Linear Time Series Analysis

Jorge Caiado
CEMAPRE/ISEG, University of Lisbon
Email: jcaiado@iseg.ulisboa.pt
Web: http://pascal.iseg.utl.pt/~jcaiado/

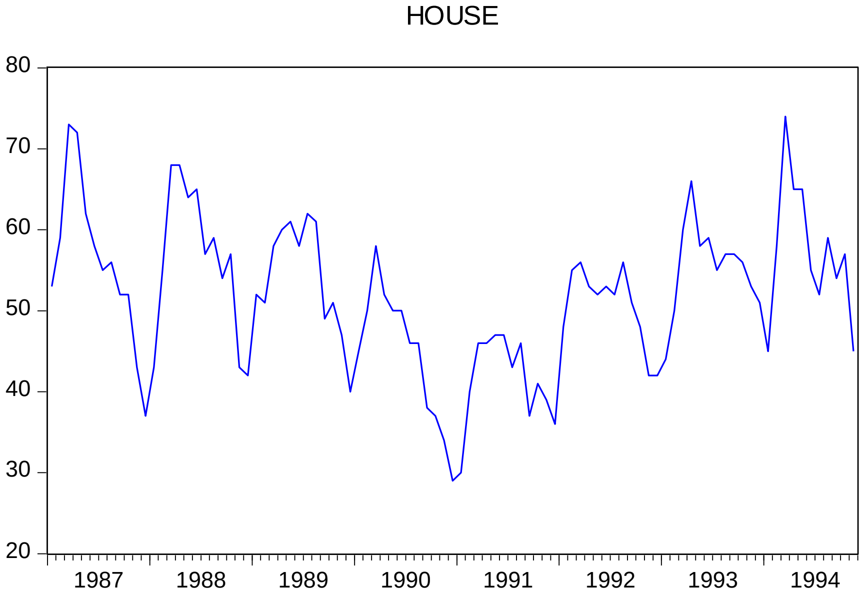
Outline

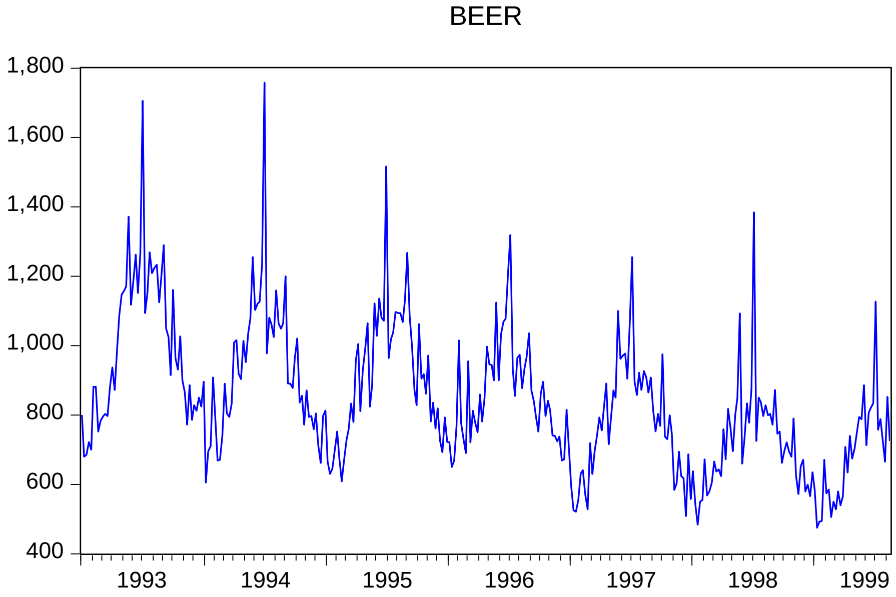
- ❑ Stationarity
- ❑ Autocorrelations and partial autocorrelations
- ❑ White noise
- ❑ Estimation of the autocorrelations and partial autocorrelations
- ❑ Moving average and autoregressive representations of time series processes
- ❑ Stationary time series models: AR, MA, ARMA, SAR, SMA and SARMA models
- ❑ Nonstationary time series models: nonstationary in mean, nonstationary in the variance, ARIMA model and SARIMA model
- ❑ Model identification, parameter estimation, diagnostic checking and model selection
- ❑ Forecasting

Stationarity

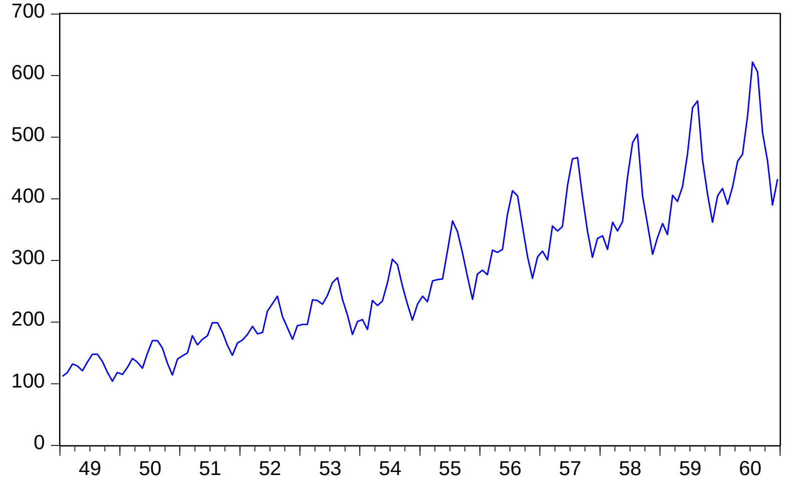
Definition:

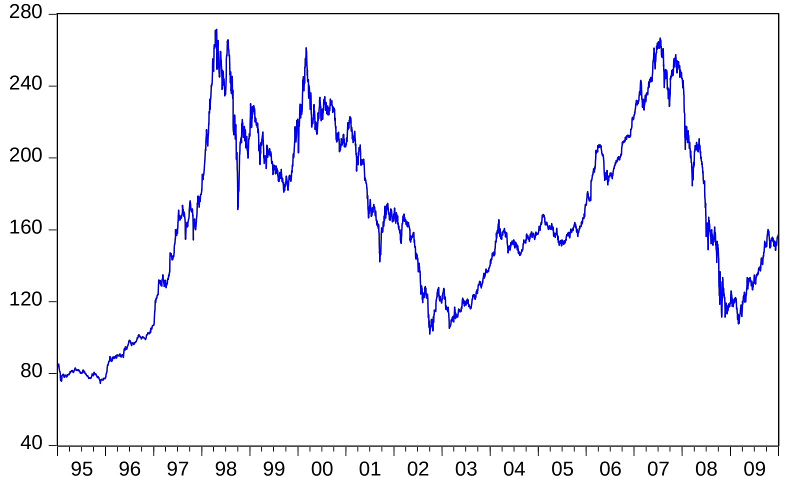
A **time series** is a sequence of observations over time. For example: monthly sales of new one-family houses; Daily stock indices; Weekly beer consumption; daily average temperature; Annual electricity production.

