/a(s)$$

Métodos de Alisamento Exponencial

Método de Holt-Winters



Métodos de Alisamento Exponencial

Método de Holt-Winters

Exponential Smoothing

Smoothing method # of params

- Single 1
- Double 1
- Holt-Winters - No seasonal 2
- Holt-Winters - Additive 3
- Holt-Winters - Multiplicative 3

Smoothed series

passag_hwm

Series name for smoothed and forecasted values.

Estimation sample

1949m01 1962m12

Forecasts begin in period following estimation endpoint.

Smoothing parameters

Alpha: (mean) E Enter number between 0 and 1, or E to estimate.

Beta: (trend) E

Gamma: (seasonal) E

Cycle for seasonal

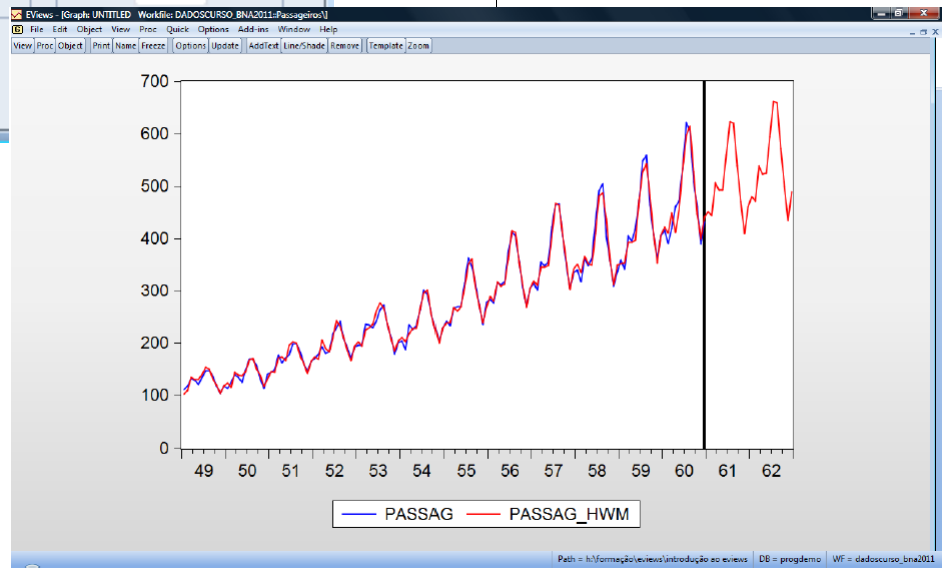
12

OK

Date: 10/10/11 Time: 19:47
 Sample: 1949M01 1960M12
 Included observations: 144
 Method: Holt-Winters Multiplicative Seasonal
 Original Series: PASSAG
 Forecast Series: PASSAG_HWM

Parameters:	Alpha	0.8500
	Beta	0.0000
	Gamma	0.0000
	Sum of Squared Residuals	16429.93
	Root Mean Squared Error	10.68161

End of Period Levels:	Mean	488.6778
	Trend	2.647727
	Seasonals:	
	1960M01	0.917865
	1960M02	0.897986
	1960M03	1.020369
	1960M04	0.986309
	1960M05	0.982004
	1960M06	1.107922
	1960M07	1.228576
	1960M08	1.215925
	1960M09	1.050468
	1960M10	0.914270
	1960M11	0.791584
	1960M12	0.886722



Métodos de Alisamento Exponencial

Método de Holt-Winters

O método aditivo de Holt-Winters baseia-se nas seguintes equações de actualização:

$$a(t) = \alpha(Y_t - S_{t-s}) + (1 - \alpha)[a(t-1) + b(t-1)] \quad , 0 < \alpha < 1$$

$$b(t) = \beta[a(t) - a(t-1)] + (1 - \beta)b(t-1) \quad , 0 < \beta < 1$$

$$S_t = \gamma[Y_t - a(t)] + (1 - \gamma)S_{t-s} \quad , 0 < \gamma < 1$$

As previsões a h passos à frente do modelo de sazonalidade aditiva são obtidas através da função:

$$P_{T+h} = a(T) + b(T) \times h + S_{T+h-s} \quad , h = 1, 2, \dots$$

A inicialização do método de Holt-Winters aditivo é igual ao método multiplicativo, exceptuando os índices sazonais, cujos valores iniciais são obtidos por

$$S_1 = Y_1 - a(s), S_2 = Y_2 - a(s), \dots, S_s = Y_s - a(s).$$

Métodos de Alisamento Exponencial

Outras formas de alisamento

Tendência	Sazonalidade		
	Nula	Aditiva	Multiplicativa
Nula	$P_{t+h} = a_t$	$P_{t+h} = a_t + S_{t+h-s}$	$P_{t+h} = a_t S_{t+h-s}$
Aditiva	$P_{t+h} = a_t + hb_t$	$P_{t+h} = a_t + hb_t + S_{t+h-s}$	$P_{t+h} = (a_t + hb_t) S_{t+h-s}$
Aditiva amortecida	$P_{t+h} = a_t + \sum_{i=1}^h \phi^i b_t$	$P_{t+h} = a_t + \sum_{i=1}^h \phi^i b_t + S_{t+h-s}$	$P_{t+h} = \left(a_t + \sum_{i=1}^h \phi^i b_t \right) S_{t+h-s}$
Multiplicativa	$P_{t+h} = a_t b_t^h$	$P_{t+h} = a_t b_t^h + S_{t+h-s}$	$P_{t+h} = a_t b_t^h S_{t+h-s}$
Multiplicativa amortecida	$P_{t+h} = a_t b_t^{\sum_{i=1}^h \phi^i}$	$P_{t+h} = a_t b_t^{\sum_{i=1}^h \phi^i} + S_{t+h-s}$	$P_{t+h} = a_t b_t^{\sum_{i=1}^h \phi^i} S_{t+h-s}$

Nota: P_t representa a previsão para o instante t , a_t o nível, b_t a tendência, S_t a sazonalidade, h o horizonte temporal e ϕ a constante de amortecimento.

Métodos de Alisamento Exponencial

Exercício de aplicação

Com base no ficheiro de EViews “Dados_ST.wk1”, na página “Taxas_POR”, efectue as seguintes análises:

- a) Altere a amostra (sample) para o período de 2003m1 a 2012m4.
- b) Construa as séries da taxa de inflação homóloga e mensal do sector de lazer (“ipc_lazer”) e represente-as graficamente.
- c) Ajuste aos dados taxa de inflação homóloga o método de alisamento exponencial duplo com $\alpha=0,1$ e α óptimo. Comente as previsões obtidas.
- d) Ajuste aos dados da taxa de inflação homóloga o método de Holt com $\alpha=0,15$ e $\beta=0,3$ e α e β óptimos. Comente as previsões obtidas.
- e) Ajuste aos dados da taxa de inflação mensal o método de Holt-Winters aditivo com $\alpha=0,1$, $\beta=0,3$ e $\gamma=0,15$ e com α , β e γ óptimos. Comente as previsões obtidas.
- f) Calcule as funções dos erros de previsão EQM, EAM e EPAM para as previsões passo a passo calculadas entre 1994m8 e 1995m7 com as duas especificações do método estudado na alínea anterior.
- g) Estabeleça previsões da série de inflação mensal para o período de 1996m8 a 1998m7 com base no método de Holt-Winters aditivo (com pesquisa óptima de parâmetros).

Modelos ARIMA

Estacionaridade, FAC e FACP

Um processo estocástico pode definir-se como uma sucessão (finita ou infinita) de variáveis aleatórias:

$$\{Y_1, Y_2, \dots, Y_n, \dots\} \text{ ou } \{Y_t, t = 1, 2, \dots, n, \dots\}, \text{ em tempo discreto;}$$

ou

$$\{Y_t : t \in \tau\}, \tau = [0, \infty[, \text{ em tempo contínuo.}$$

Uma série temporal pode entender-se como uma das infinitas realizações de um processo estacionário. Os processos estacionários baseiam-se no pressuposto de que o sistema se encontre num determinado estado de equilíbrio estatístico.

Modelos ARIMA

Estacionaridade, FAC e FACP

Um processo $Y_t, t = 1, 2, \dots, n$ diz-se estacionário até à 2ª ordem (ou estacionário em covariância) se tem média constante,

$$\mu_t = E(Y_t) = \mu,$$

variância constante,

$$\sigma_t^2 = \text{Var}(Y_t) = E(Y_t - \mu_t)^2 = \sigma^2,$$

a covariância entre Y_{t_1} e Y_{t_2} ,

$$\gamma(t_1, t_2) = E(Y_{t_1} - \mu_{t_1})(Y_{t_2} - \mu_{t_2}) = \gamma(t_1 + k, t_2 + k), \quad \forall t_1, t_2, k$$

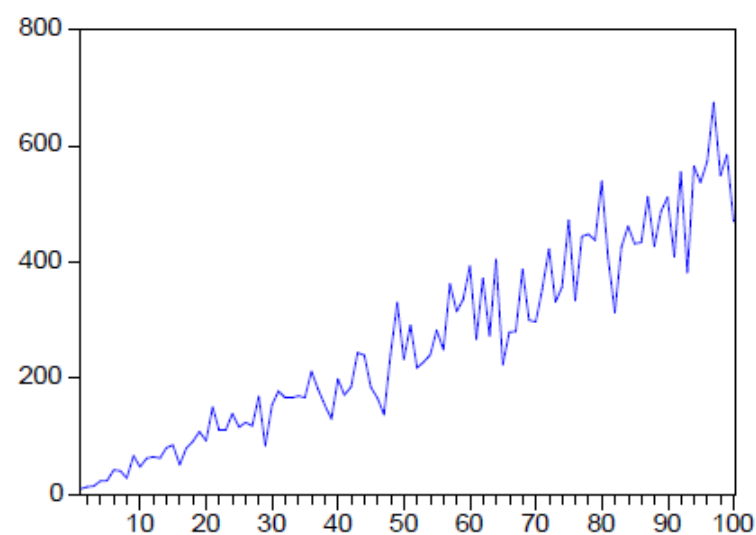
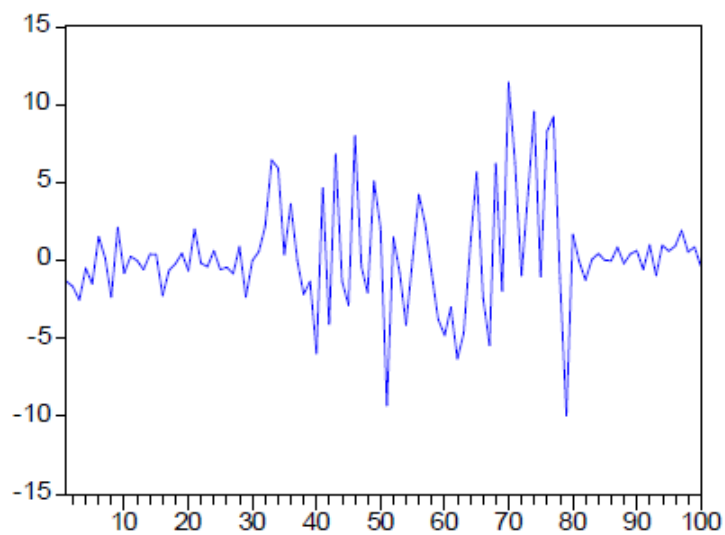
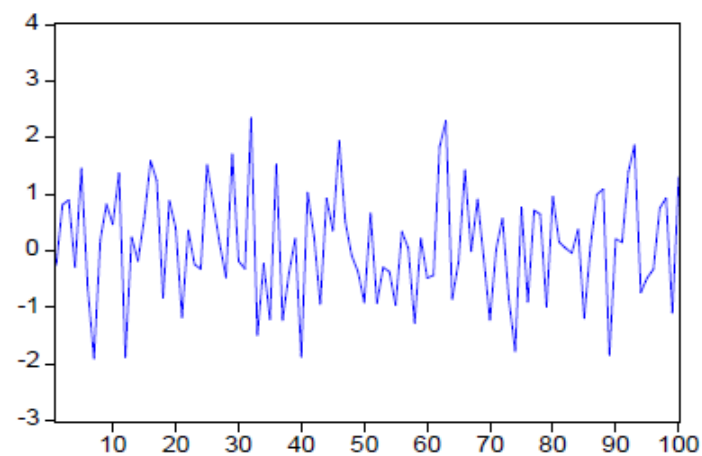
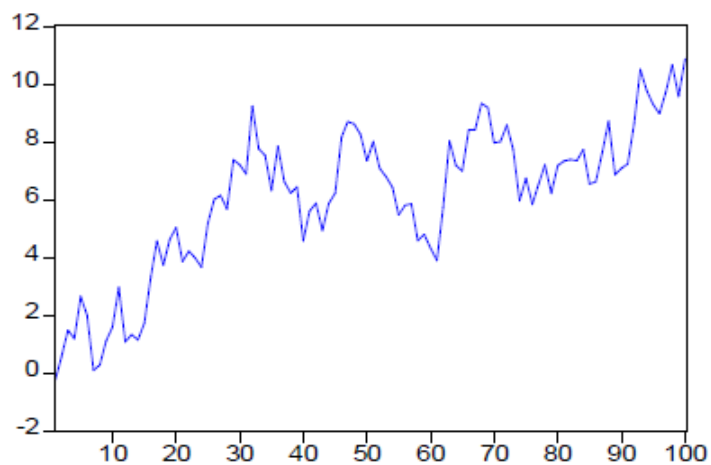
e a correlação entre Y_{t_1} e Y_{t_2} ,

$$\rho(t_1, t_2) = \frac{\gamma(t_1, t_2)}{\sqrt{\sigma_{t_1}^2} \sqrt{\sigma_{t_2}^2}} = \rho(t_1 + k, t_2 + k), \quad \forall t_1, t_2, k,$$

são independentes do tempo t , dependendo apenas da diferença de tempo k entre Y_t e Y_{t+k} .

Modelos ARIMA

Estacionaridade, FAC e FACP



Modelos ARIMA

Estacionaridade, FAC e FACP

Função autocovariância

$$\gamma_k = \text{Cov}(Y_t, Y_{t+k}) = E[(Y_t - \mu)(Y_{t+k} - \mu)]$$

Função autocorrelação (FAC)

$$\rho_k = \frac{\text{Cov}(Y_t, Y_{t+k})}{\sqrt{[\text{Var}(Y_t)][\text{Var}(Y_{t+k})]}} = \frac{\gamma_k}{\gamma_0}$$

As principais propriedades das funções autocovariância e autocorrelação são as seguintes:

- 1) $\gamma_0 = \text{Var}(Y_t)$; $\rho_0 = 1$;
- 2) $|\gamma_k| \leq \gamma_0$; $|\rho_k| \leq 1$;
- 3) $\gamma_k = \gamma_{-k}$; $\rho_k = \rho_{-k}$ (γ_k e ρ_k são simétricas em torno da origem, $k = 0$);
- 4) γ_k e ρ_k são semidefinidas positivas.

Modelos ARIMA

Estacionaridade, FAC e FACP

Função de autocorrelação parcial (FACP)

Suponha-se um processo estacionário Y_t , com $E(Y_t) = 0$ e considere-se a regressão linear múltipla de Y_{t+k} sobre as k variáveis desfasadas $Y_{t+k-1}, Y_{t+k-2}, \dots, Y_t$:

$$Y_{t+k} = \phi_{k1} Y_{t+k-1} + \phi_{k2} Y_{t+k-2} + \dots + \phi_{kk} Y_t + e_{t+k},$$

onde ϕ_{kj} , $j = 1, 2, \dots, k$ são os coeficientes de regressão e e_{t+k} o erro não correlacionado com Y_{t+k-j} para $j \geq 1$. A FACP está relacionada com a FAC através do sistema de equações de Yule-Walker que se passa a descrever.

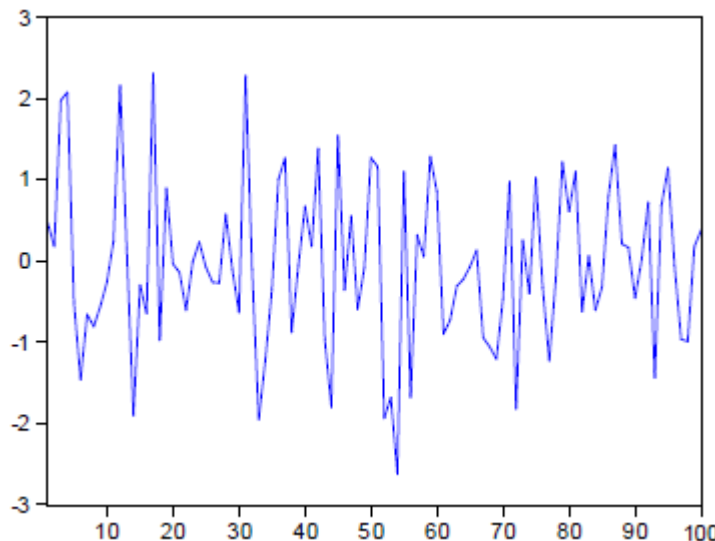
Modelos ARIMA

Ruído branco

Um processo ruído branco (“*white noise*”) é constituído por uma sucessão de valores não correlacionados ao longo do tempo e expressa-se por:

$$Y_t = \varepsilon_t$$

onde o termo aleatório ε_t tem média constante, $E(\varepsilon_t) = \mu_\varepsilon$, variância constante, $Var(\varepsilon_t) = \sigma_\varepsilon^2$ e covariância nula, $Cov(\varepsilon_t, \varepsilon_{t-k}) = 0$, para todo o $k \neq 0$. As FAC e FACP de um ruído branco são igualmente nulas para todo o $k \neq 0$.



Modelos ARIMA

Estimação das FAC e FACP

No caso da função de autocorrelação (FAC), esta pode ser estimada através da expressão:

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} = \frac{\sum_{t=k+1}^n (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2}, \quad k = 0, 1, 2, \dots$$

No caso da função de autocorrelação parcial (FACP), a estimação pode ser feita através de um método recursivo, através da expressão:

$$\hat{\phi}_{kk} = \frac{\hat{\rho}_k - \sum_{j=1}^{k-1} \hat{\phi}_{k-1,j} \hat{\rho}_{k-j}}{1 - \sum_{j=1}^{k-1} \hat{\phi}_{k-1,j} \hat{\rho}_j},$$

com $\hat{\phi}_{11} = \hat{\rho}_1$ (inicialização) e $\hat{\phi}_{kj} = \hat{\phi}_{k-1,j} - \hat{\phi}_{kk} \hat{\phi}_{k-1,k-j}$, $j = 1, 2, \dots, k-1$.

Modelos ARIMA

Estimação das FAC e FACP - Exemplo

Taxa de juro

t	X_t	X_{t-1}	X_{t-2}	X_{t-3}	D	D1	D2	D3	P1	P2	P3
1	2,36				-0,65						
2	2,58	2,36			-0,43	-0,65			0,28		
3	2,68	2,58	2,36		-0,33	-0,43	-0,65		0,15	0,22	
4	2,73	2,68	2,58	2,36	-0,28	-0,33	-0,43	-0,65	0,09	0,12	0,19
5	2,84	2,73	2,68	2,58	-0,17	-0,28	-0,33	-0,43	0,05	0,06	0,08
6	2,98	2,84	2,73	2,68	-0,03	-0,17	-0,28	-0,33	0,01	0,01	0,01
7	2,94	2,98	2,84	2,73	-0,07	-0,03	-0,17	-0,28	0,00	0,01	0,02
8	3,05	2,94	2,98	2,84	0,04	-0,07	-0,03	-0,17	0,00	0,00	-0,01
9	3,21	3,05	2,94	2,98	0,20	0,04	-0,07	-0,03	0,01	-0,01	-0,01
10	3,44	3,21	3,05	2,94	0,43	0,20	0,04	-0,07	0,08	0,02	-0,03
11	3,60	3,44	3,21	3,05	0,59	0,43	0,20	0,04	0,25	0,11	0,02
12	3,76	3,60	3,44	3,21	0,75	0,59	0,43	0,20	0,44	0,32	0,15
Méd	3,01										

	FAC	FACP
1	0,690	0,690
2	0,435	-0,078
3	0,212	-0,110

Correlogram of TAXA

Date: 22/09/11 Time: 22:13
 Sample: 2010M07 2011M06
 Included observations: 12

	Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1	0,690	0,690	7.2654	0,007		
2	0,435	-0,078	10.439	0,005		
3	0,212	-0,110	11.277	0,010		
4	0,040	-0,081	11.310	0,023		
5	-0,083	-0,067	11.474	0,043		

Modelos ARIMA

Representação médias móveis e autoregressiva

Processo médias móveis infinito

O processo médias móveis infinito, ou processo $MA(\infty)$, descreve o processo Y_t como combinação linear de uma série de variáveis aleatórias não correlacionadas:

$$Y_t = \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \dots = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j},$$

onde $\psi_0 = 1$, ε_t é um processo ruído branco com média zero, $E(\varepsilon_t) = 0$, e variância constante, $Var(\varepsilon_t) = \sigma_\varepsilon^2$, e $\sum_{j=0}^{\infty} \psi_j^2 < \infty$. Se definirmos o operador atraso pela relação, $B^j Y_t = Y_{t-j}$, podemos escrever o processo $MA(\infty)$ na seguinte forma compactada:

$$Y_t = (1 + \psi_1 B + \psi_2 B^2 + \dots) \varepsilon_t = \psi(B) \varepsilon_t,$$

onde $\psi(B) = \sum_{j=0}^{\infty} \psi_j B^j$.

Modelos ARIMA

Representação médias móveis e autoregressiva

Pode facilmente mostrar-se que:

$$E(Y_t) = 0,$$

$$\text{Var}(Y_t) = \sigma_\varepsilon^2 \sum_{j=0}^{\infty} \psi_j^2,$$

e

$$E(\varepsilon_t Y_{t-k}) = \begin{cases} \sigma_\varepsilon^2, & k = 0 \\ 0, & k > 0, \end{cases}$$

o que permite obter a FAC do processo:

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{E(Y_t Y_{t+k})}{\text{Var}(Y_t)} = \frac{\sum_{j=0}^{\infty} \psi_j \psi_{j+k}}{\sum_{j=0}^{\infty} \psi_j^2}.$$

Modelos ARIMA

Representação médias móveis e autoregressiva

Processo autoregressivo infinito

Outra forma de escrever o processo Y_t é na representação autoregressiva infinita, ou $AR(\infty)$, que consiste na regressão de Y_t sobre os seus valores passados mais um choque aleatório:

$$Y_t = \pi_1 Y_{t-1} + \pi_2 Y_{t-2} + \dots + \varepsilon_t = \sum_{j=1}^{\infty} \pi_j Y_{t-j} + \varepsilon_t,$$

ou,

$$(1 - \pi_1 B - \pi_2 B^2 - \dots) Y_t = \varepsilon_t,$$

ou ainda,

$$\pi(B) Y_t = \varepsilon_t,$$

onde $\pi(B) = 1 - \pi_1 B - \pi_2 B^2 - \dots = 1 - \sum_{j=1}^{\infty} \pi_j B^j$ e $1 + \sum_{j=1}^{\infty} |\pi_j| < \infty$.

Modelos ARIMA

Processos estacionários

Box e Jenkins (1970, 1976) introduziram os modelos lineares de séries estacionárias não sazonais, designados por modelos ARMA(p,q), através da representação genérica:

$$Y_t = \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}, \phi_p \neq 0 \text{ e } \theta_q \neq 0,$$

onde a série Y_t é função dos seus valores passados e da combinação linear de uma sucessão de choques aleatórios; $\phi_1, \phi_2, \dots, \phi_p$ são os parâmetros autoregressivos e p a ordem da componente autoregressiva; $\theta_1, \theta_2, \dots, \theta_q$ são os parâmetros de médias móveis e q a ordem da componente de médias móveis. Na modelação empírica de séries estacionárias, são frequentemente utilizados modelos parcimoniosos do tipo ARMA(p,q), $p, q = 0,1,2$. Em seguida, descreve-se a sua estrutura e as suas principais propriedades estatísticas.

Modelos ARIMA

Processos estacionários – AR(1)

Modelo AR(1)

O modelo autoregressivo de 1ª ordem ou AR(1) tem a representação,

$$Y_t = \phi Y_{t-1} + \varepsilon_t,$$

onde ε_t é um ruído branco de média zero. Para que o processo seja estacionário deve satisfazer a condição $|\phi| < 1$. Para melhor ilustrar esta condição de estacionaridade, escreva-se o processo na forma compactada (usando o já conhecido operador atraso B):

$$(1 - \phi B)Y_{t-1} = \varepsilon_t,$$

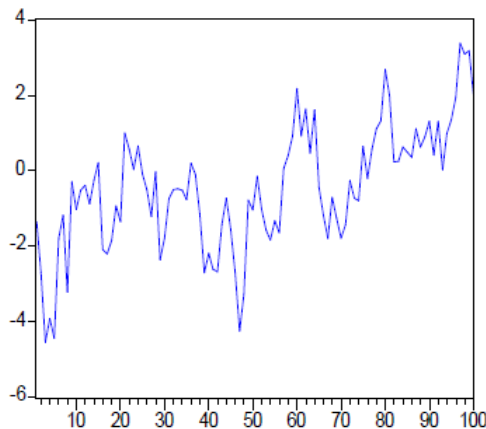
onde $\phi(B) = (1 - \phi B)$ é uma função polinomial de Y_t . A estacionaridade do processo AR(1) exige que a raiz da equação polinomial $\phi(B) = 0$, dada por $B = \phi^{-1}$, tenha módulo superior a um, o que equivale a pedir que:

$$|B| = \left| \frac{1}{\phi} \right| > 1 \Leftrightarrow |\phi| < 1.$$

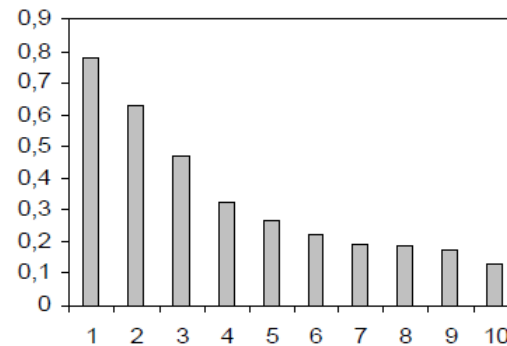
Modelos ARIMA

Processos estacionários – Simulação de AR(1)

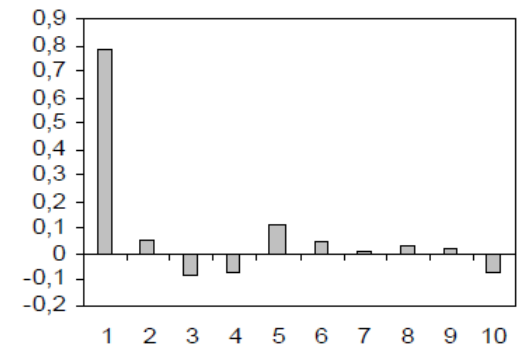
(a) $Y_t = 0,7Y_{t-1} + \varepsilon_t$



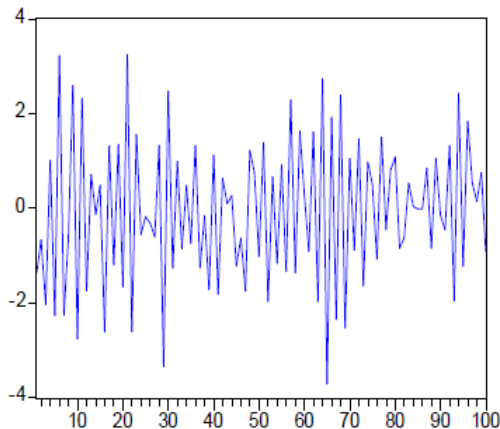
FAC do AR(1): $Y_t = 0,7Y_{t-1} + \varepsilon_t$



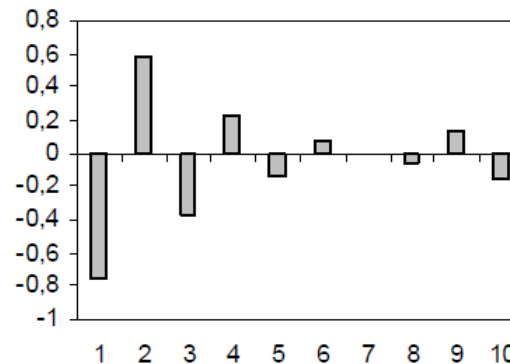
FACP do AR(1): $Y_t = 0,7Y_{t-1} + \varepsilon_t$



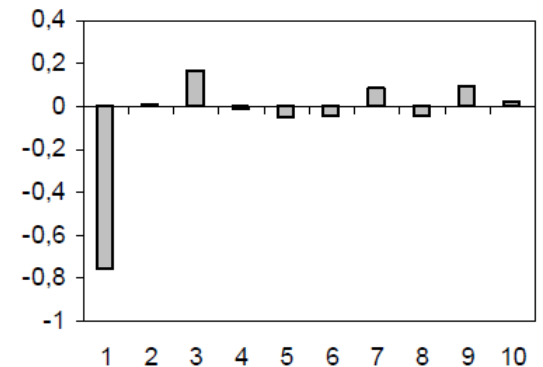
(b) $Y_t = -0,7Y_{t-1} + \varepsilon_t$



FAC do AR(1): $Y_t = -0,7Y_{t-1} + \varepsilon_t$



FACP do AR(1): $Y_t = -0,7Y_{t-1} + \varepsilon_t$



Modelos ARIMA

Processos estacionários – AR(2)

Modelo AR(2)

O modelo autoregressivo de 2ª ordem ou AR(2) tem a forma,

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t,$$

ou, utilizando o operador atraso,

$$(1 - \phi_1 B - \phi_2 B^2) Y_t = \varepsilon_t,$$

ou,

$$\phi_2(B) Y_t = \varepsilon_t,$$

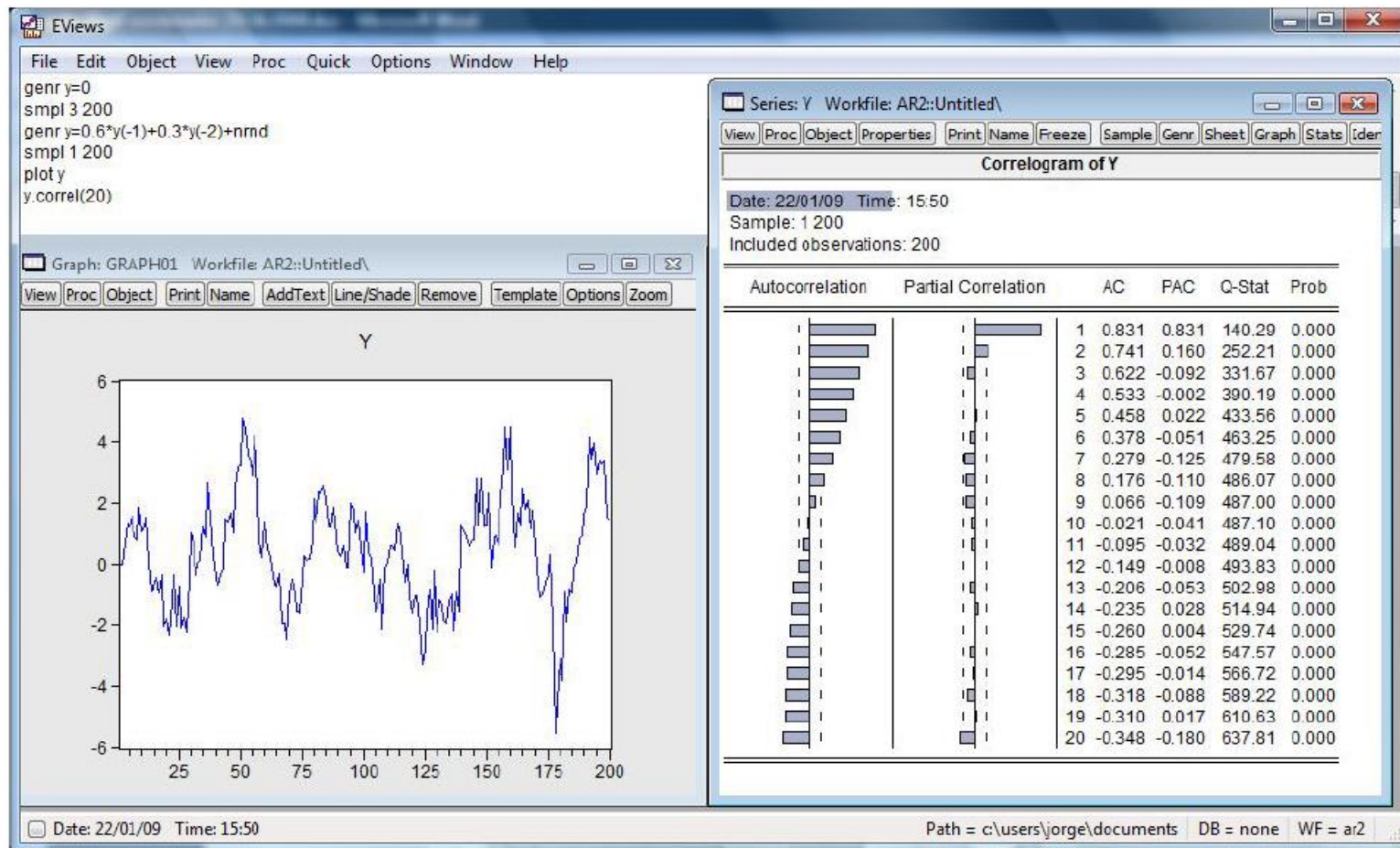
onde $\phi_2(B) = 1 - \phi_1 B - \phi_2 B^2$ é um polinómio autoregressivo de 2ª ordem e ε_t é um ruído branco de média zero. Para ser estacionário, é necessário que as raízes de $\phi_2(B) = 0$ tenham módulo superior a um.