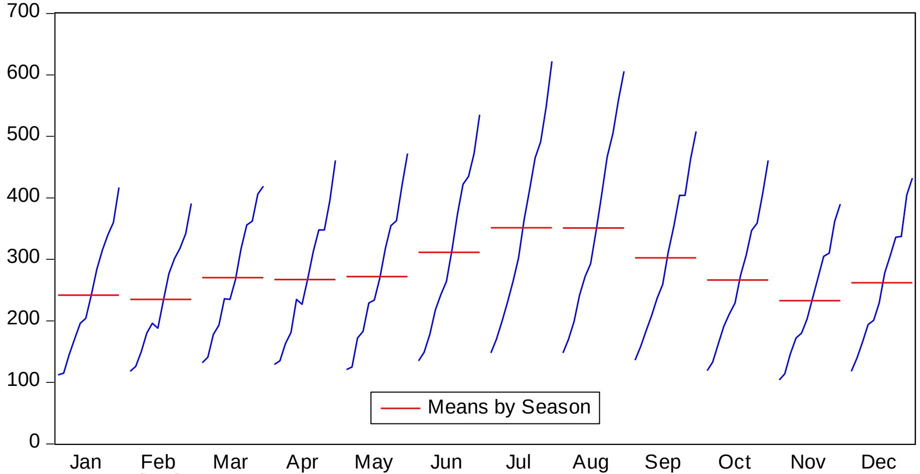


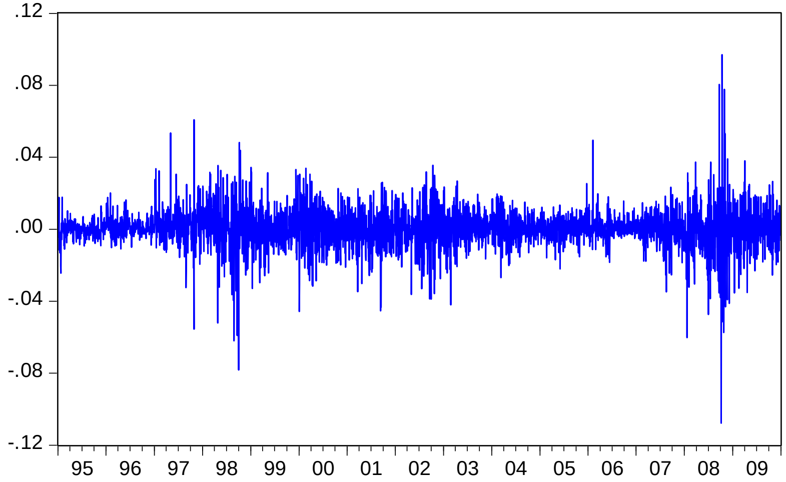


Stationarity


PASSAG


POR


PASSAG by Season


Log Differenced POR

Stationarity

Definitions:

A **stochastic process** is a family of time indexed random variables, $Z(w,t)$: $t=0,\pm1, \pm2, \ldots$, where w is the sample space and t is the index set.

A time series is a realization (or sample function) from a certain stochastic process, Yt, $t=1,2,\ldots,n$.

A process Yt, $t=1,2,\ldots,n$ is said to be **strictly stationary** if it has constant mean, constant variance, and the covariance and the correlation between Yt and $Yt+k$ depend only on time difference k.

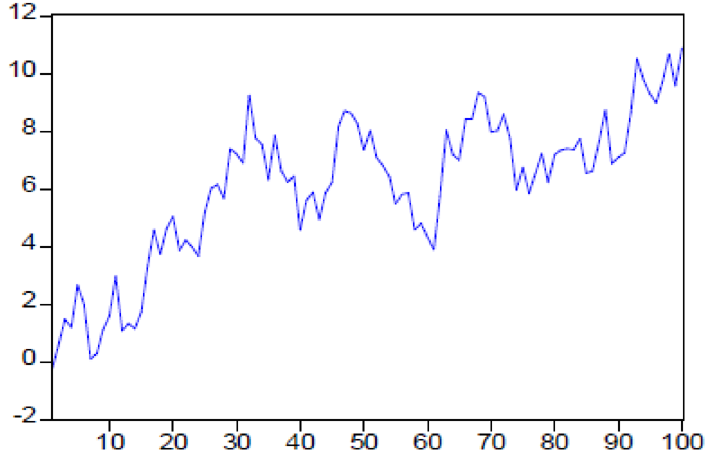
$$\mu_t = E(Y_t) = \mu \,,$$

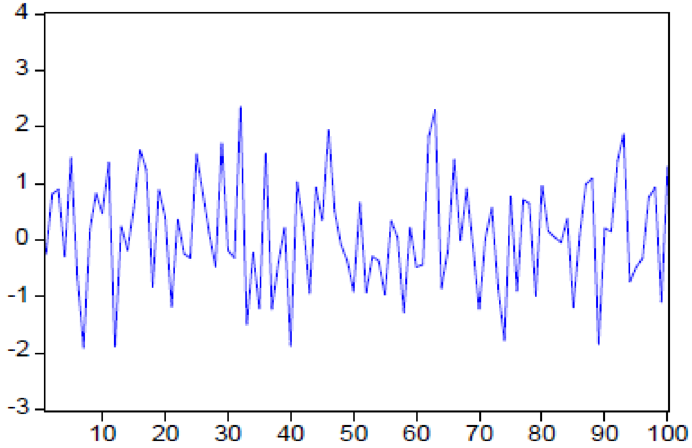
$$\sigma_t^2 = Var(Y_t) = E(Y_t - \mu_t)^2 = \sigma^2 \,,$$

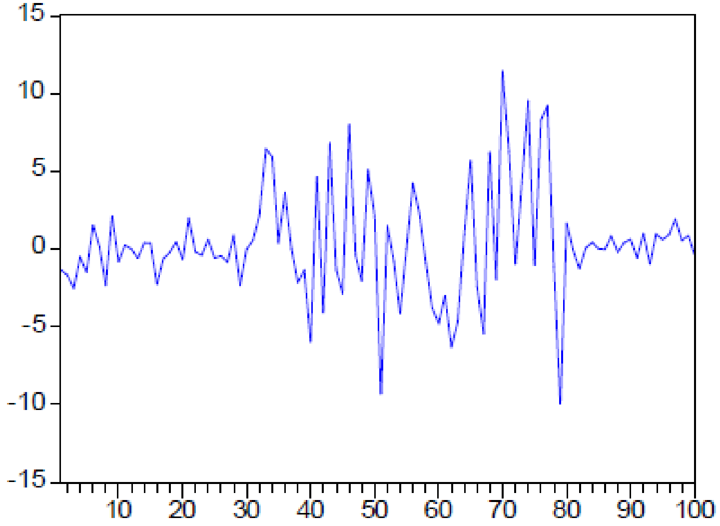
$$\gamma(t_1,t_2) = E(Y_{t_1} - \mu_{t_1})(Y_{t_2} - \mu_{t_2}) = \gamma(t_1 + k, t_2 + k) \,, \ \forall t_1, t_2, k$$

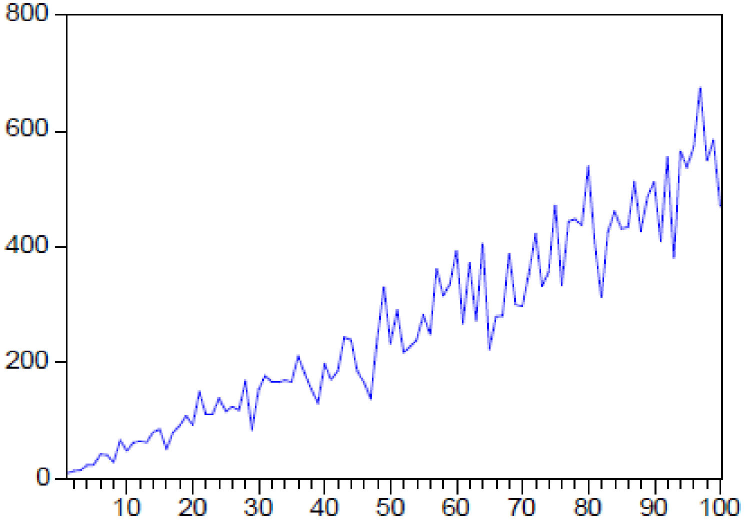
$$\rho(t_1,t_2) = \frac{\gamma(t_1,t_2)}{\sqrt{\sigma_{t_1}^2}\sqrt{\sigma_{t_2}^2}} = \rho(t_1 + k, t_2 + k) \,, \ \forall t_1, t_2, k \,,$$

Stationary?


12
10
8
6
4
2
0
-2
10
20
30
40
50
60
70
80
90
100


4
3
2
1
0
-1
-2
-3
10
20
30
40
50
60
70
80
90
100


15
10
5
0
-5
-10
-15
10
20
30
40
50
60
70
80
90
100


800
600
400
200
0
10
20
30
40
50
60
70
80
90
100

Autocorrelation

The **autocovariance function (ACOVF) and autocorrelation function (ACF)** represent the covariance and correlation between *Yt* and *Yt+k*, from the same process *Y* separated only by *k* time lags.

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

The autocovariance function and the autocorrelation function have the following properties:

1) $\gamma_0 = Var(Y_t)$; $\rho_0 = 1$;

2) $|\gamma_k| \leq \gamma_0$; $|\rho_k| \leq 1$;

3) $\gamma_k = \gamma_{-k}$; $\rho_k = \rho_{-k}$ for all k;

4) γ_k and ρ_k are positive semidefinite.

Partial autocorrelation

The **partial autocorrelation function (PACF)** measures the correlation between Y_t and Y_{t-k}, when the effects of intervening variables Y_{t-1}, Y_{t-2}, …, Y_{t-k+1} are removed. The partial autocorrelation coefficient of order k is denoted by ϕ_{kk} and can be derived by regressing Y_{t+k} against Y_{t+k-1}, Y_{t+k-2}, …, Y_t:

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

Multiplying Y_{t+k-j} on both sides of the equation and taking expected values, we get

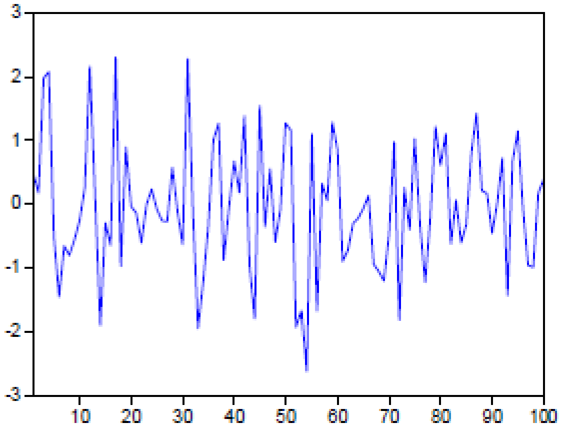
$$\phi_{11} = \rho_1 , \; \phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} , \; \phi_{33} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}} , \; \ldots , \; \phi_{kk} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_{k-2} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & 1 \end{vmatrix}}$$

White noise

A process is called a “white noise” process if it is a sequence of uncorrelated random variables:

$$Y_t = \varepsilon_t \, ,$$

where ε_t has constant mean $E(\varepsilon_t) = \mu_\varepsilon$ (usually assumed to be 0) , constant variance $Var(\varepsilon_t) = \sigma_\varepsilon^2$ and null covariance $Cov(\varepsilon_t, \varepsilon_{t-k}) = 0$ for all $k \neq 0$. The ACF and PACF of a white noise process are null for all $k \neq 0$.



Simulation of a white noise process with zero mean and unit variance

Sample ACF and PACF

For a given observed time series, $Y_t, t = 1,2,\ldots,n$, the sample autocorrelation function (ACF) is defined as

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

The sample partial autocorrelation function (PACF) is obtained by a recursive method as follows:

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

with $\hat{\phi}_{11} = \hat{\rho}_1$ and $\hat{\phi}_{kj} = \hat{\phi}_{k-1,j} - \hat{\phi}_{kk}\hat{\phi}_{k-1,k-j}$, $j = 1,2,\ldots,k-1$.

Backshift notation

A very useful notation in time series analysis is the backshift operator *B*, which is used as follows:

$$BY_t = Y_{t-1}.$$

In other words, *B* has the effect of shifting the data back one period.
For *k* periods, the notation is

$$B^k Y_t = Y_{t-k}.$$

For monthly data, B^{12} is used to shift attention to the same month last year, $B^{12}Y_t = Y_{t-12}$.

For quarterly data, the backshift operator is used as follows: $B^4 Y_t = Y_{t-4}$.

MA(∞) representation

The process Y_t can be expressed as a linear combination of a sequence of uncorrelated random variables:

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

where $\psi_0 = 1$, ε_t is a zero mean white noise with constant variance and $\sum_{j=0}^{\infty} \psi_j^2 < \infty$.

It can be shown that

$$E(Y_t) = 0\,,\quad Var(Y_t) = \sigma_\varepsilon^2 \sum_{j=0}^{\infty} \psi_j^2\,,\quad E(\varepsilon_t Y_{t-k}) = \begin{cases} \sigma_\varepsilon^2, & k = 0 \\ 0, & k > 0, \end{cases}$$

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

AR(∞) representation

Another useful form is to write Y_t in an autoregressive representation, as follows:

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

or, equivalently,

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

where $\pi(B) = 1 - \pi_1 B - \pi_2 B^2 - \cdots = 1 - \sum_{j=1}^{\infty} \pi_j B^j$ and $1 + \sum_{j=1}^{\infty} \left| \pi_j \right| < \infty$.

Autoregressive models

The finite-order representation of the autoregressive model described earlier, if only a finite number of π weights are nonzero, is given by

$$Y_t = \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + \varepsilon_t ,$$

where ε_t is a zero mean white noise series. Because $\sum_{j=1}^{\infty} |\pi_j| = \sum_{j=1}^{p} |\phi_j| < \infty$, the process is always invertible. To be stationaty, the roots of $(1 - \phi_1 B - \cdots - \phi_p B^p) = 0$ must be outside of the unit circle.