Deste modo, as condições necessárias e suficientes de estacionaridade do processo AR(2) são:

$$\phi_2 + \phi_1 < 1 \wedge \phi_2 - \phi_1 < 1 \wedge -1 < \phi_2 < 1.$$

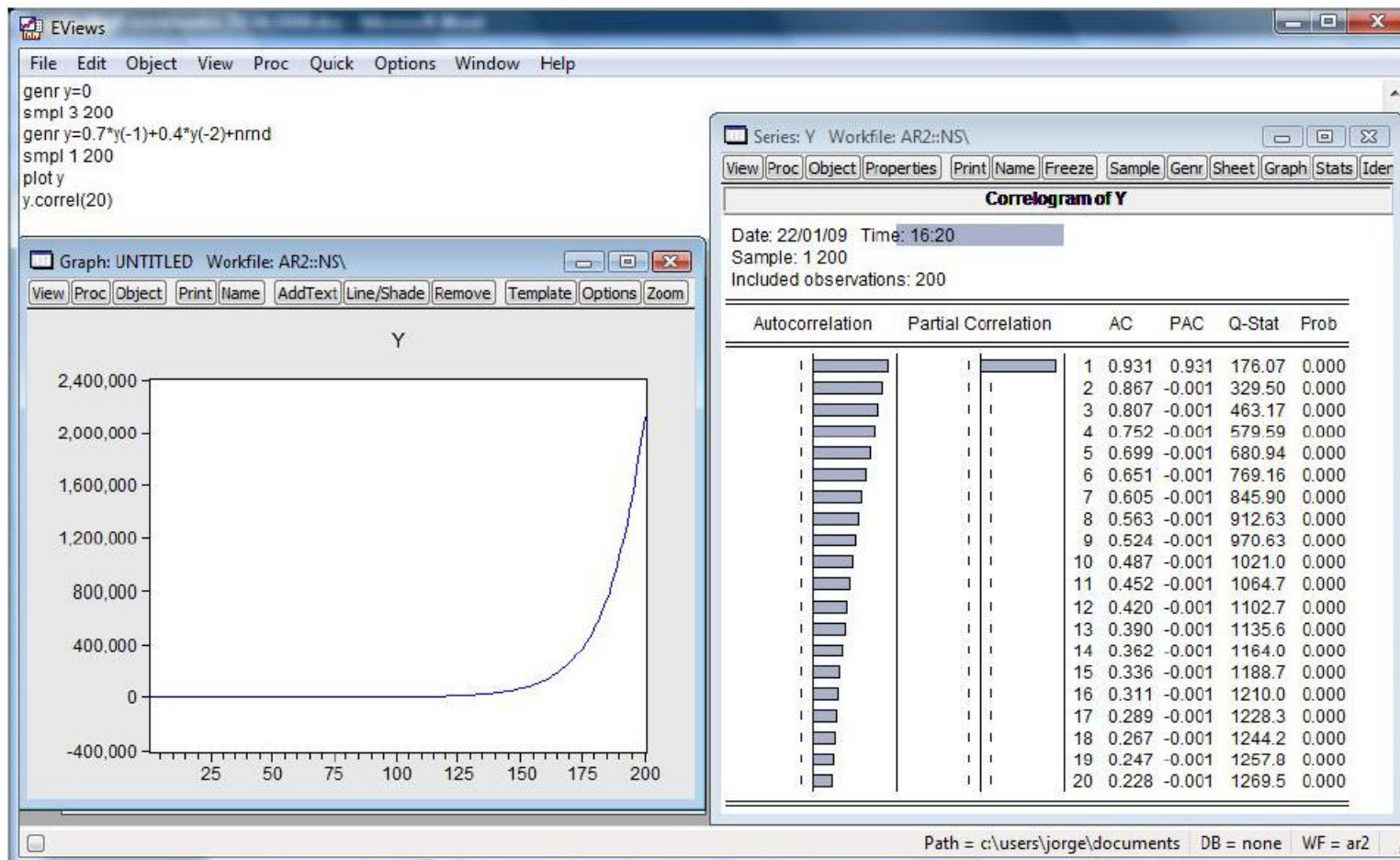
Modelos ARIMA

Processos estacionários – Simulação de AR(2) estacionário



Modelos ARIMA

Processos estacionários – Simulação de AR(2) não estac.



Modelos ARIMA

Processos estacionários – AR(p)

Modelo AR(p)

O modelo autoregressivo de ordem p ou AR(p) tem a forma,

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t, \text{ com } \phi_p \neq 0,$$

ou, utilizando o operador atraso,

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) Y_t = \varepsilon_t,$$

ou ainda,

$$\phi_p(B) Y_t = \varepsilon_t,$$

onde $\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$ é um polinómio autoregressivo de ordem p e ε_t é um ruído branco de média zero. A estacionaridade do processo exige que as p raízes de $\phi_p(B) = 0$ tenham módulo superior a um.

Modelos ARIMA

Processos estacionários – MA(1)

Modelo MA(1)

O modelo médias móveis de 1ª ordem ou MA(1) descreve a série Y_t através de um choque aleatório, ε_t , e também do seu valor desfasado um período, ε_{t-1} . Define-se através da relação,

$$Y_t = \varepsilon_t - \theta\varepsilon_{t-1},$$

ou,

$$Y_t = (1 - \theta B)\varepsilon_t,$$

ou ainda,

$$Y_t = \theta(B)\varepsilon_t,$$

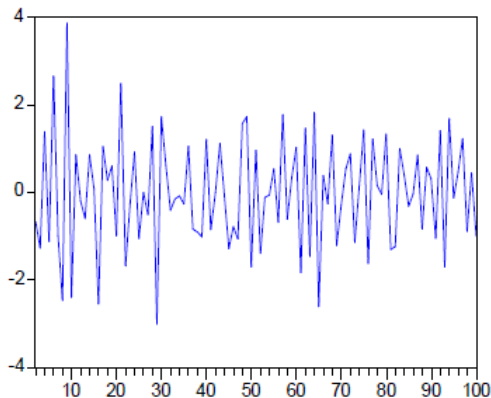
onde $\theta(B) = 1 - \theta B$ é um polinómio médias móveis de 1ª ordem e ε_t é um ruído branco de média zero. O processo MA(1) é sempre estacionário, porquanto $1 + \theta^2 < \infty$ (veja-se Wei, 2007). Para ser invertível, a raiz de $\theta(B) = 0$ deve ter módulo superior a um:

$$|B| = \left| \frac{1}{\theta} \right| > 1 \Leftrightarrow |\theta| < 1.$$

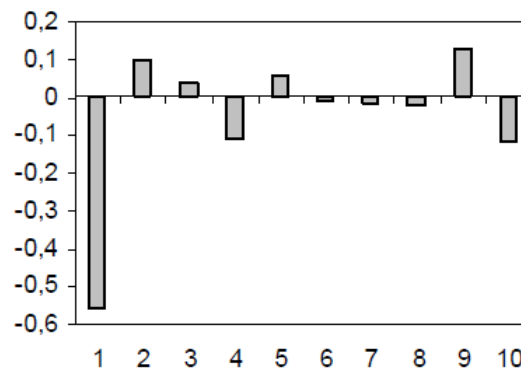
Modelos ARIMA

Processos estacionários – Simulação de MA(1)

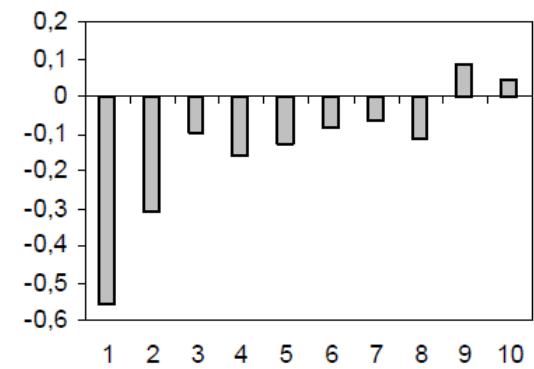
(a) $Y_t = \varepsilon_t - 0.75\varepsilon_{t-1}$



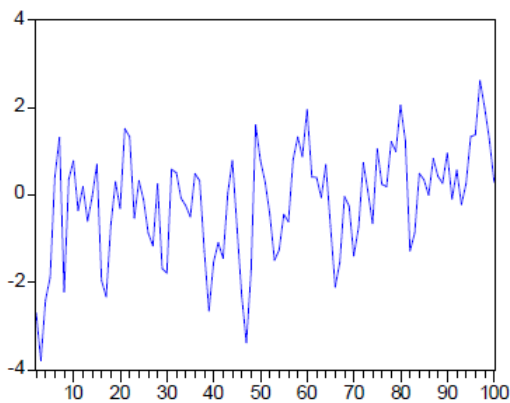
FAC do MA(1): $Y_t = \varepsilon_t - 0.75\varepsilon_{t-1}$



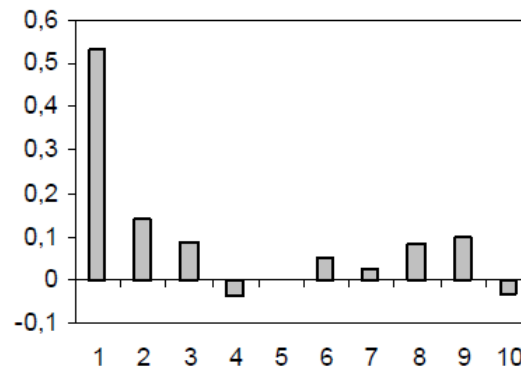
FACP do MA(1): $Y_t = \varepsilon_t - 0.75\varepsilon_{t-1}$



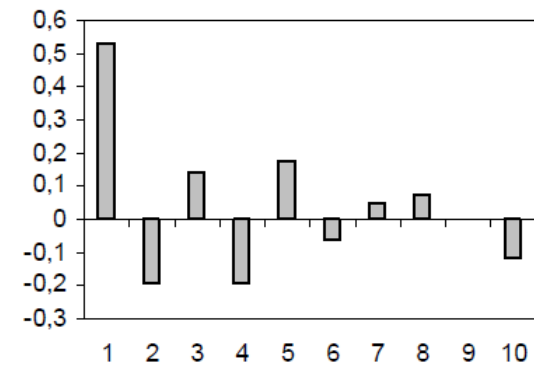
(b) $Y_t = \varepsilon_t + 0.75\varepsilon_{t-1}$



FAC do MA(1): $Y_t = \varepsilon_t + 0.75\varepsilon_{t-1}$



FACP do MA(1): $Y_t = \varepsilon_t + 0.75\varepsilon_{t-1}$



Modelos ARIMA

Processos estacionários – MA(2)

Modelo MA(2)

O modelo médias móveis de 2ª ordem ou MA(2) assume a expressão,

$$Y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2},$$

ou,

$$Y_t = (1 - \theta_1 B - \theta_2 B^2) \varepsilon_t,$$

ou ainda,

$$Y_t = \theta_2(B) \varepsilon_t,$$

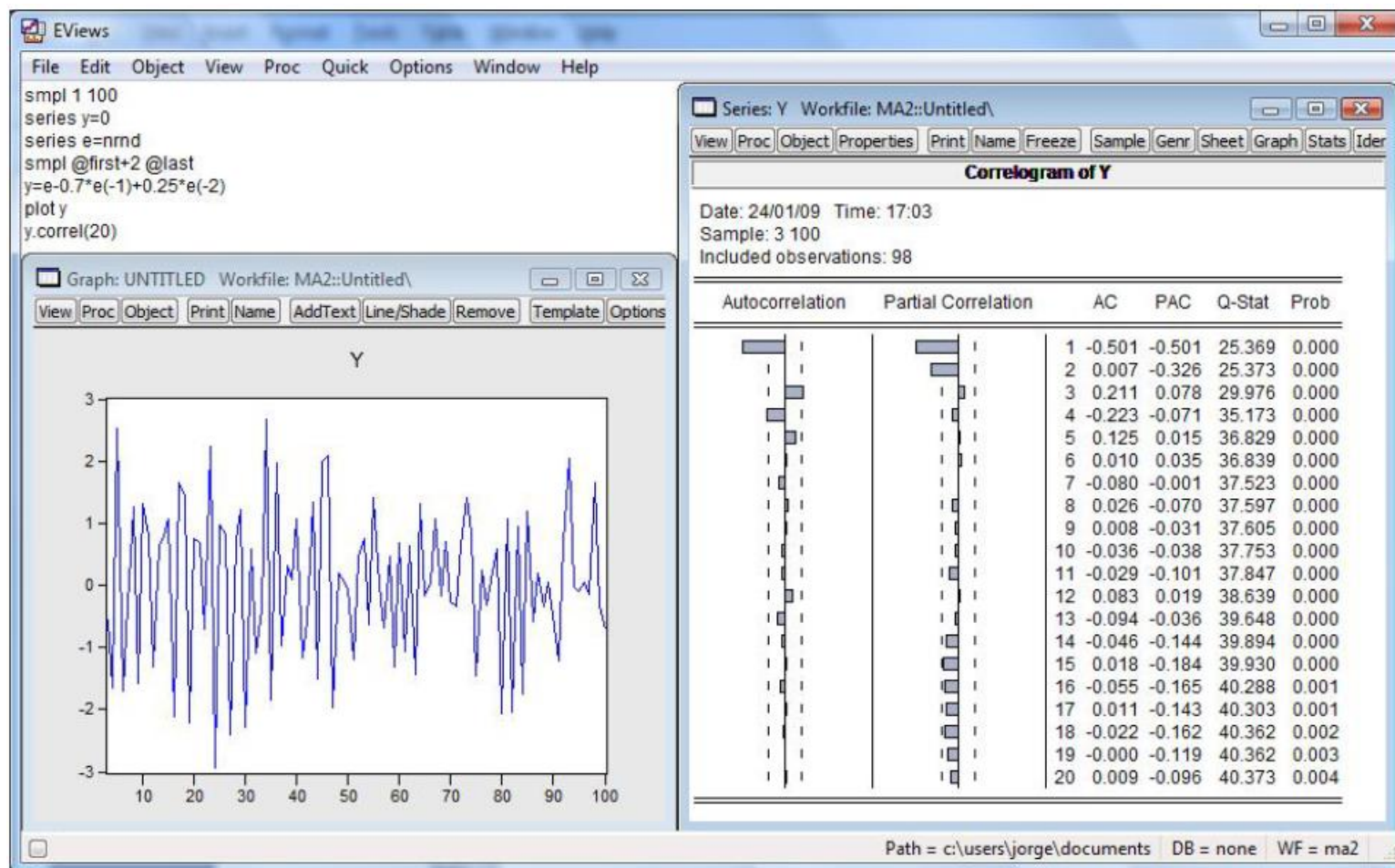
onde $\theta_2(B) = 1 - \theta_1 B - \theta_2 B^2$ é um polinómio médias móveis de 2ª ordem e ε_t é um ruído branco de média zero. Para ser invertível, as raízes de $\theta_2(B) = 0$ devem ter módulo superior a um, o que equivale a pedir que os coeficientes de médias móveis θ_1 e θ_2 têm que satisfazer as condições:

$$\theta_2 + \theta_1 < 1 \wedge \theta_2 - \theta_1 < 1 \wedge -1 < \theta_2 < 1.$$

Modelos ARIMA

Processos estacionários – Simulação de MA(2) estacionário

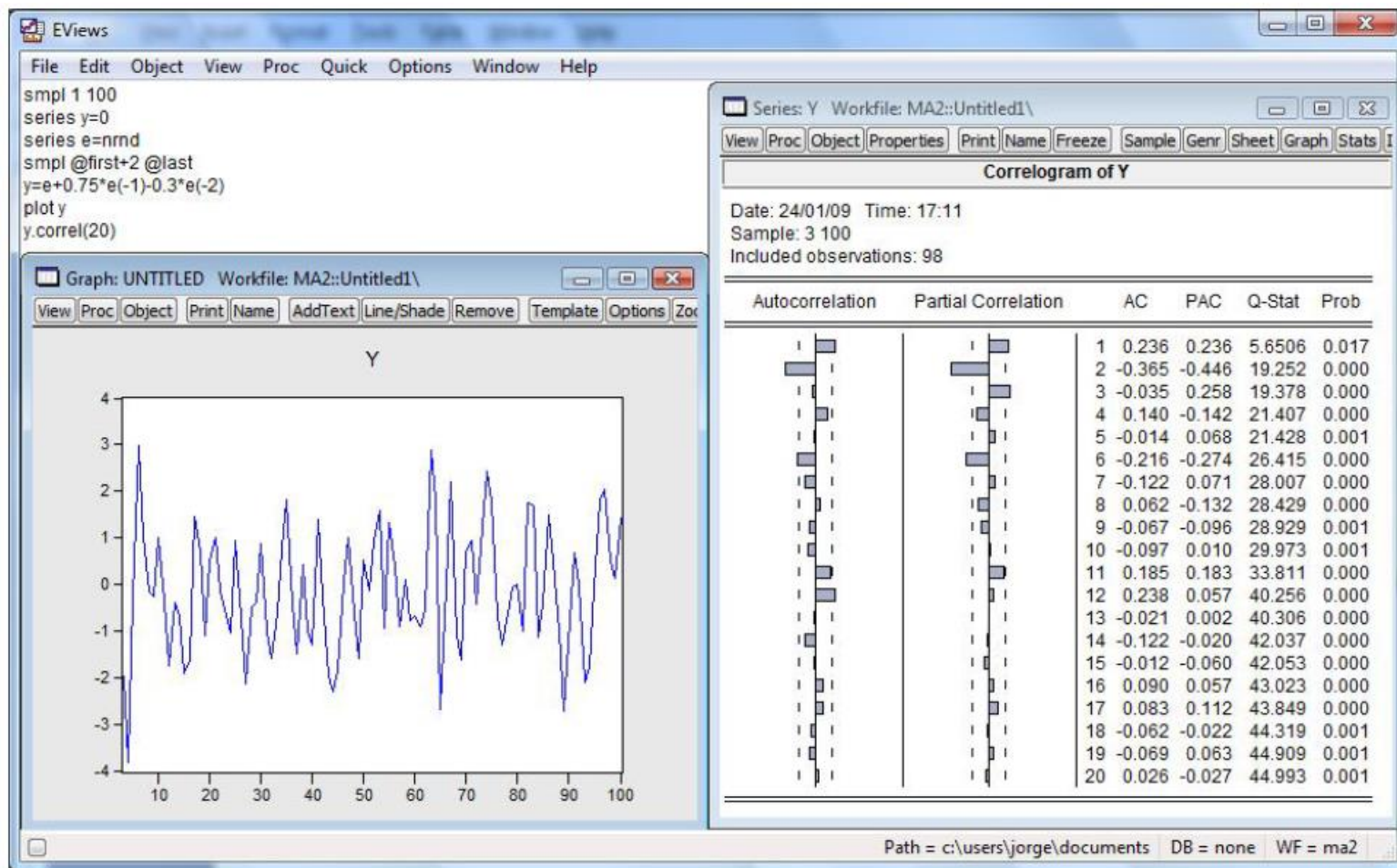
$$Y_t = \varepsilon_t - 0,7\varepsilon_{t-1} + 0,25\varepsilon_{t-2}$$



Modelos ARIMA

Processos estacionários – Simulação de MA(2) estacionário

$$Y_t = \varepsilon_t + 0,75\varepsilon_{t-1} - 0.3\varepsilon_{t-2}$$



Modelos ARIMA

Processos estacionários – MA(q)

Modelo MA(q)

O modelo médias móveis de ordem q ou MA(q) assume a expressão:

$$Y_t = \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \dots - \theta_q\varepsilon_{t-q}, \text{ com } \theta_q \neq 0,$$

ou,

$$Y_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)\varepsilon_t,$$

ou ainda,

$$Y_t = \theta_q(B)\varepsilon_t,$$

onde $\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$ é um polinómio de médias móveis de ordem q e ε_t é um ruído branco com média zero e variância constante. Para ser invertível, as raízes de $\theta_q(B) = 0$ devem ter módulo superior a um.