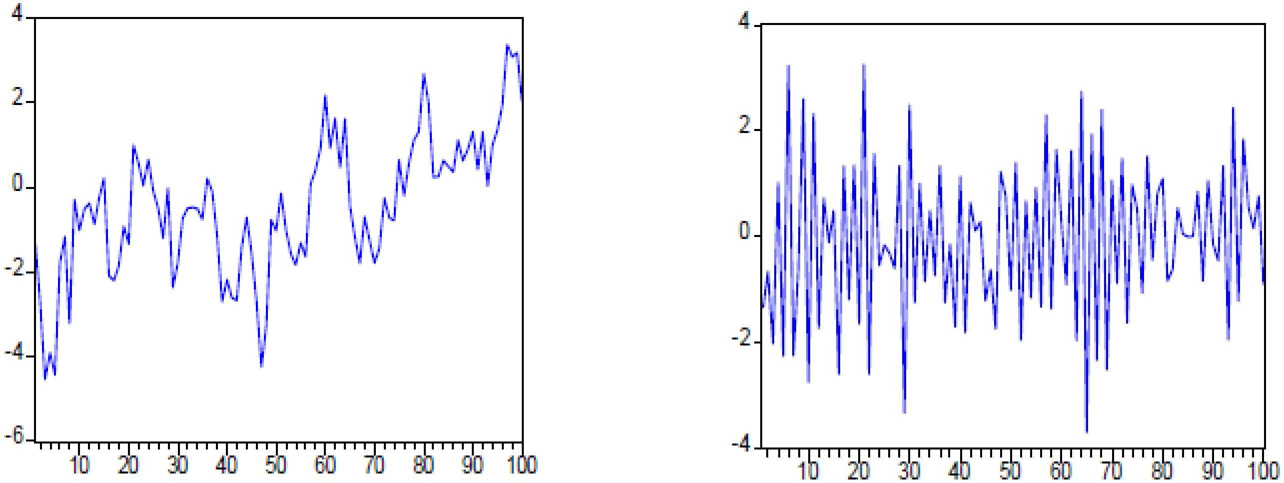
AR(1) model

The first-order autoregressive model or AR(1) model is given by

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

where ε_t is a zero mean white noise series. The model is always invertible. To be stationary, the roots of $(1 - \phi B) = 0$ must be outside of the unit circle. Because the root $B = 1/\phi$, for a stationary model, we have $|\phi| < 1$.

Autoregressive models

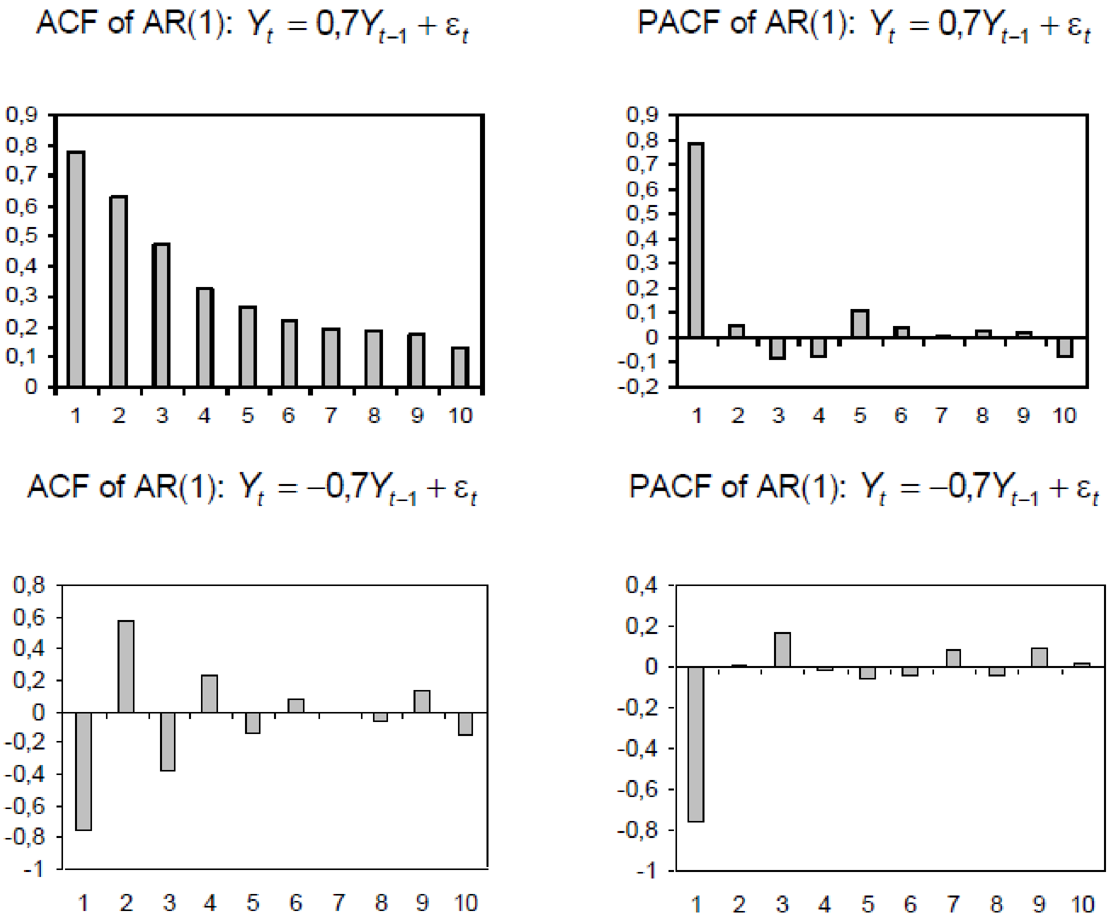


(a) $Y_t = 0{,}7Y_{t-1} + \varepsilon_t$

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

Simulated AR(1) models with $\phi = 0.7$ and $\phi = -0.7$

Autoregressive models



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

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

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

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

ACF and PACF of the simulated AR(1) models

Jorge Caiado

Autoregressive models

AR(2) model

The second-order autoregressive AR(2) models is

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

or

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

where ε_t is a zero mean white noise series. To be stationary, the roots of $\phi_2(B) = 1 - \phi_1 B - \phi_2 B^2 = 0$ must be outside of the unit circle. Thus, we have the following necessary and sufficient conditions for stationarity:

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

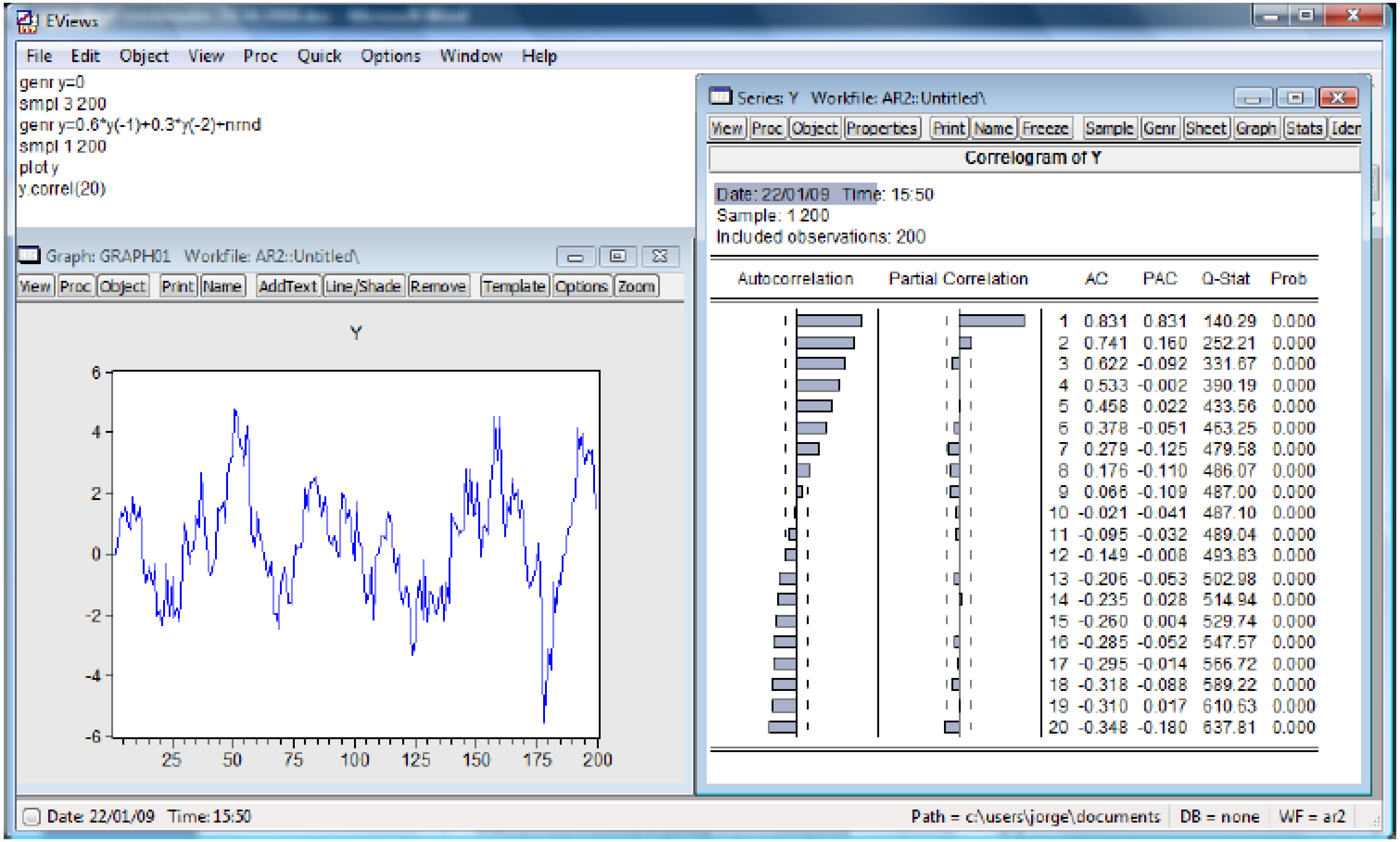
The ACF tails off as an exponential decay or a damped sine waves depending on the roots of $\phi(B) = 0$, and the PACF cuts off after lag 2, $\phi_{kk} = 0$ for $k \geq 3$.

AR(*p*) model

More complicated conditions hold for AR(*p*) models with $p \geq 3$.

Econometric software (EViews among others) takes care of this.

Autoregressive models

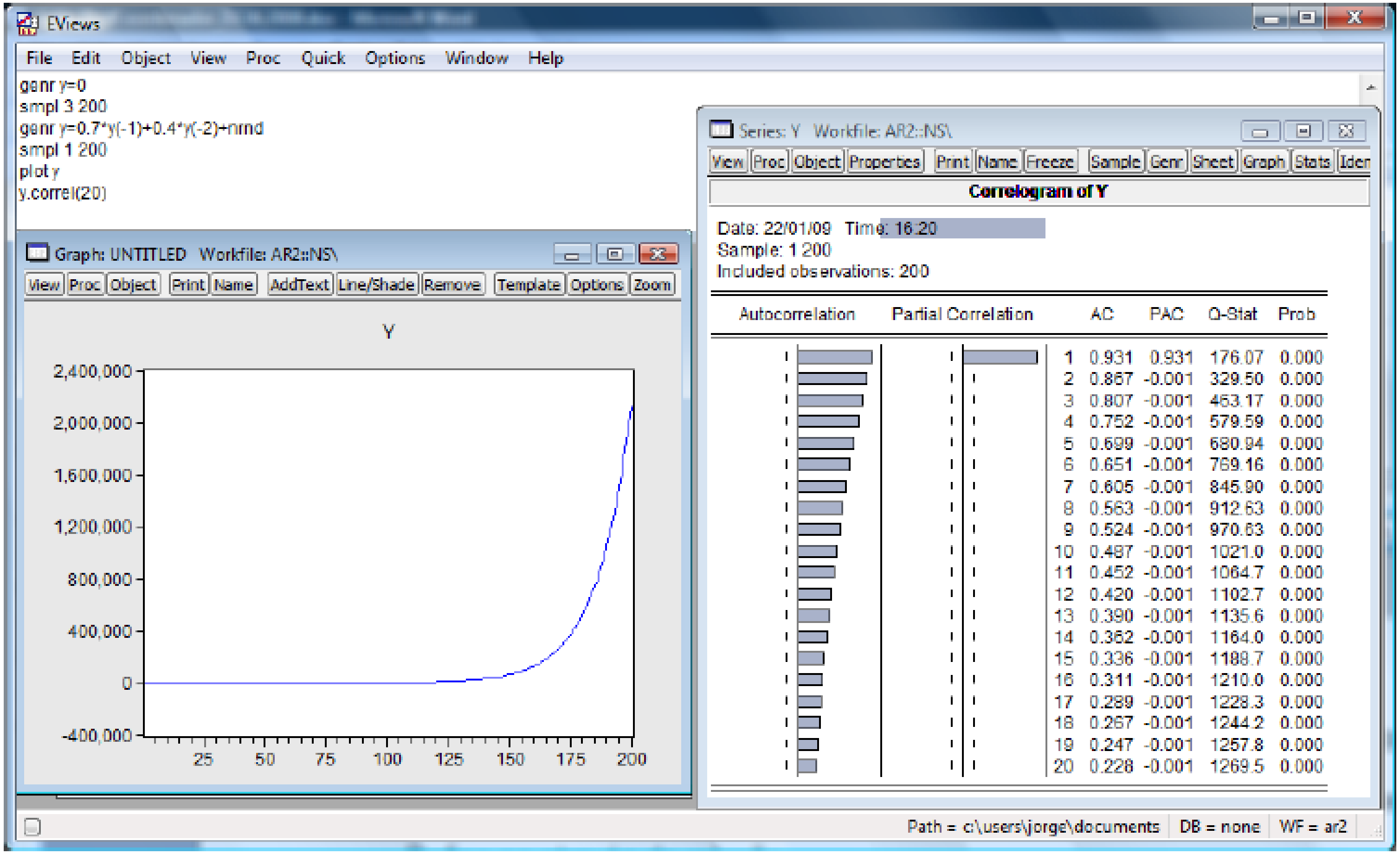


```
genr y=0
smpl 3 200
genr y=0.6*y(-1)+0.3*y(-2)+nrnd
smpl 1 200
plot y
y.correl(20)
```

```
Correlogram of Y

Date: 22/01/09   Time: 15:50
Sample: 1 200
Included observations: 200

      AC     PAC    Q-Stat  Prob
 1  0.831   0.831   140.29  0.000
 2  0.741   0.160   252.21  0.000
 3  0.622  -0.092   331.67  0.000
 4  0.533  -0.002   390.19  0.000
 5  0.458   0.022   433.56  0.000
 6  0.378  -0.051   463.25  0.000
 7  0.279  -0.125   479.58  0.000
 8  0.176  -0.110   486.07  0.000
 9  0.066  -0.109   487.00  0.000
10 -0.021  -0.041   487.10  0.000
11 -0.095  -0.032   489.04  0.000
12 -0.149  -0.008   493.83  0.000
13 -0.206  -0.053   502.98  0.000
14 -0.235   0.028   514.94  0.000
15 -0.260   0.004   529.74  0.000
16 -0.285  -0.052   547.57  0.000
17 -0.295  -0.014   566.72  0.000
18 -0.318  -0.088   589.22  0.000
19 -0.310   0.017   610.63  0.000
20 -0.348  -0.180   637.81  0.000
```