Modelos ARIMA

Processos estacionários – ARMA(1,1)

Modelo ARMA(1,1)

O modelo ARMA(1,1) inclui um factor autoregressivo de 1ª ordem e um factor de médias móveis de 1ª ordem e tem a representação,

$$Y_t = \phi Y_{t-1} + \varepsilon_t - \theta \varepsilon_{t-1}, \text{ com } \phi \neq \theta,$$

ou,

$$(1 - \phi B)Y_t = (1 - \theta B)\varepsilon_t,$$

ou ainda,

$$\phi(B)Y_t = \theta(B)\varepsilon_t,$$

onde $\phi(B) = 1 - \phi B$ é um polinómio autoregressivo de 1ª ordem, $\theta(B) = 1 - \theta B$ é um polinómio médias móveis de 1ª ordem, e ε_t é um ruído branco. Para ser estacionário, a raiz de $\phi(B) = 0$ deve ter módulo superior a um, o que acontece para $-1 < \phi < 1$. Para ser invertível, a raiz de $\theta(B) = 0$ deve ter módulo superior a um, o que equivale a pedir que $-1 < \theta < 1$.

Modelos ARIMA

Processos estacionários – ARMA(1,1)

O modelo ARMA(1,1) pode igualmente representar-se em termos de um processo de médias móveis infinito:

$$Y_t = \psi(B)\varepsilon_t,$$

com

$$\psi(B) = (1 + \psi_1 B + \psi_2 B^2 + \dots) = \frac{1 - \theta B}{1 - \phi B}.$$

Daqui resulta que:

$$(1 - \phi B)(1 + \psi_1 B + \psi_2 B^2 + \dots) = 1 - \theta B.$$

Recorrendo ao método dos coeficientes indeterminados (Wei, 2007), sai que:

$$\psi_j = \phi^{j-1}(\phi - \theta), j \geq 1.$$

Modelos ARIMA

Processos estacionários – ARMA(1,1)

O modelo ARMA(1,1) pode igualmente escrever-se em termos de uma representação autoregressiva:

$$\pi(B)Y_t = \varepsilon_t,$$

com

$$\pi(B) = 1 - \pi_1 B - \pi_2 B^2 - \dots = \frac{1 - \phi B}{1 - \theta B}.$$

Daqui resulta que:

$$(1 - \theta B)(1 - \pi_1 B - \pi_2 B^2 - \dots) = 1 - \phi B,$$

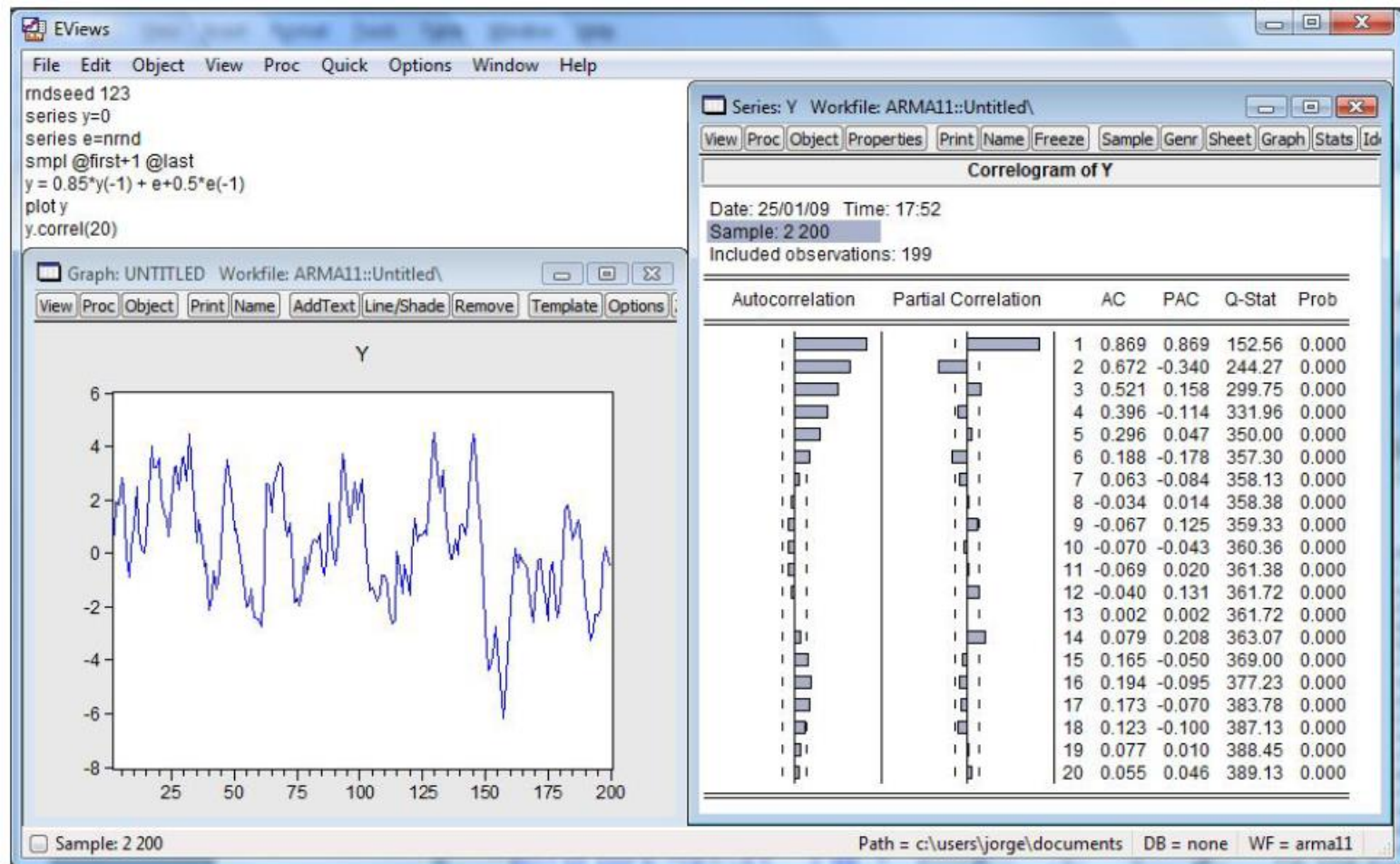
o que equivale a:

$$\pi_j = \theta^{j-1}(\phi - \theta), j \geq 1.$$

Modelos ARIMA

Processos estacionários – ARMA(1,1)

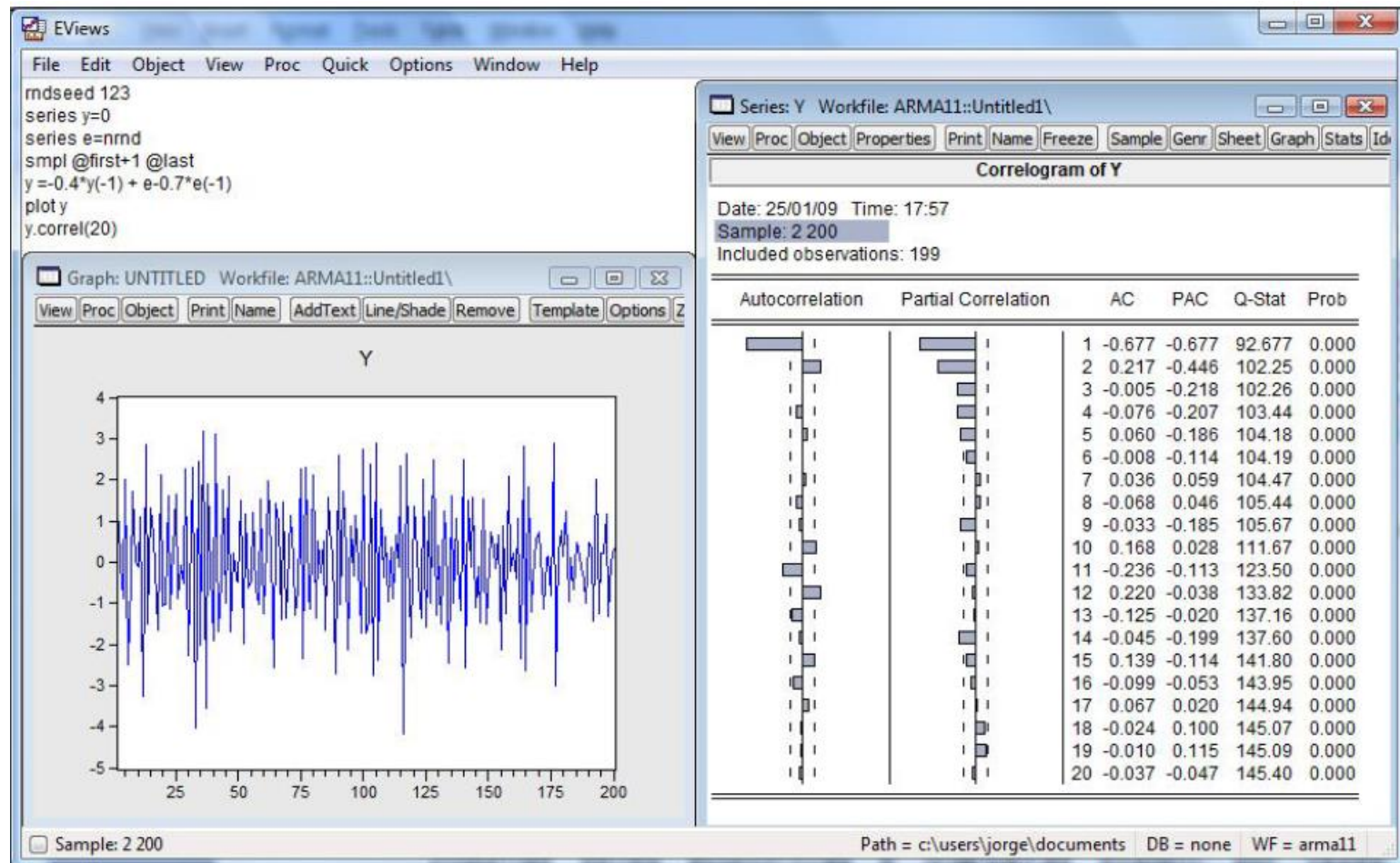
$$Y_t = 0,85Y_{t-1} + \varepsilon_t + 0,5\varepsilon_{t-1}$$



Modelos ARIMA

Processos estacionários – ARMA(1,1)

$$Y_t = -0,4Y_{t-1} + \varepsilon_t - 0,7\varepsilon_{t-1}$$



Modelos ARIMA

Processos estacionários – ARMA(p,q)

Modelo ARMA(p,q)

O modelo misto autoregressivo e médias móveis ARMA(p,q) tem a representação,

$$Y_t = \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q},$$

ou, usando o operador atraso,

$$(1 - \phi_1 B - \dots - \phi_p B^p) Y_t = (1 - \theta_1 B - \dots - \theta_q B^q) \varepsilon_t,$$

ou ainda,

$$\phi_p(B) Y_t = \theta_q(B) \varepsilon_t,$$

onde $\phi_p(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ é um polinómio autoregressivo de ordem p e

$\theta_q(B) = 1 - \theta_1 B - \dots - \theta_q B^q$ é um polinómio médias móveis de ordem q , e ε_t é um

ruído branco. A estacionaridade do processo exige que as raízes de $\phi_p(B) = 0$

tenham módulo superior a um e a invertibilidade do processo requer que as raízes

de $\theta_q(B) = 0$ tenham módulo superior a um.

Modelos ARIMA

Processos estacionários – SARMA(P,Q)_s

O modelo misto autoregressivo e de médias móveis estritamente sazonal de ordens P e Q , ou simplesmente modelo SARMA(P,Q)_s tem a forma,

$$Y_t = \Phi_1 Y_{t-s} + \dots + \Phi_P Y_{t-Ps} + \varepsilon_t - \Theta_1 \varepsilon_{t-s} - \dots - \Theta_Q \varepsilon_{t-Qs}, \text{ com } \Phi_P \neq 0 \text{ e } \Theta_Q \neq 0,$$

ou

$$(1 - \Phi_1 B^s - \dots - \Phi_P B^{Ps}) Y_t = (1 - \Theta_1 B^s - \dots - \Theta_Q B^{Qs}) \varepsilon_t,$$

ou ainda,

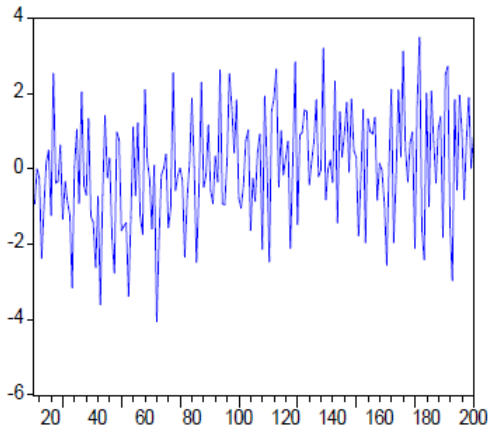
$$\Phi_P(B^s) Y_t = \Theta_Q(B^s) \varepsilon_t,$$

onde $\Phi_P(B^s) = 1 - \Phi_1 B^s - \dots - \Phi_P B^{Ps}$ é um polinómio autoregressivo estritamente sazonal em B^s de grau P , $\Theta_Q(B^s) = 1 - \Theta_1 B^s - \dots - \Theta_Q B^{Qs}$ é um polinómio médias móveis estritamente sazonal em B^s de grau Q , e ε_t é um ruído branco de média zero. Para que o processo seja estacionário e invertível, as raízes de $\Phi_P(B^s) = 0$ e $\Theta_Q(B^s) = 0$ devem ter módulo superior a um, respectivamente.

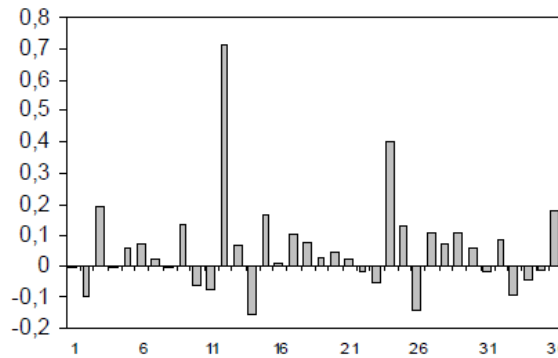
Modelos ARIMA

Processos estacionários – Simulação de SARMA(P,Q)s

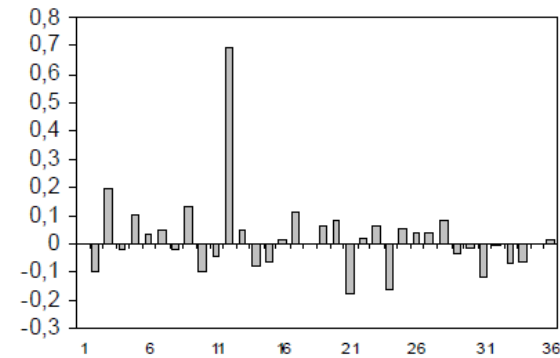
(i) $(1 - 0.65B^{12})Y_t = (1 + 0.25B^{12})\varepsilon_t$



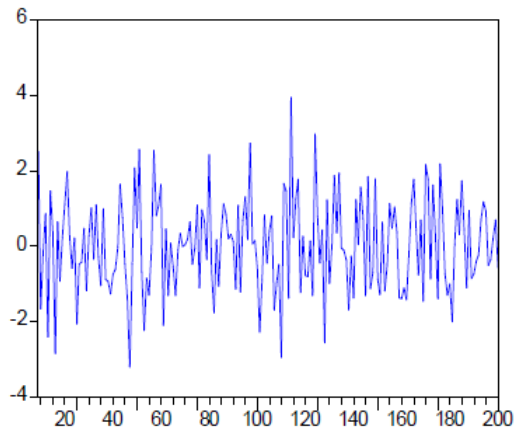
FAC do SARMA(1,1)₁₂



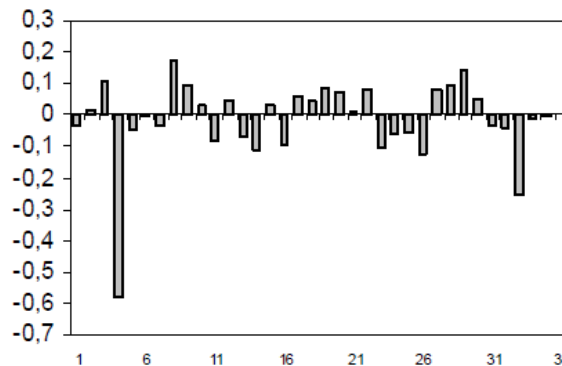
FACP do SARMA(1,1)₁₂



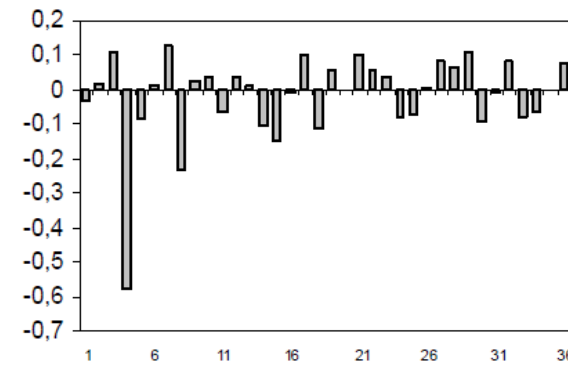
(ii) $(1 - 0.3B^4)Y_t = (1 - 0.4B^4 + 0.15B^8)\varepsilon_t$