ACF and PACF of a simulated stationary AR(2) model: $Y_t = 0{,}6Y_{t-1} + 0{,}3Y_{t-2} + \varepsilon_t$

Autoregressive models



```
genr y=0
smpl 3 200
genr y=0.7*y(-1)+0.4*y(-2)+nrnd
smpl 1 200
plot y
y.correl(20)
```

```
                         Correlogram of Y

Date: 22/01/09   Time: 16:20
Sample: 1 200
Included observations: 200

   Autocorrelation    Partial Correlation      AC     PAC    Q-Stat   Prob

                                           1  0.931   0.931  176.07   0.000
                                           2  0.867  -0.001  329.50   0.000
                                           3  0.807  -0.001  463.17   0.000
                                           4  0.752  -0.001  579.59   0.000
                                           5  0.699  -0.001  680.94   0.000
                                           6  0.651  -0.001  769.16   0.000
                                           7  0.605  -0.001  845.90   0.000
                                           8  0.563  -0.001  912.63   0.000
                                           9  0.524  -0.001  970.63   0.000
                                          10  0.487  -0.001  1021.0   0.000
                                          11  0.452  -0.001  1064.7   0.000
                                          12  0.420  -0.001  1102.7   0.000
                                          13  0.390  -0.001  1135.6   0.000
                                          14  0.362  -0.001  1164.0   0.000
                                          15  0.336  -0.001  1188.7   0.000
                                          16  0.311  -0.001  1210.0   0.000
                                          17  0.289  -0.001  1228.3   0.000
                                          18  0.267  -0.001  1244.2   0.000
                                          19  0.247  -0.001  1257.8   0.000
                                          20  0.228  -0.001  1269.5   0.000
```

ACF and PACF of a simulated nonstationary AR(2) model: $Y_t = 0,7Y_{t-1} + 0,4Y_{t-2} + \varepsilon_t$

Moving average models

The finite-order representation of the moving average model described earlier, if only a finite number of ψ weights are nonzero, is given by

$$Y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q},$$

where ε_t is a zero mean white noise series. Because $1 + \theta_1^2 + \cdots + \theta_q^2 < \infty$, the process is always stationary. To be invertible, the roots of $(1 - \theta_1 B - \cdots - \theta_q B^q) = 0$ must be outside of the unit circle.

MA(1) model

The first-order moving average model or MA(1) model is

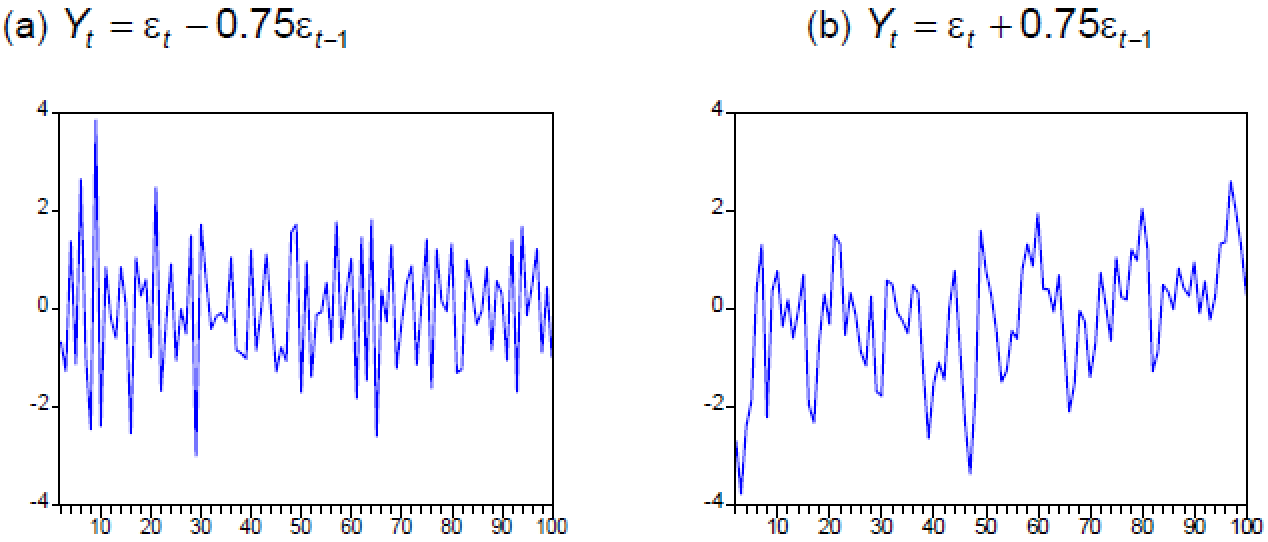
$$Y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1},$$

or

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

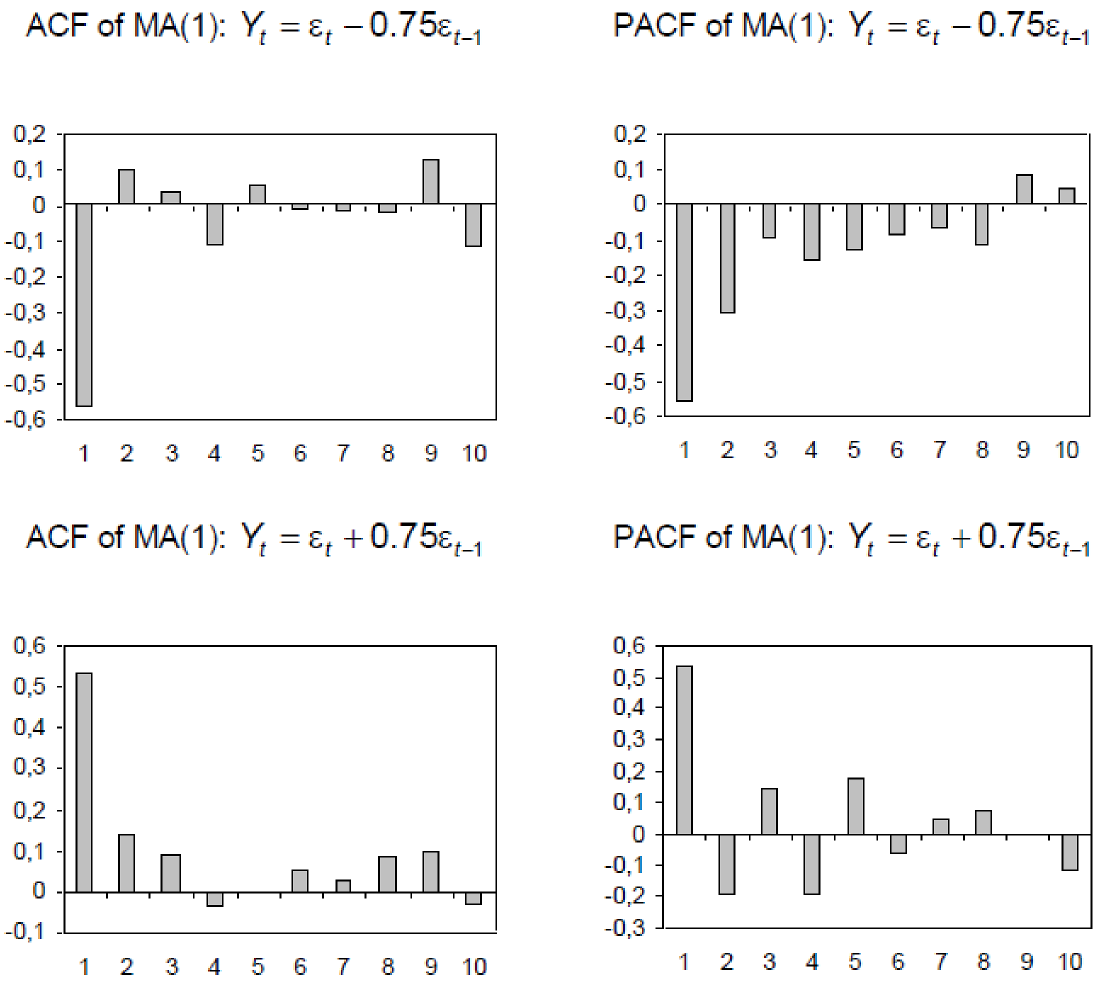
where $\theta(B) = 1 - \theta_1 B$ and ε_t is white noise. To be invertible, the root of $\theta(B) = 0$ must lie outside the unit circle. Thus, we require $|\theta_1| < 1$.

Moving average models



Simulated MA(1) models

Moving average models



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

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

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

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

ACF and PACF of simulated MA(1) models

Moving average models

LISBOA SCHOOL OF ECONOMICS & MANAGEMENT

MA(2) model

The second-order moving average MA(2) model is given by

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

or

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

where $\theta_2(B) = 1 - \theta_1 B - \theta_2 B^2$ and ε_t is white noise. To be invertible, the roots of $\theta_2(B) = 0$ must lie outside the unit circle. Hence, we have the following conditions:

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

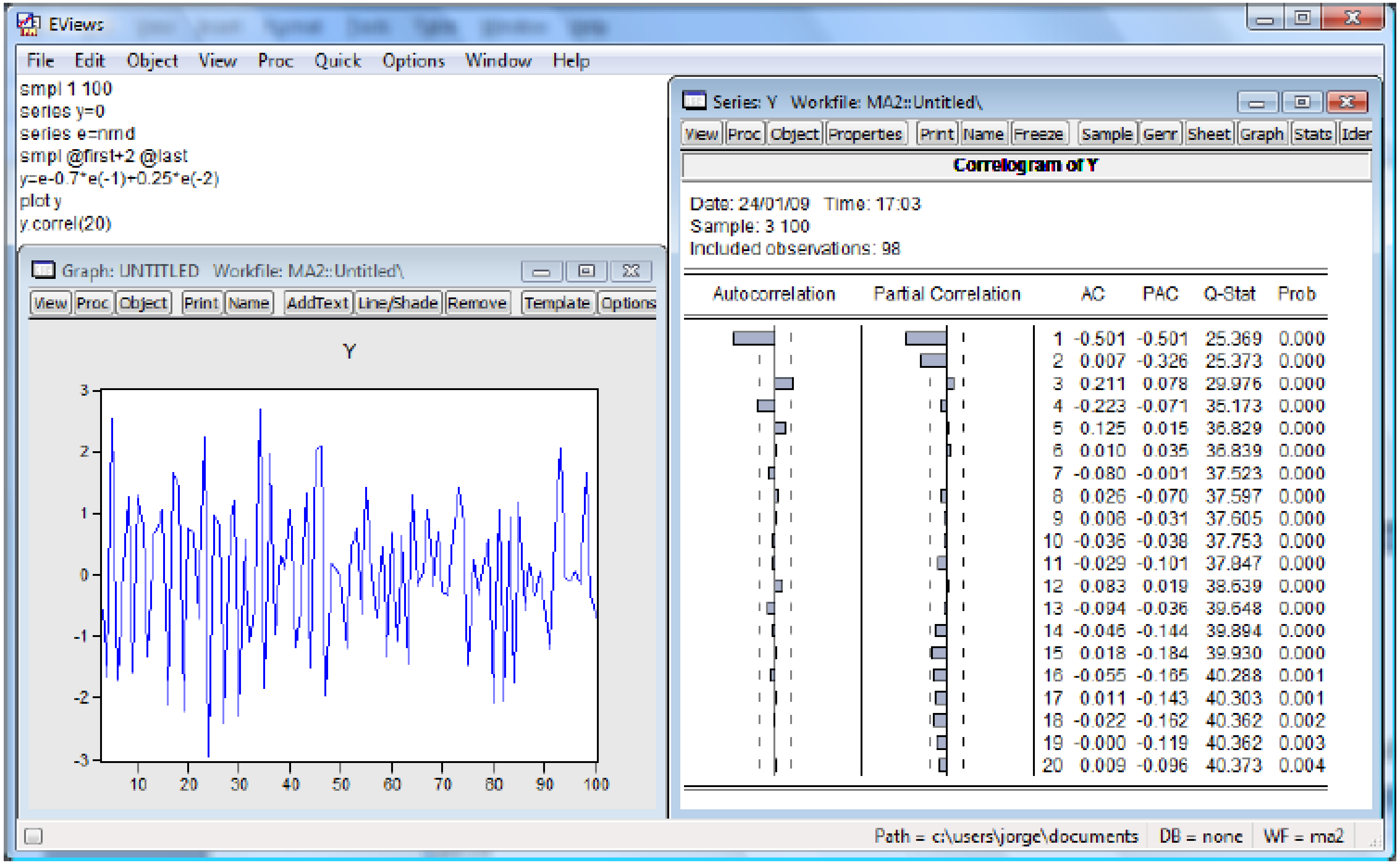
ACF of the MA(2) model cuts off after lag 2 and PACF tails off as an exponential decay or a damped sine wave depending on the roots of $\theta_2(B) = 0$.

MA(q) model

More complicated conditions hold for MA(q) models with $q \geq 3$.

Econometric software (EViews among others) takes care of this.