FAC do SARMA(1,2)₄



FACP do SARMA(1,2)₄



Modelos ARIMA

Processos estacionários – SARMA(p,q)P,Q)s

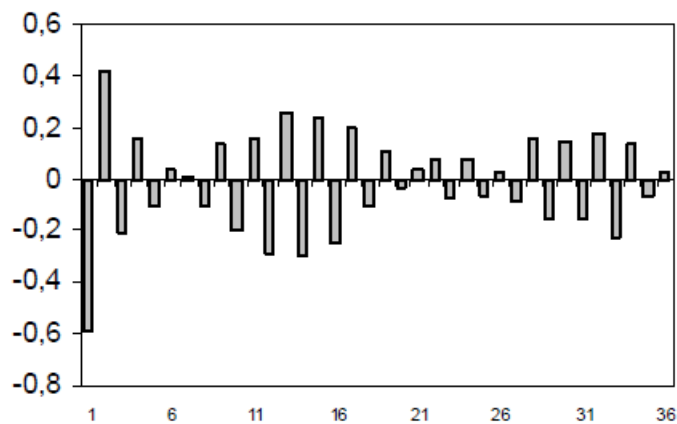
A combinação do modelo ARMA(p,q) e SARMA(P,Q)_s permite obter o modelo multiplicativo, com componente sazonal e não sazonal, SARMA(p,q)(P,Q)_s, através da expressão,

$$(1 - \phi_1 B - \dots - \phi_p B^p)(1 - \Phi_1 B^s - \dots - \Phi_P B^{Ps}) Y_t = (1 - \theta_1 B - \dots - \theta_q B^q)(1 - \Theta_1 B^s - \dots - \Theta_Q B^{Qs}) \varepsilon_t,$$

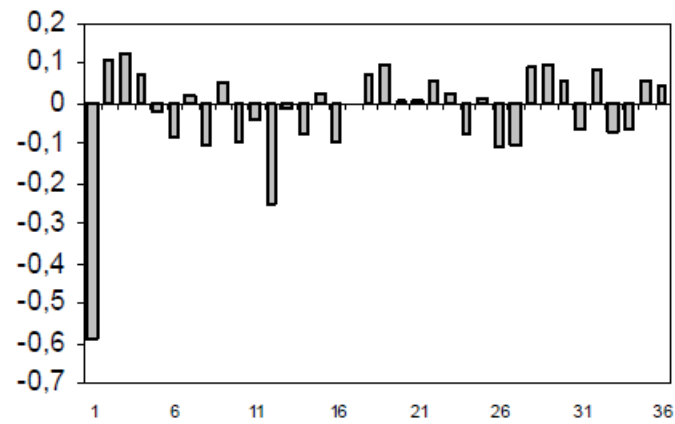
ou,

$$\phi_p(B)\Phi_P(B^s)Y_t = \theta_q(B)\Theta_Q(B^s)\varepsilon_t.$$

FAC



FACP



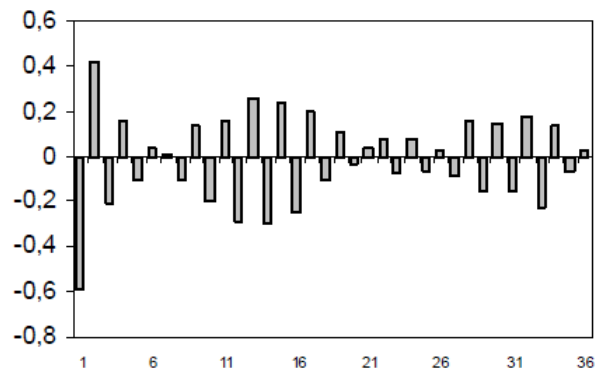
$$\text{SARMA}(1,0)(1,0)_{12}: (1 - 0.7B)(1 + 0.25B^{12})Y_t = \varepsilon_t$$

Processos estacionários

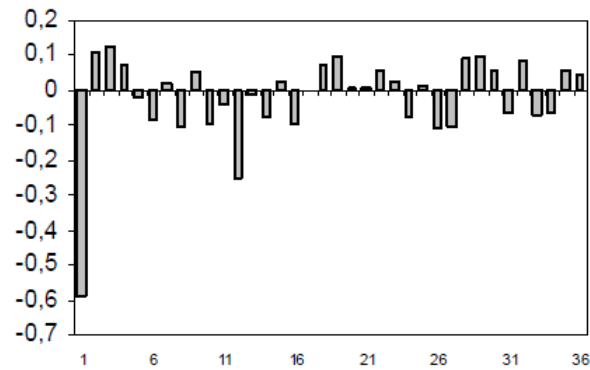
Modelo SARMA(p,q)(P,Q)s

FAC e FACP do modelo SARMA(1,0)(1,0)₁₂: $(1 - 0.7B)(1 + 0.25B^{12})Y_t = \varepsilon_t$

FAC



FACP



Processos estacionários

Figurinos teóricos da FAC e FACP

Modelo	FAC	FACP
$AR(p)$	Decaimento exponencial ou sinusoidal amortecido para zero	Queda brusca para zero a partir do lag $p+1$
$MA(q)$	Queda brusca para zero a partir do lag $q+1$	Decaimento exponencial ou sinusoidal amortecido para zero
$ARMA(p,q)$	Decaimento exponencial ou sinusoidal amortecido para zero	Decaimento exponencial ou sinusoidal amortecido para zero
$SAR(P)$	Decaimento exponencial ou sinusoidal amortecido para zero sobre os lags $s, 2s, \dots$	Queda brusca para zero a partir do lag $(P+1)s$
$SMA(Q)$	Queda brusca para zero a partir do lag $(Q+1)s$	Decaimento exponencial ou sinusoidal amortecido para zero sobre os lags $s, 2s, \dots$
$SARMA(P,Q)$	Decaimento exponencial ou sinusoidal amortecido para zero sobre os lags $s, 2s, \dots$	Decaimento exponencial ou sinusoidal amortecido para zero sobre os lags $s, 2s, \dots$
$SARMA(p,q)(P,Q)_s$	Decaimento exponencial ou sinusoidal amortecido para zero	Decaimento exponencial ou sinusoidal amortecido para zero

Processos estacionários

Exercício de aplicação

Com o auxílio do software Eviews, construa séries simuladas e estime as FAC e FACP dos seguintes processos:

- a) ARMA(1,2) com $\phi_1=0.4$, $\theta_1=-0.6$ e $\theta_2=-0.2$.
- b) SAR(3)₄ com $\Phi_1=0.55$ e $\Phi_3=-0.25$
- c) SARIMA(1,1)(0,1)₁₂ com $\phi_1=-0.45$, $\theta_1=0.2$ e $\Theta_1=0.35$

Processos não estacionários

Não estacionaridade em média

Algumas séries apresentam movimentos de tendência determinística que se manifestam consistentemente durante um período longo de tempo. Por exemplo, se a média do processo é uma função de tendência linear do tipo $\mu_t = a + bt$ (onde a e b são parâmetros e t a variável de tendência), poderá utilizar-se o seguinte modelo de tendência linear determinística,

$$Y_t = a + bt + \varepsilon_t$$

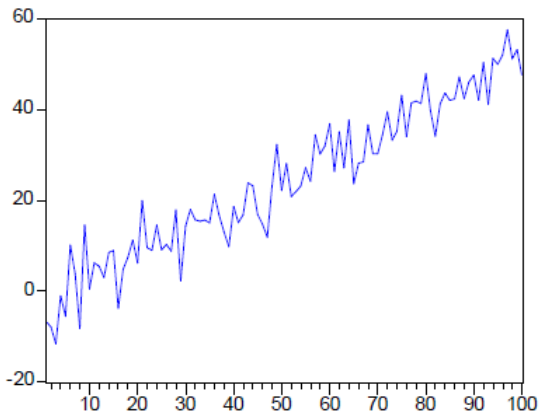
onde ε_t é um ruído branco com média zero. Suponhamos agora que a média do processo tem tendência quadrática determinística do tipo $\mu_t = a + bt + ct^2$, o que nos leva a propor o seguinte modelo de tendência quadrática,

$$Y_t = a + bt + ct^2 + \varepsilon_t.$$

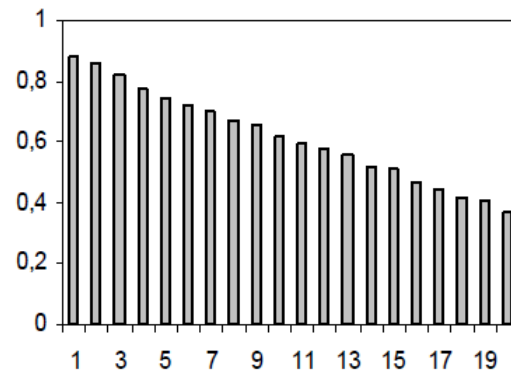
Processos não estacionários

Não estacionariedade em média

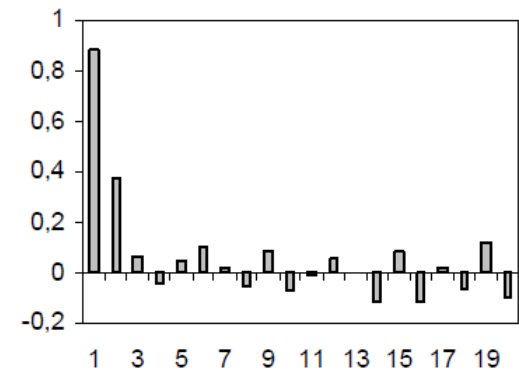
Modelo de tendência linear



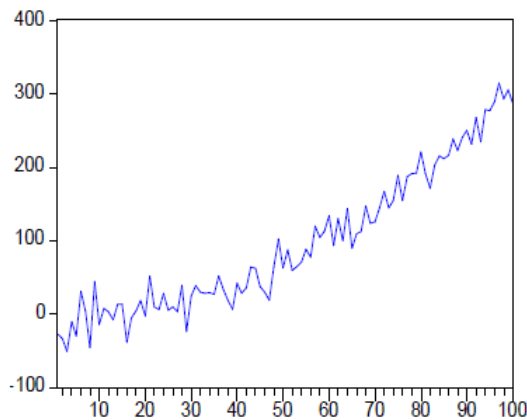
FAC do modelo de tendência linear



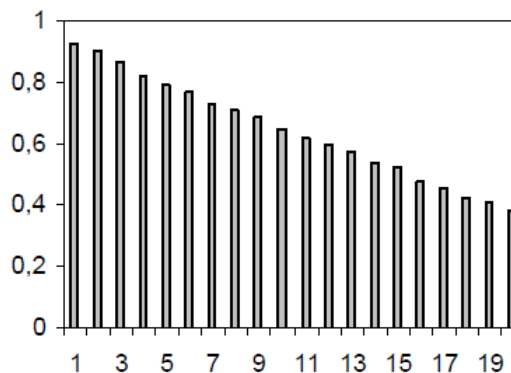
FACP do modelo de tendência linear



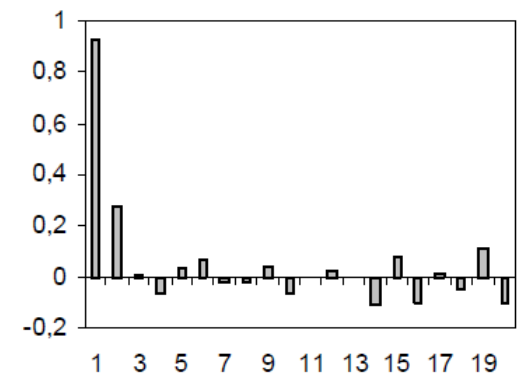
Modelo de tendência quadrática



FAC do modelo de tendência quadrática



FACP do modelo de tendência quadrática



Processos não estacionários

Diferenciação

A diferenciação simples de uma série temporal consiste em obter a diferença entre as observações da série nos momentos t e $t-1$, isto é:

$$\nabla Y_t = Y_t - Y_{t-1}$$

ou, utilizando o já conhecido operador atraso ($BY_t = Y_{t-1}$),

$$\nabla Y_t = Y_t - Y_{t-1} = (1-B)Y_t.$$

O operador de diferenciação de ordem d , para qualquer inteiro $d \geq 1$, consiste em diferenciar a série d vezes e define-se por:

$$\nabla^d Y_t = (1-B)^d Y_t.$$

Por exemplo, as diferenças de 2ª ordem de uma série Y_t correspondem às diferenças das primeiras diferenças, como se mostra de seguida:

$$\begin{aligned}\nabla^2 Y_t &= (1-B)^2 Y_t \\ &= (1-2B+B^2)Y_t \\ &= Y_t - 2Y_{t-1} + Y_{t-2} \\ &= (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}).\end{aligned}$$

Processos não estacionários

Diferenciação

Quando uma série apresenta um comportamento periódico repetitivo, pode-se aplicar uma transformação de diferenciação sazonal, definida por:

$$\nabla_s Y_t = Y_t - Y_{t-s} = (1 - B^s)Y_t.$$

Por exemplo, uma diferenciação sazonal seguida de uma diferenciação simples não sazonal pode ser expressa por:

$$\begin{aligned}\nabla(\nabla_s Y_t) &= (1 - B)(1 - B^s)Y_t \\ &= (1 - B - B^s + B^{s+1})Y_t \\ &= Y_t - Y_{t-1} - Y_{t-s} + Y_{t-s-1}.\end{aligned}$$

Por último, o operador de diferenciação sazonal de ordem D , para qualquer inteiro $D \geq 1$, consiste em diferenciar sazonalmente a série D vezes e define-se por:

$$\nabla_s^D Y_t = (1 - B^s)^D Y_t.$$

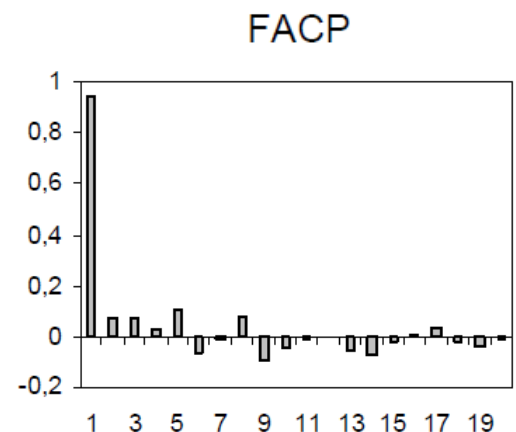
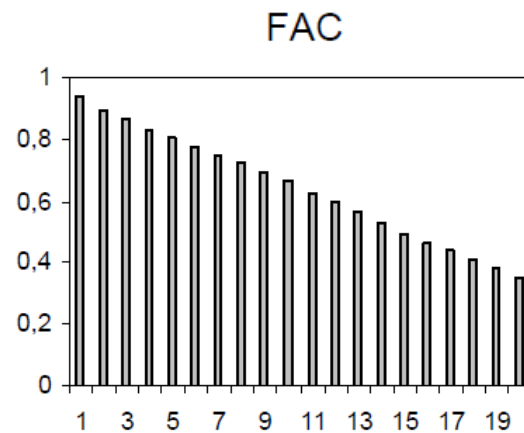
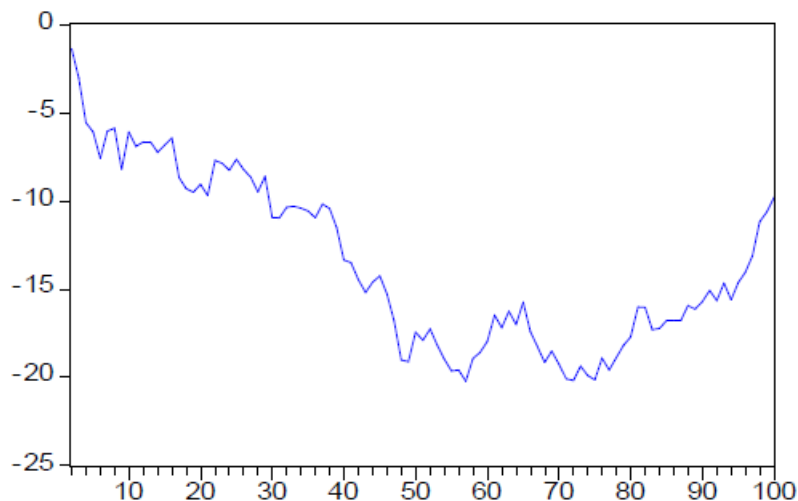
Processos não estacionários

Passeio aleatório

Para analisar a implicação do operador de diferenciação na estacionarização de uma série, considere-se o seguinte modelo de tendência estocástica,

$$Y_t = Y_{t-1} + \varepsilon_t,$$

onde ε_t é um ruído branco. Este modelo é vulgarmente conhecido por modelo de passeio aleatório (“*random walk*”), pois descreve a série no momento t com base no seu valor passado (no momento $t - 1$) mais um choque aleatório.

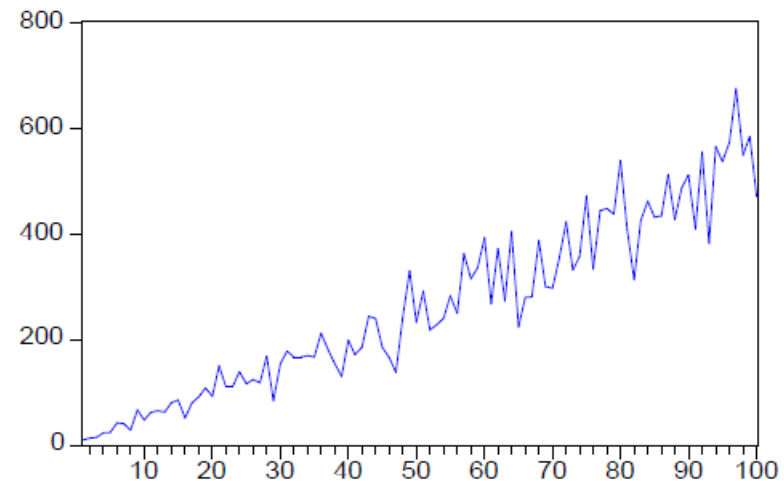
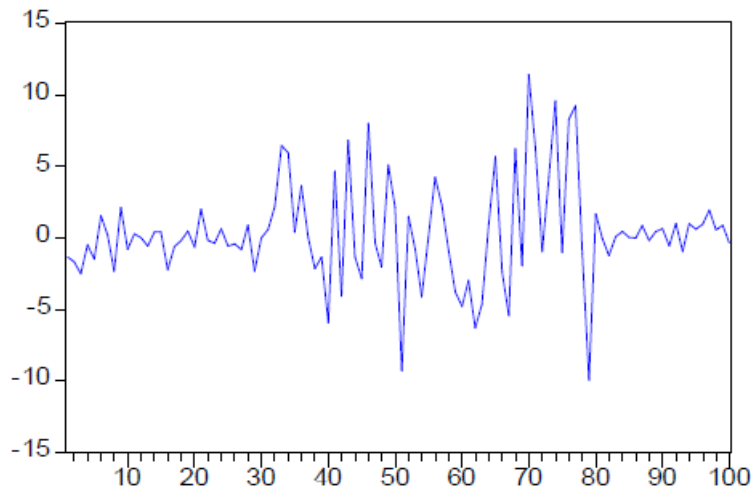


Processos não estacionários

Não estacionaridade em variância

Um processo estacionário em média não é necessariamente estacionário em variância e covariância. Contudo, um processo que não é estacionário em média também não é estacionário em variância e covariância.

Série não estacionária em variância mas estacionária em média



Série não estacionária em variância nem em média

Processos não estacionários

Estabilização da variância

Para estabilizar a variância de uma série não estacionária em variância, podemos utilizar um método de transformação paramétrica baseado na seguinte expressão:

$$X_t = T(Y_t) = \begin{cases} Y_t^\lambda, & \lambda \neq 0 \\ \log(Y_t), & \lambda = 0 \end{cases},$$

onde os valores de λ são escolhidos no intervalo $[-1,1]$, sendo os mais correntes $-1, -0.5, 0, 0.5$ e 1 , a que correspondem as transformações $X_t = 1/Y_t$, $X_t = 1/\sqrt{Y_t}$, $X_t = \log Y_t$, $X_t = \sqrt{Y_t}$ e $X_t = Y_t$, respectivamente.

Processos não estacionários

Modelo ARIMA(p,d,q)

O modelo ARIMA(p,d,q) assume a expressão:

$$(1 - \phi_1 B - \dots - \phi_p B^p)(1 - B)^d Y_t = (1 - \theta_1 B - \dots - \theta_q B^q) \varepsilon_t$$

onde $(1 - B)^d Y_t$, com $d \geq 1$ é a série estacionária depois de diferenciada d vezes, $\phi_1, \phi_2, \dots, \phi_p$ são os parâmetros autoregressivos e $\theta_1, \theta_2, \dots, \theta_q$ os parâmetros médias móveis.

As formas mais simples e frequentes do modelo ARIMA são ARIMA(0,1,0), ARIMA(1,1,0), ARIMA(0,1,1) e ARIMA(1,1,1). Por exemplo, o modelo ARIMA(1,1,0) tem a representação:

$$(1 - \phi_1 B)(1 - B)Y_t = \varepsilon_t$$

ou, multiplicando os factores,

$$(1 - \phi_1 B + B + \phi_1 B^2)Y_t = \varepsilon_t,$$

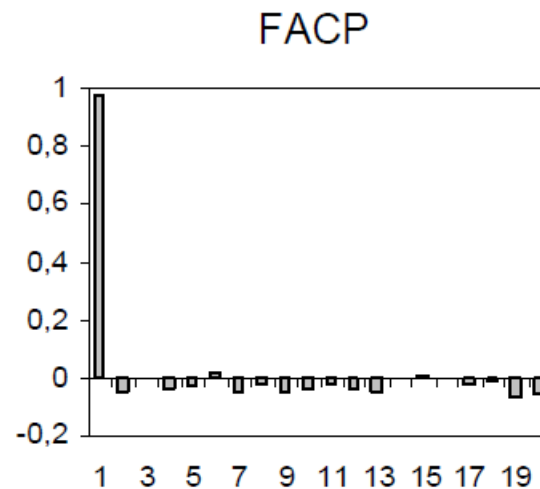
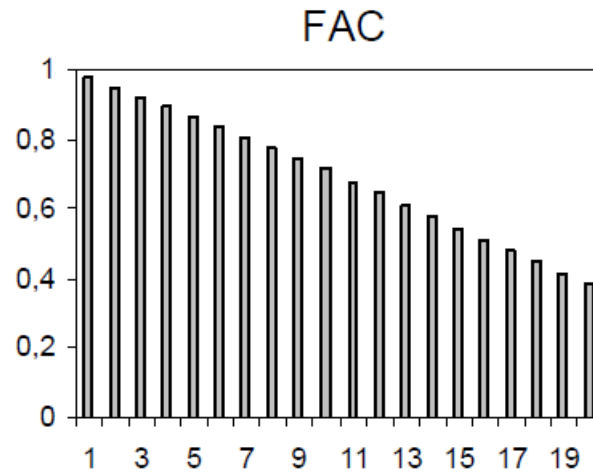
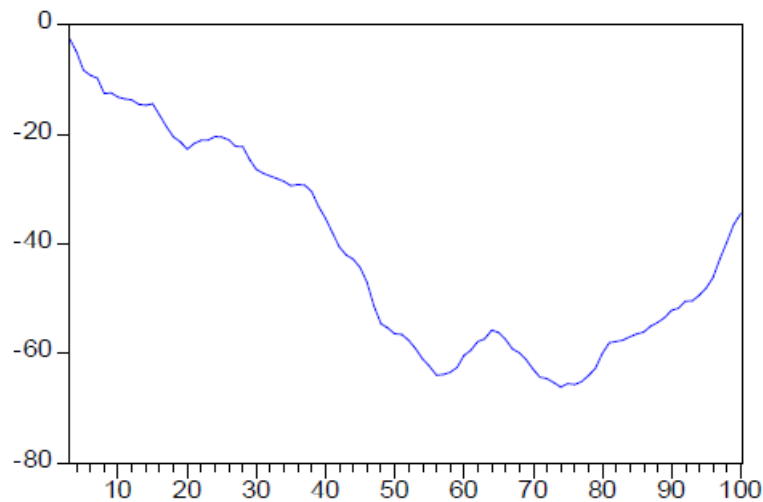
donde sai,

$$Y_t = (1 + \phi_1)Y_{t-1} - \phi_1 Y_{t-2} + \varepsilon_t.$$

Processos não estacionários

Modelo ARIMA(1,1,0) - simulação

$$(1 - 0,75B)(1 - B)Y_t = \varepsilon_t$$



Processos não estacionários

Modelo ARIMA(p, d, q)(P, D, Q) $_s$

Assim sendo, pode extender-se o modelo ARIMA a um modelo multiplicativo integrado sazonal representado simbolicamente por modelo SARIMA(p, d, q)(P, D, Q) $_s$. Por exemplo, o modelo SARIMA(0,1,1)(0,1,1) $_{12}$ tem a representação:

$$(1-B)(1-B^{12})Y_t = (1-\theta_1 B)(1-\Theta_1 B^{12})\varepsilon_t$$

ou, multiplicando os factores,

$$(1-B-B^{12}+B^{13})Y_t = (1-\theta_1 B-\Theta_1 B^{12}+\theta_1 \Theta_1 B^{13})\varepsilon_t,$$

tem-se a relação equivalente,

$$Y_t = Y_{t-1} - Y_{t-12} + Y_{t-13} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \Theta_1 \varepsilon_{t-12} + \theta_1 \Theta_1 \varepsilon_{t-13}.$$

Na maioria das aplicações práticas, os valores de p , q , P e Q são quase sempre inferiores ou iguais a 2, enquanto que os valores de d e D usualmente assumem os valores inteiros 0 ou 1. As FAC e FACP dos modelos SARIMA apresentam um decaimento exponencial ou sinusoidal amortecido sobre os *lags* não sazonais e sazonais, respectivamente.

Metodologia de análise

Identificação do modelo

A primeira etapa de modelação de uma série temporal consiste na identificação de um modelo $SARIMA(p,d,q)(P,D,Q)_s$ que descreva a relação existente entre as suas observações. Passo a dar:

- Construção do cronograma da série e sua estacionarização
- Estimação das FAC e FACP da série original
- Teste de raízes unitárias
- Estimação das FAC e FACP da série estacionária e identificação dos inteiros p , q , P e Q

Metodologia de análise

Estimação dos parâmetros

Uma vez identificados os modelos candidatos a descrever a série, segue-se a etapa de estimação dos seus parâmetros. Os dois principais métodos de estimação dos parâmetros do modelo ARIMA são:

Método da máxima verosimilhança, que consiste em determinar os valores dos parâmetros que tornam mais verosímil a ocorrência de um conjunto de observações idênticas aquelas de que efectivamente se dispõe. Este método obtém estimativas dos parâmetros através de um processo iterativo em que se maximiza a função de verosimilhança dos estimadores.

Método dos mínimos quadrados. Este método não permite obter estimadores consistentes com os verdadeiros parâmetros. Excepção feita nos casos em que se tem modelos com apenas componente autoregressiva, em que os estimadores se podem obter por minimização da soma de quadrados dos resíduos.

Metodologia de análise

Avaliação do diagnóstico

Esta etapa refere-se à avaliação da qualidade estatística das estimativas obtidas e da qualidade do ajustamento do modelo às observações da série em estudo.

- Avaliação da qualidade estatística das estimativas obtidas:
 - Teste de significância individual
 - Análise das correlações entre estimativas
- Avaliação da qualidade do ajustamento do modelo ARIMA (análise dos resíduos)
 - Teste de significância individual da FAC e da FACP
 - Teste de significância global da FAC
 - Teste de normalidade dos resíduos

Metodologia de análise

Seleção de modelos

Existem dois critérios de seleção de modelos que tomem em consideração as estatísticas baseadas nos resíduos do modelo ajustado:

- Critério de Informação de Akaike (AIC)

$$AIC = -2\ln L + 2m$$

No caso do Eviews: $AIC = n \ln \hat{\sigma}_\varepsilon^2 + 2m$

- Critério Bayesiano de Schwartz (BIC)

$$BIC = n \ln \hat{\sigma}_\varepsilon^2 + m \ln n$$

Metodologia de análise

Previsão

Suponhamos que temos no momento T as observações $Y_T, Y_{T-1}, Y_{T-2}, \dots$ de uma série temporal e pretendemos, com base nestas, prever o valor futuro no momento $T + m, Y_{T+m}$. A previsão para o instante $T + m$ será função dos valores presentes e passados da série, isto é:

$$\hat{Y}_T(m) = P_T(m) = E(Y_{T+m} | Y_T, Y_{T-1}, Y_{T-2}, \dots),$$

onde $P_T(m)$ é o preditor de Y_{T+m} , T é a origem da previsão e m é o horizonte da previsão.

No que se refere ao horizonte temporal, o estabelecimento de previsões reveste-se de duas formas: previsão estática (passo-a-passo) e previsão dinâmica (a múltiplos passos à frente).

Metodologia de análise

Previsão

Erros de previsão para modelos ARMA

Erros de previsão para modelos ARIMA

Previsão com o modelo AR(1)

$$\text{Estática: } \hat{Y}_{T+m-1}(1) = P_{T+m-1}(1) = E(Y_{T+m} | Y_{T+m-1}, \dots, Y_1) = \phi Y_{T+m-1}$$

$$\text{Dinâmica: } \hat{Y}_T(m) = P_T(m) = E(Y_{T+m} | Y_T, \dots, Y_1) = \phi \hat{Y}_{T+m-1} = \phi P_{T+m-1} = \phi(\phi^{m-1} Y_T) = \phi^m Y_T$$

$$\text{Variância do erro de previsão: } V[e_T(m)] = E[e_T^2(m)] = (1 + \phi^2 + \dots + \phi^{2m-2}) \sigma_\varepsilon^2$$

Metodologia de análise

Previsão

Previsão com o modelo MA(1)

$$\text{Estática: } \hat{Y}_{T+m-1}(1) = P_{T+m-1}(1) = E(Y_{T+m} | Y_{T+m-1}, \dots, Y_1) = E(c + \varepsilon_{T+m} - \theta\varepsilon_{T+m-1}) = c - \theta\hat{\varepsilon}_{T+m-1}$$

$$\text{Dinâmica: } \hat{Y}_T(m) = P_T(m) = E(Y_{T+m} | Y_T, \dots, Y_1) = E(c + \varepsilon_{T+m} - \theta\varepsilon_{T+m-1}) = c$$

$$\text{Variância do erro de previsão: } V[e_T(m)] = E[e_T^2(m)] = E[(Y_{T+m} - P_T(m))^2] = (1 + \theta^2)\sigma_\varepsilon^2$$

Previsão com o modelo ARMA(1,1)

$$\text{Estática: } \hat{Y}_{T+m-1}(1) = P_{T+m-1}(1) = E(Y_{T+m} | Y_{T+m-1}, \dots, Y_1) = c + \phi Y_{T+m-1} - \theta\varepsilon_{T+m-1}$$

$$\begin{aligned} \text{Dinâmica: } \hat{Y}_T(m) = P_T(m) &= E(Y_{T+m} | Y_T, \dots, Y_1) = E(c + \phi Y_{T+m-1} + \varepsilon_{T+m} - \theta\varepsilon_{T+m-1}) \\ &= c(\phi^{m-1} + \dots + \phi + 1) + \phi^m Y_T - \phi^{m-1} \theta \varepsilon_T. \end{aligned}$$

$$\text{Variância do erro de previsão: } V[e_T(m)] = E[e_T^2(m)] = \left\{ 1 + \sum_{j=1}^{m-1} [\phi^{m-1}(\phi - \theta)]^2 \right\} \sigma_\varepsilon^2$$

Metodologia de análise

Intervalos de previsão

Os intervalos de previsão são baseados na variância dos erros de previsão. Para construir um intervalo de previsão a m passos à frente (admitindo a normalidade dos erros de previsão), devemos utilizar a seguinte expressão:

$$\{P_T(m) - z\sqrt{V(e_T(m))}, P_T(m) + z\sqrt{V(e_T(m))}\},$$

onde z é o valor que limita o intervalo de previsão e corresponde a uma determinada probabilidade ou nível de confiança. Os graus de confiança mais utilizados em intervalos de previsão são 90%, 95% e 99%, a que correspondem valores de z iguais a 1,645, 1,96 e 2,576, respectivamente.

Aplicações do modelo ARIMA

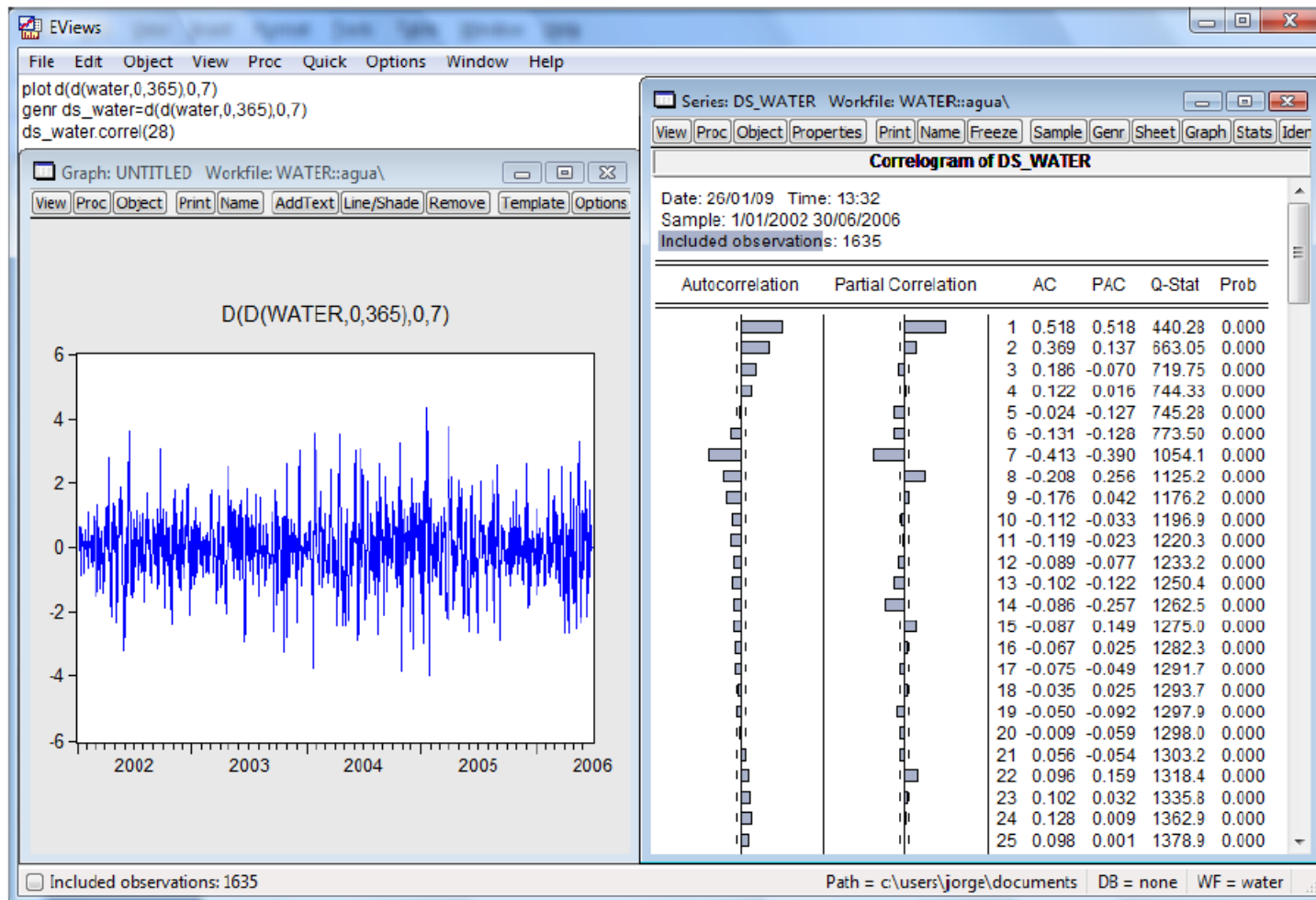
Consumo de água

Para ilustrar a metodologia de modelação e previsão ARIMA de séries temporais, considerou-se uma série de consumo diário de água em Granada (Espanha) observada no período de 1 de Janeiro de 2002 a 30 de Junho de 2006, num total de 2006 observações. Os dados foram obtidos através da Spanish IEEE Computational Intelligence Society (Fonte: <http://www.congresocedi.es/2007>) e utilizados pelo autor num estudo de análise comparativa de modelos de previsão estocásticos e determinísticos com dupla sazonalidade para previsão do consumo de água (veja-se Caiado, 2009).

A observação correspondente ao dia 29 de Fevereiro do ano bissexto 2004 foi eliminada da amostra de modo a manter o mesmo número de dias (365) em cada ano. Para a amostra de estimação (“training sample”), foram utilizadas as primeiras 1976 observações entre 1 de Janeiro de 2002 e 31 de Maio de 2006 e para a amostra pós-estimação de avaliação de previsões (“post-sample”), foram utilizadas as restantes 30 observações entre 1 e 30 de Junho de 2006. Para o estudo de modelação e previsão foi utilizado o programa EVIEWS.

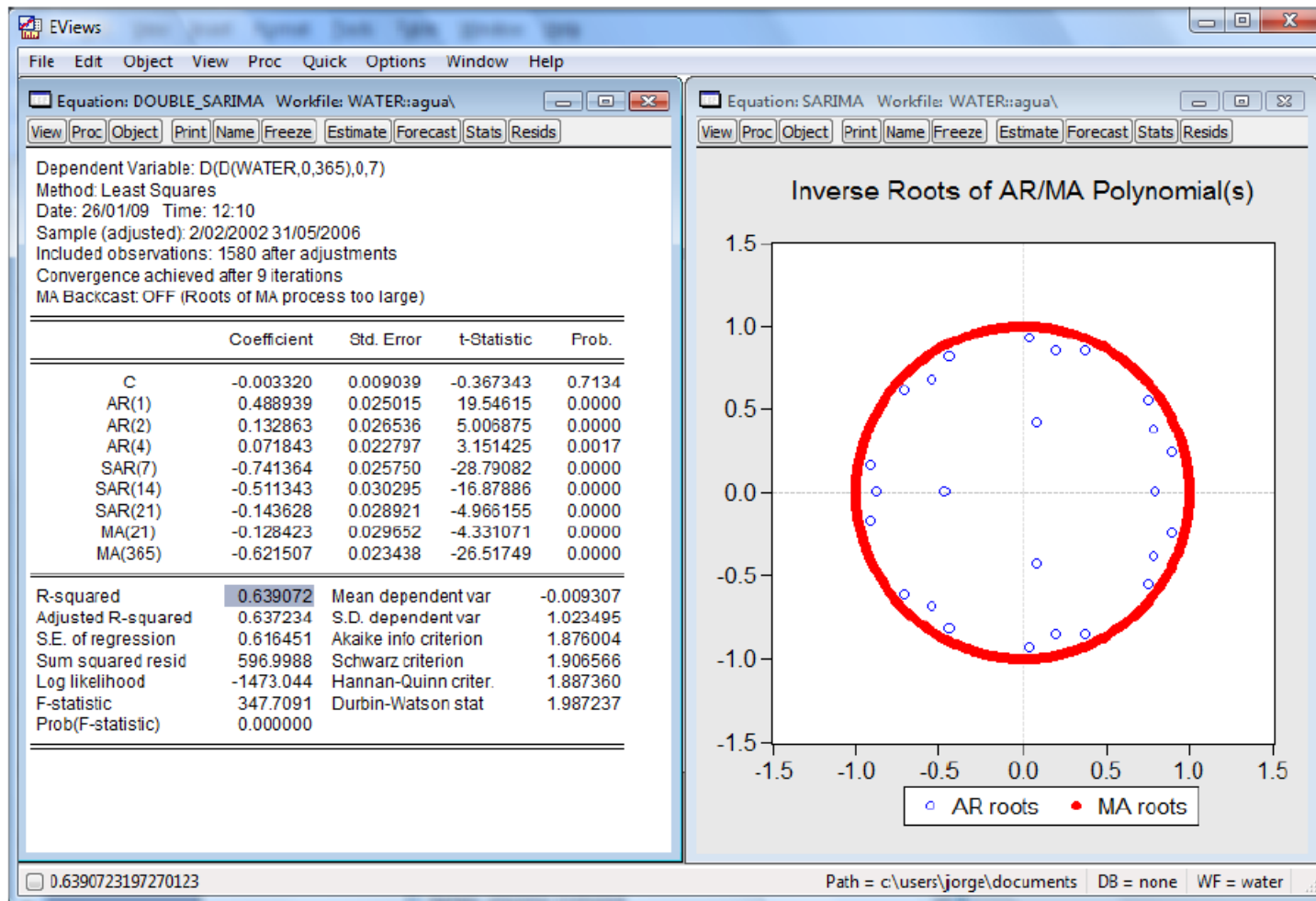
Aplicações do modelo ARIMA

Consumo de água



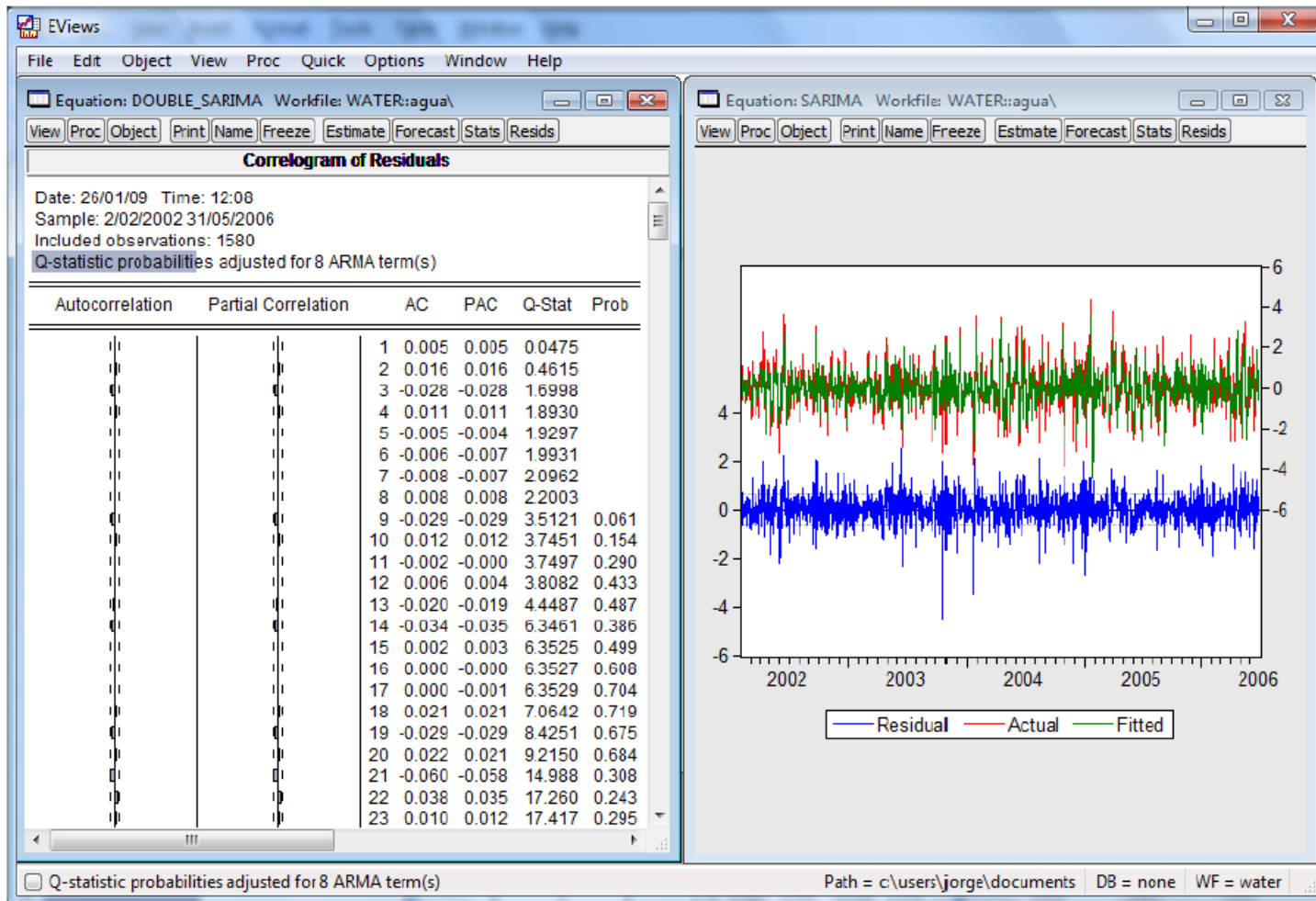
Aplicações do modelo ARIMA

Consumo de água



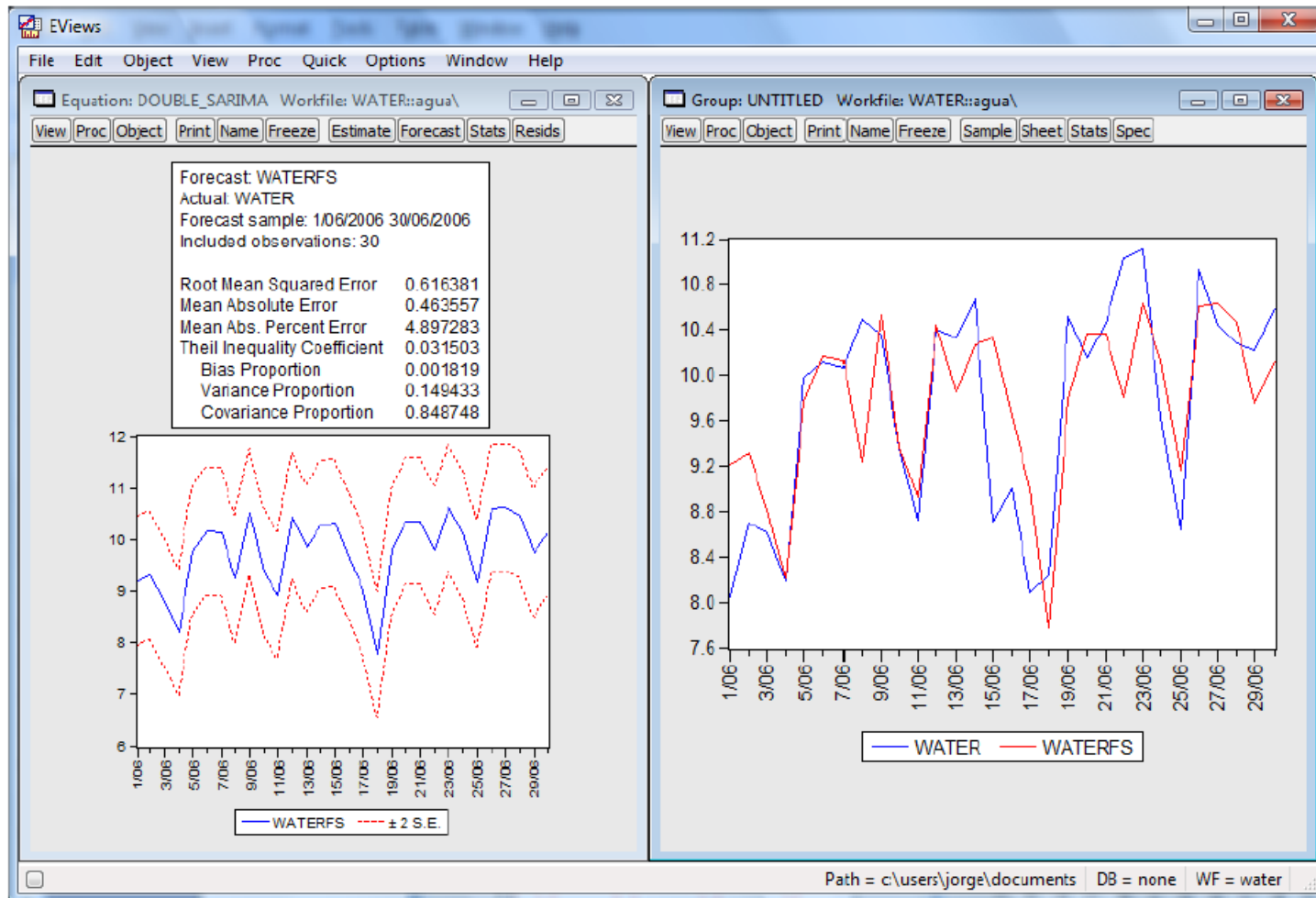
Aplicações do modelo ARIMA

Consumo de água



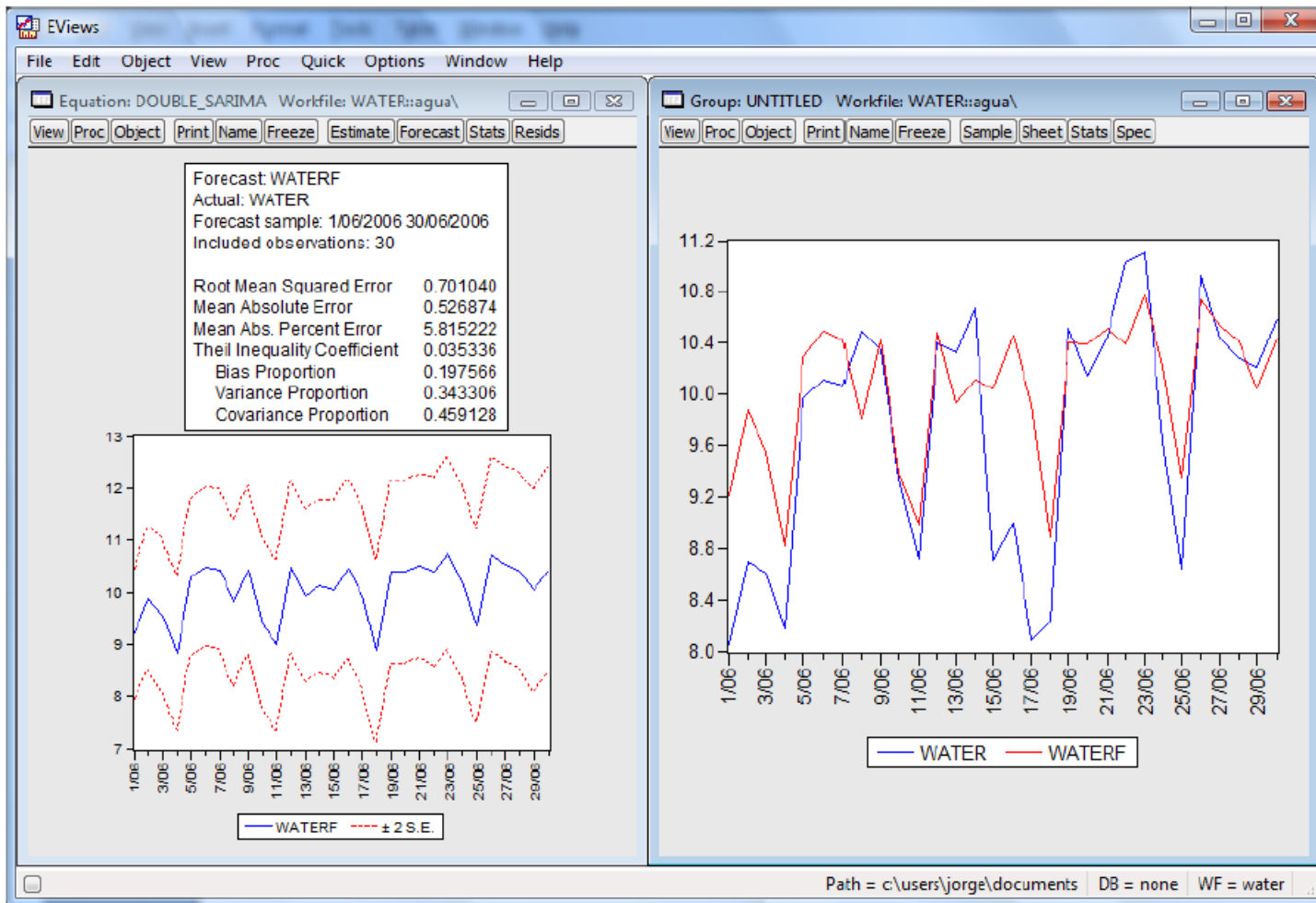
Aplicações do modelo ARIMA

Consumo de água



Aplicações do modelo ARIMA

Consumo de água



Modelos ARMA

Exercícios de aplicação

No ficheiro em EViews “Dados_ST.wk1” – page “COFFEE”, encontram-se os dados do consumo de café nos EUA entre 1910 e 1967.

- a) Construa o cronograma e analise o andamento da série.
- b) Ajuste um modelo ARMA aos dados observados entre 1910 e 1963, seguindo as etapas habituais de identificação, estimação, avaliação do diagnóstico e selecção de modelos.
- c) Estabeleça previsões dinâmicas e estáticas para os 4 anos seguintes e avalie os erros de previsão EQM, EAM e EPAM.

No ficheiro em EViews “Dados_ST.wk1” – page “HOUSE”, encontra-se uma série temporal de vendas de moradias unifamiliares entre 1987M1 e 1995M11.

- a) Construa o cronograma e analise o andamento da série.
- b) Ajuste um modelo SARMA aos dados observados entre 1987M1 e 1994M11, seguindo as etapas habituais de identificação, estimação, avaliação do diagnóstico e selecção de modelos.
- c) Estabeleça previsões dinâmicas e estáticas para o período 1994M12-1994M11 e avalie os erros de previsão EQM, EAM e EPAM.

Modelos de intervenção e detecção de outliers

Modelo de intervenção: Função degrau e impulso

O objectivo da análise de intervenção consiste em avaliar o impacto de um ou mais choques exógenos no comportamento de uma determinada série cronológica, mais concretamente, série *output*.

Para caracterizar e quantificar a natureza e a magnitude desses efeitos, há que distinguir entre dois tipos de variáveis de intervenção:

a) Variável “degrau” (*step variable*), representa uma intervenção que ocorre no momento $t = T$ e que permanece actuante desde então,

$$S_t^{(T)} = \begin{cases} 0, & t < T, \\ 1, & t \geq T. \end{cases} \quad (1)$$

b) Variável “impulso” (*pulse variable*), correspondente a uma intervenção que ocorre no momento $t = T$ e que tem efeito apenas nesse instante,

$$P_t^{(T)} = \begin{cases} 1, & t = T, \\ 0, & t \neq T. \end{cases} \quad (2)$$

Modelos de intervenção e detecção de outliers

Função degrau e impulso

Para apenas uma variável de intervenção — “degrau” ou “impulso”, designando genericamente por I_t , o modelo que relaciona o *output* com o *input* é definido pela relação:

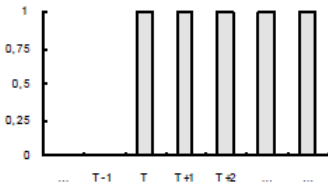
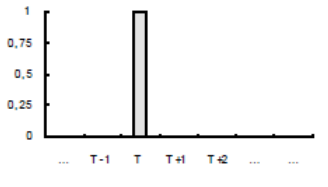
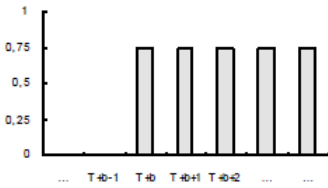
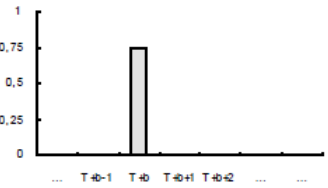
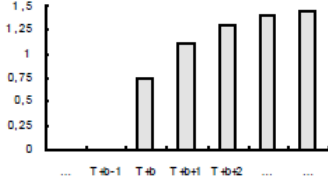
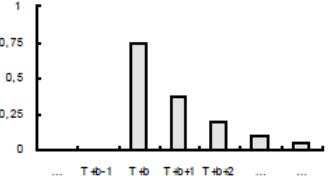
$$Y_t = \frac{\omega_s(B)}{\delta_r(B)} B^b I_t + N_t, \quad (3)$$

onde $[\omega_s(B)/\delta_r(B)]B^b$ representa a resposta do *output* à intervenção (vulgarmente conhecida por função transferência), sendo $\omega_s(B) = \omega_0 - \omega_1 B - \dots - \omega_s B^s$ e $\delta_r(B) = 1 - \delta_1 B - \dots - \delta_r B^r$ polinómios de graus s e r , respectivamente, e b representa o número de períodos que leva a intervenção a produzir um efeito no *output*; e N_t a série residual ou *noise* (ou a série do *output* sem a presença da intervenção) e pode representar-se por um processo ARMA(p, q) estacionário e invertível: $\phi_p(B)N_t = \theta_q(B)\varepsilon_t$.

Ainda na relação (3), os coeficientes $\omega_0, \omega_1, \dots, \omega_s$ do polinómio $\omega_s(B)$ representam os **efeitos iniciais** da intervenção e os coeficientes $\delta_1, \delta_2, \dots, \delta_r$ do polinómio $\delta_r(B)$ os **efeitos permanentes** da intervenção.

Modelos de intervenção e detecção de outliers

Respostas às intervenções "degrau" e "impulso"

Intervenção "Degrau"		Intervenção "Impulso"	
$S_t^{(I)}$		$P_t^{(I)}$	
Respostas			
$\omega B^b S_t^{(I)},$ $\omega = 0,75$		$\omega B^b P_t^{(I)},$ $\omega = 0,75$	
$\frac{\omega B^b}{1-\delta B} S_t^{(I)},$ $\omega = 0,75,$ $\delta = 0,5$		$\frac{\omega B^b}{1-\delta B} P_t^{(I)},$ $\omega = 0,75,$ $\delta = 0,5$	

Modelos de intervenção e detecção de outliers

Respostas às intervenções "degrau" e "impulso"

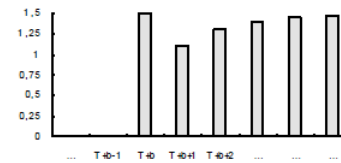
Respostas			
$\frac{\omega B^b}{1-\delta B} S_t^{(T)},$ $\omega = -0,75,$ $\delta = 0,5$		$\frac{\omega B^b}{1-\delta B} P_t^{(T)},$ $\omega = -0,75,$ $\delta = 0,5$	
$\frac{\omega B^b}{1-\delta B} S_t^{(T)},$ $\omega = 0,75,$ $\delta = 1$		$\frac{\omega B^b}{1-\delta B} P_t^{(T)},$ $\omega = 0,75,$ $\delta = 1$	
$\frac{\omega B^b}{1-\delta B} S_t^{(T)},$ $\omega = -0,75,$ $\delta = 1$		$\frac{\omega B^b}{1-\delta B} P_t^{(T)},$ $\omega = -0,75,$ $\delta = 1$	

Modelos de intervenção e detecção de outliers

Respostas às intervenções "degrau" e "impulso" combinadas

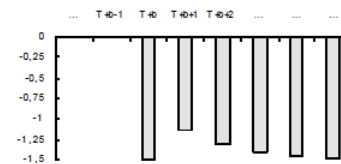
$$\frac{\omega_0 B^b}{1-\delta B} S_t^{(T)} + \omega_1 B^b P_t^{(T)}$$

$\omega_0 = 0,75, \omega_1 = 0,75, \delta = 0,5$



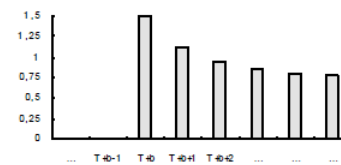
$$\frac{\omega_0 B^b}{1-\delta B} S_t^{(T)} + \omega_1 B^b P_t^{(T)}$$

$\omega_0 = -0,75, \omega_1 = -0,75, \delta = 0,5$



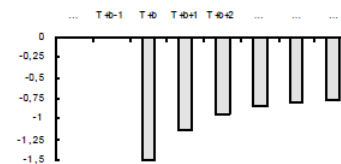
$$\frac{\omega_0 B^b}{1-\delta B} P_t^{(T)} + \omega_1 B^b S_t^{(T)}$$

$\omega_0 = 0,75, \omega_1 = 0,75, \delta = 0,5$



$$\frac{\omega_0 B^b}{1-\delta B} P_t^{(T)} + \omega_1 B^b S_t^{(T)}$$

$\omega_0 = -0,75, \omega_1 = -0,75, \delta = 0,5$



Modelos de intervenção e detecção de outliers

Metodologia de análise de intervenção

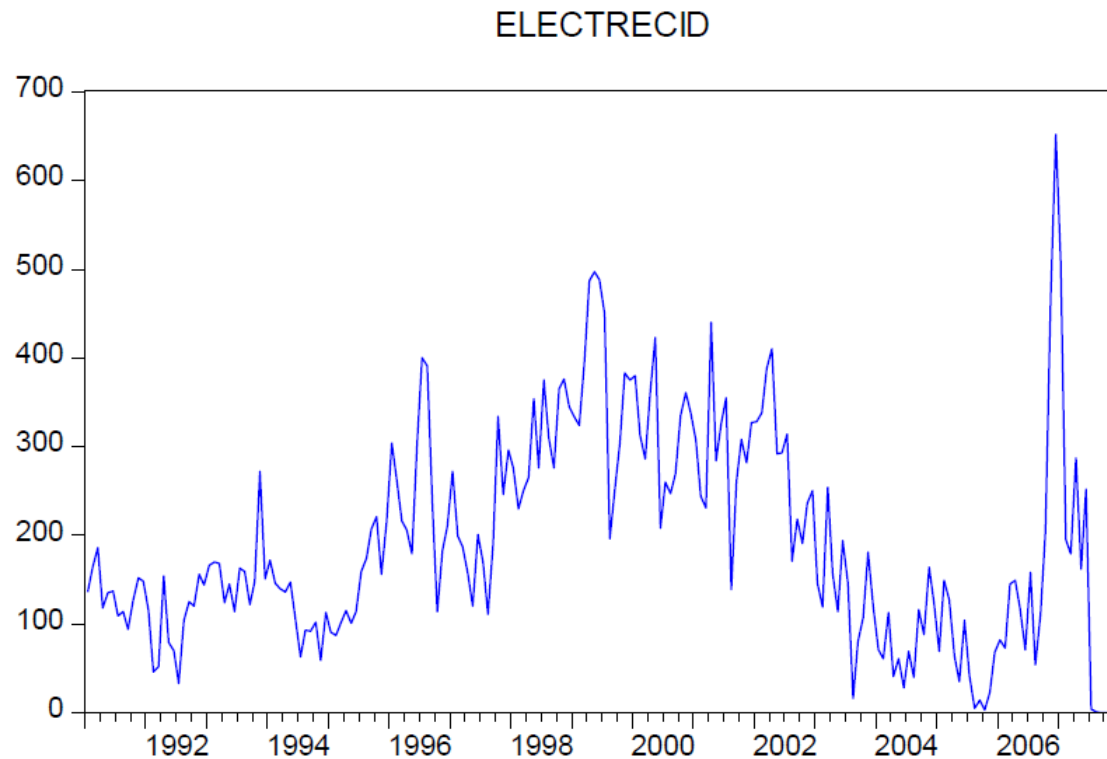
A especificação de um modelo particular de intervenção pode começar pela especificação do modelo ARIMA (p,d,q) representativo da série residual N_t . Na metodologia de análise, são conhecidas três formas alternativas de identificação do *noise*:

- A primeira, introduzida por Box e Tiao (1975), consiste em utilizar as observações da série *output* Y_t anteriores à data de ocorrência da primeira intervenção, $\{Y_t : t < T\}$, uma vez que, assumindo a variável de intervenção o valor zero nesse período, $I_t^T = 0, t < T$, a série do *noise* vem dada pela série do *output*, isto é, $Y_t = N_t, t < T$;
- Outra hipótese, seria utilizar todas as observações da série *output* excluindo o intervalo (ou intervalos) de tempo em que ocorre a intervenção (ou intervenções) e se fazem sentir os seus efeitos;
- Outros autores referem uma outra alternativa de identificação, que consiste em incluir todas as observações disponíveis do *output* no modelo se estas forem em número suficientemente grande de forma a atenuar eventuais distorções nas autocorrelações e autocorrelações parciais resultantes da ocorrência da intervenção (ou intervenções).

Modelos de intervenção e detecção de outliers

Exercício de aplicação

Consumo de eletricidade, entre 1991M1 e 2007M12 (204 obs. mensais). File: Dados_ST2012 Page: Energia_PT Série: Electrecid_exp



Construa um modelo de intervenção para a série observada, avaliando o efeito da subida “abrupta” do consumo de electricidade em 2006M11 (Obs. 191)

Modelos de intervenção e detecção de outliers

Outliers em séries temporais

Quando as datas de ocorrência dos fenómenos que influenciam influenciam o comportamento da série não são conhecidos *a priori*, os *inputs* tomam o nome de **outliers**.

Existem essencialmente quatro modelos paramétricos para a detecção de *outliers* que podem ser caracterizados segundo os efeitos que produzem na série *output*.

a) **Additive Outlier**, representa uma variável do tipo “impulso” cujo efeito apenas ocorre em $t=T$, sendo nulo o efeito nos restantes momentos. É definido através do modelo,

$$Y_t = Z_t + \omega I_t^{[T]}, \quad (5)$$

onde Z_t representa a série sem a presença de *outliers*, que admite-se seguir um processo ARMA(p,q) invertível e estacionário: $\phi_p(B)Z_t = \theta_q(B)\varepsilon_t$; e $I_t^{[T]}$ é uma variável binária que indica a presença ou ausência de um *outlier* no momento T , através da relação: $I_t^{[T]} = 1$ se $t = T$ e $I_t^{[T]} = 0$ se $t \neq T$.

Modelos de intervenção e detecção de outliers

Outliers em séries temporais

b) Innovation Outlier, traduz-se num choque na série residual, cujo efeito faz-se sentir em todas as observações a partir do momento t ($Y_t, Y_{t+1}, Y_{t+2}, \dots$) segundo a estrutura do sistema descrito por $\theta_q(B)/\phi_p(B)$. Pode descrever-se através da relação,

$$Y_t = Z_t + \frac{\theta_q(B)}{\phi_p(B)} \omega I_t^{[T]},$$

c) Level Shift, corresponde a uma variável do tipo “degrau”, cujo efeito ocorre em $t = T$ e mantém-se actuante depois desse momento. Define-se através da expressão,

$$Y_t = Z_t + \frac{\omega}{1-B} I_t^{[T]},$$

d) Temporary (ou Transient) Change, refere-se a um choque na série para $t \geq T$, cujo efeito tende a diminuir a um ritmo exponencial após um impacto inicial. Este outlier é definido através do modelo,

$$Y_t = Z_t + \frac{\omega}{1-\delta B} I_t^{[T]},$$

Modelos de intervenção e detecção de outliers

Metodologia de detecção de outliers

A detecção de *outliers* (essencialmente de dois tipos, *Additive Outlier* e *Level Step*) pode ser feita através do seguinte procedimento: começa-se por construir um modelo ARIMA para a série original do *output* Y_t , supondo que não existe nenhum *outlier*, seguindo-se um processo iterativo de regressão dos resíduos obtidos sobre potenciais *outliers*, até conseguir-se, avaliando os efeitos dos seus choques, encontrar “todos”² os *Additive Outlier* e *Level Step*.

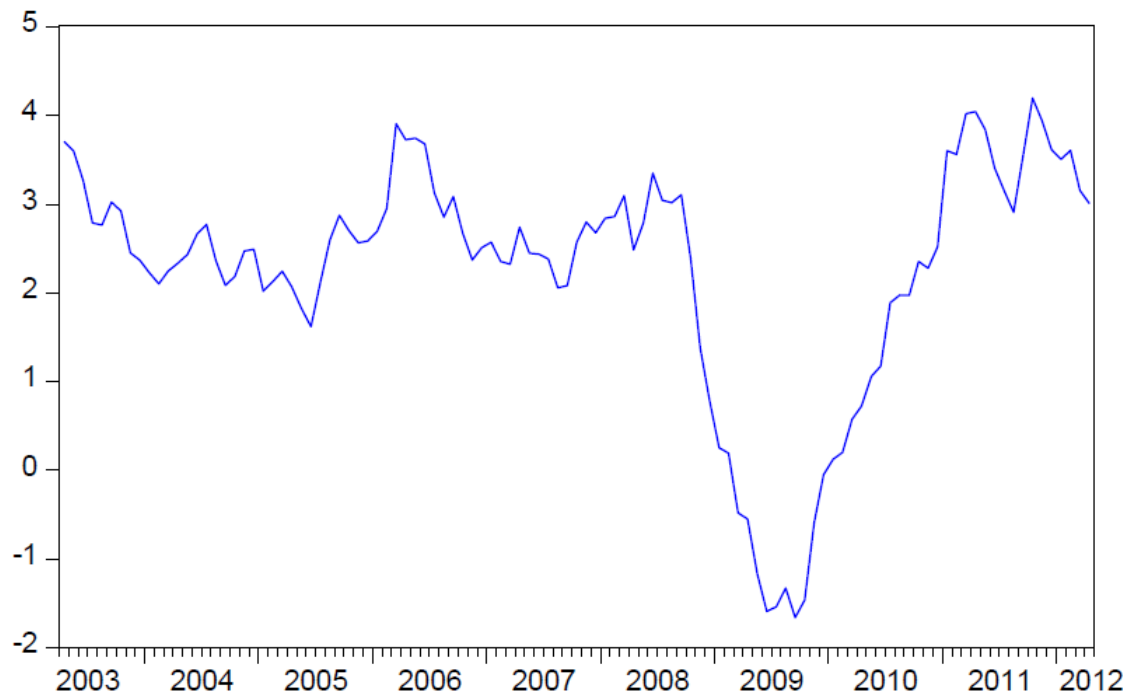
Uma vez encontrados, os *outliers* (cujos momentos de ocorrência passam a ser conhecidos) são incluídos como variáveis de intervenção (“Impulso” e “Degrau”) nos modelos de intervenção introduzidos por Box e Tiao (1975), procedendo-se à identificação da função transferência ou da resposta do *output* às intervenções e de um modelo ARIMA para a série residual com coeficientes estatisticamente significativos. Uma vez especificado a forma do modelo passa-se às etapas habituais de estimação e avaliação do diagnóstico.

Víctor Gomez e Augustín Maravall (Gomez e Maravall, 1994) desenvolveram o software TRAMO (“Time series regression with ARIMA noise, missing observations and outliers”) para a modelação e detecção automática de outliers de vários tipos. Este software está disponível no EViews.

Modelos de intervenção e detecção de outliers

Exercício de aplicação

Taxa de variação homóloga do IPC (taxa de inflação) em Portugal, 2003m4-2012m4 (109 observações mensais) File: Dados_ST2012 Page: Taxas_POR
Série: INFL



Utilizando o software TRAMO, proceda à detecção de outliers na série observada.

Referências bibliográficas

- Box, G. P. E., Jenkins, G. M. e Reinsel, G. (1994). Time Series Analysis: Forecasting and Control, 3rd ed., Prentice-Hall.
- Brockwell, P. J. e Davis, R. A. (1991). Time Series: Theory and Methods. 2nd ed., Springer, New York.
- Caiado, J. (2010). "[Performance of combined double seasonal univariate time series models for forecasting water demand](#)", *Journal of Hydrologic Engineering*, 15, 215-222.
- Hamilton, J. D. (1994). *Time Series Analysis*, Princeton University Press.
- Peña, D., Tiao, G. e Tsay, R. (2001). *A Course in Time Series Analysis*, Wiley, New York.
- Shumway, R. H. and Stoffer, D. S. (2000). *Time Series Analysis and Its Applications*. Springer-Verlag, New York.
- Wei, W. W. S. (2007). *Time Series Analysis: Univariate and Multivariate Methods*. 2nd ed., Addison-Wesley Publishing Company, Redwood City, California