Moving average models

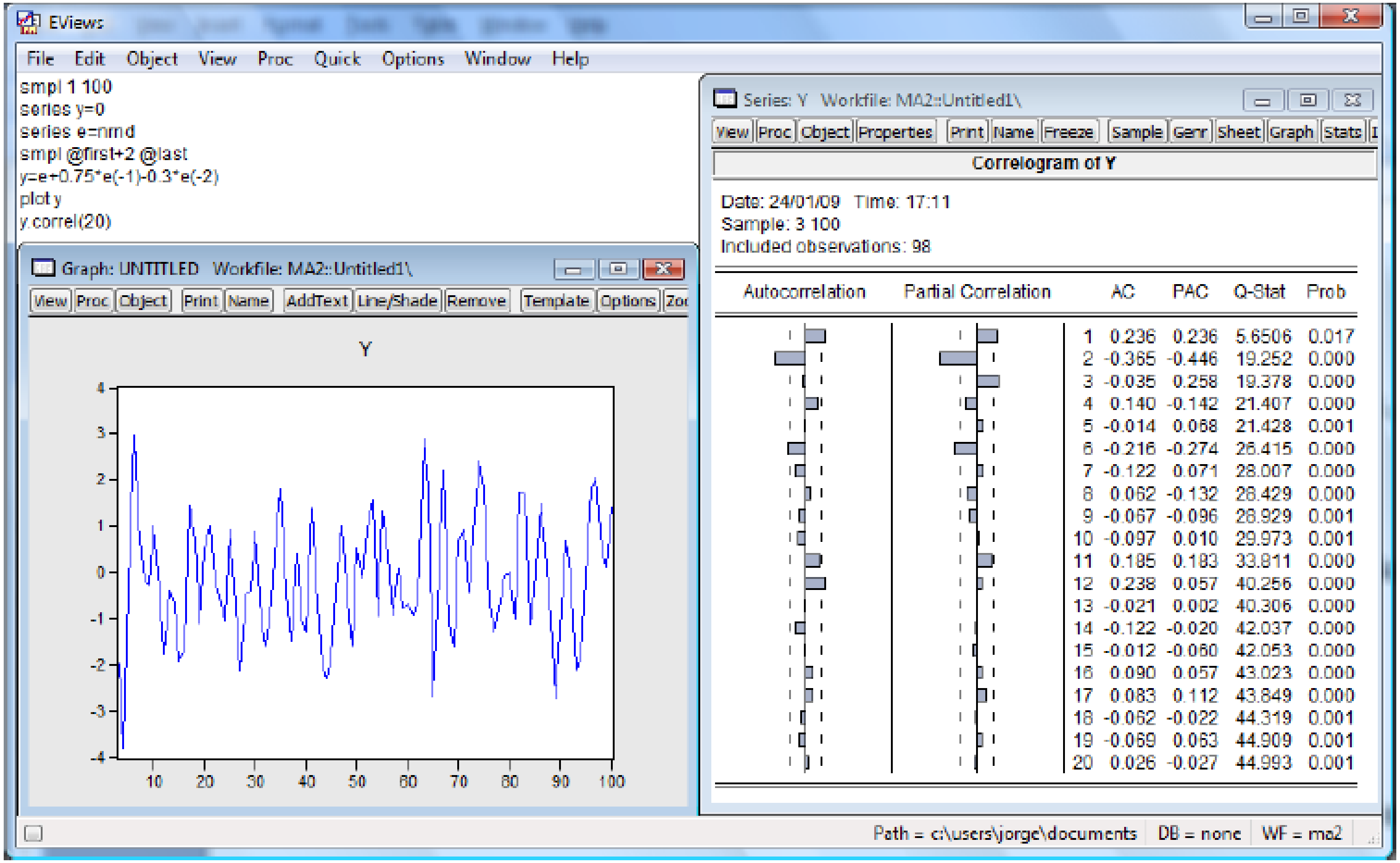


```
smpl 1 100
series y=0
series e=nrnd
smpl @first+2 @last
y=e-0.7*e(-1)+0.25*e(-2)
plot y
y.correl(20)
```

```
                         Correlogram of Y

Date: 24/01/09   Time: 17:03
Sample: 3 100
Included observations: 98

Autocorrelation   Partial Correlation      AC     PAC    Q-Stat   Prob

                                       1 -0.501 -0.501  25.369  0.000
                                       2  0.007 -0.326  25.373  0.000
                                       3  0.211  0.078  29.976  0.000
                                       4 -0.223 -0.071  35.173  0.000
                                       5  0.125  0.015  36.829  0.000
                                       6  0.010  0.035  36.839  0.000
                                       7 -0.080 -0.001  37.523  0.000
                                       8  0.026 -0.070  37.597  0.000
                                       9  0.008 -0.031  37.605  0.000
                                      10 -0.036 -0.038  37.753  0.000
                                      11 -0.029 -0.101  37.847  0.000
                                      12  0.083  0.019  38.639  0.000
                                      13 -0.094 -0.036  39.648  0.000
                                      14 -0.046 -0.144  39.894  0.000
                                      15  0.018 -0.184  39.930  0.000
                                      16 -0.055 -0.165  40.288  0.001
                                      17  0.011 -0.143  40.303  0.001
                                      18 -0.022 -0.162  40.362  0.002
                                      19 -0.000 -0.119  40.362  0.003
                                      20  0.009 -0.096  40.373  0.004
```

ACF and PACF of the simulated MA(2) model: $Y_t = \varepsilon_t - 0{,}7\varepsilon_{t-1} + 0{,}25\varepsilon_{t-2}$

Moving average models



```
smpl 1 100
series y=0
series e=nrnd
smpl @first+2 @last
y=e+0.75*e(-1)-0.3*e(-2)
plot y
y.correl(20)
```

```
                         Correlogram of Y

Date: 24/01/09   Time: 17:11
Sample: 3 100
Included observations: 98

Autocorrelation   Partial Correlation      AC     PAC    Q-Stat   Prob

                                       1  0.236   0.236  5.6506  0.017
                                       2 -0.365  -0.446  19.252  0.000
                                       3 -0.035   0.258  19.378  0.000
                                       4  0.140  -0.142  21.407  0.000
                                       5 -0.014   0.068  21.428  0.001
                                       6 -0.216  -0.274  26.415  0.000
                                       7 -0.122   0.071  28.007  0.000
                                       8  0.062  -0.132  28.429  0.000
                                       9 -0.067  -0.096  28.929  0.001
                                      10 -0.097   0.010  29.973  0.001
                                      11  0.185   0.183  33.811  0.000
                                      12  0.238   0.057  40.256  0.000
                                      13 -0.021   0.002  40.306  0.000
                                      14 -0.122  -0.020  42.037  0.000
                                      15 -0.012  -0.060  42.053  0.000
                                      16  0.090   0.057  43.023  0.000
                                      17  0.083   0.112  43.849  0.000
                                      18 -0.062  -0.022  44.319  0.001
                                      19 -0.069   0.063  44.909  0.001
                                      20  0.026  -0.027  44.993  0.001
```

ACF and PACF of the simulated MA(2) model: $Y_t = \varepsilon_t + 0{,}75\varepsilon_{t-1} - 0{,}2\varepsilon_{t-2}$

Autoregressive and moving average models

ARMA(1,1) model

The mixed autoregressive and moving average ARMA(1,1) model includes the autoregressive AR(1) and moving average MA(1) models as special cases.

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

or

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

where $\phi(B) = 1 - \phi B$, $\theta(B) = 1 - \theta B$ and ε_t is white noise. To be stationary, the root of $\phi(B) = 0$ must lie outside the unit circle, i.e., $-1 < \phi < 1$. To be invertible, the root of $\theta(B) = 0$ must lie outside the unit circle, i.e., $-1 < \theta < 1$.

The ARMA(1,1) model can be written in a pure moving average representation as

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

where

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

Autoregressive and moving average models

The ARMA(1,1) model can be written in a pure autoregressive representation as

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

where

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

Both the ACF and PACF of a mixed ARMA(1,1) model tail off as *k* incresaes, with its shape depending on the signs and magnitudes of ϕ and θ.

ARMA(p,q) model

The general mixed autoregressive and moving average ARMA(*p*,*q*) model is given by

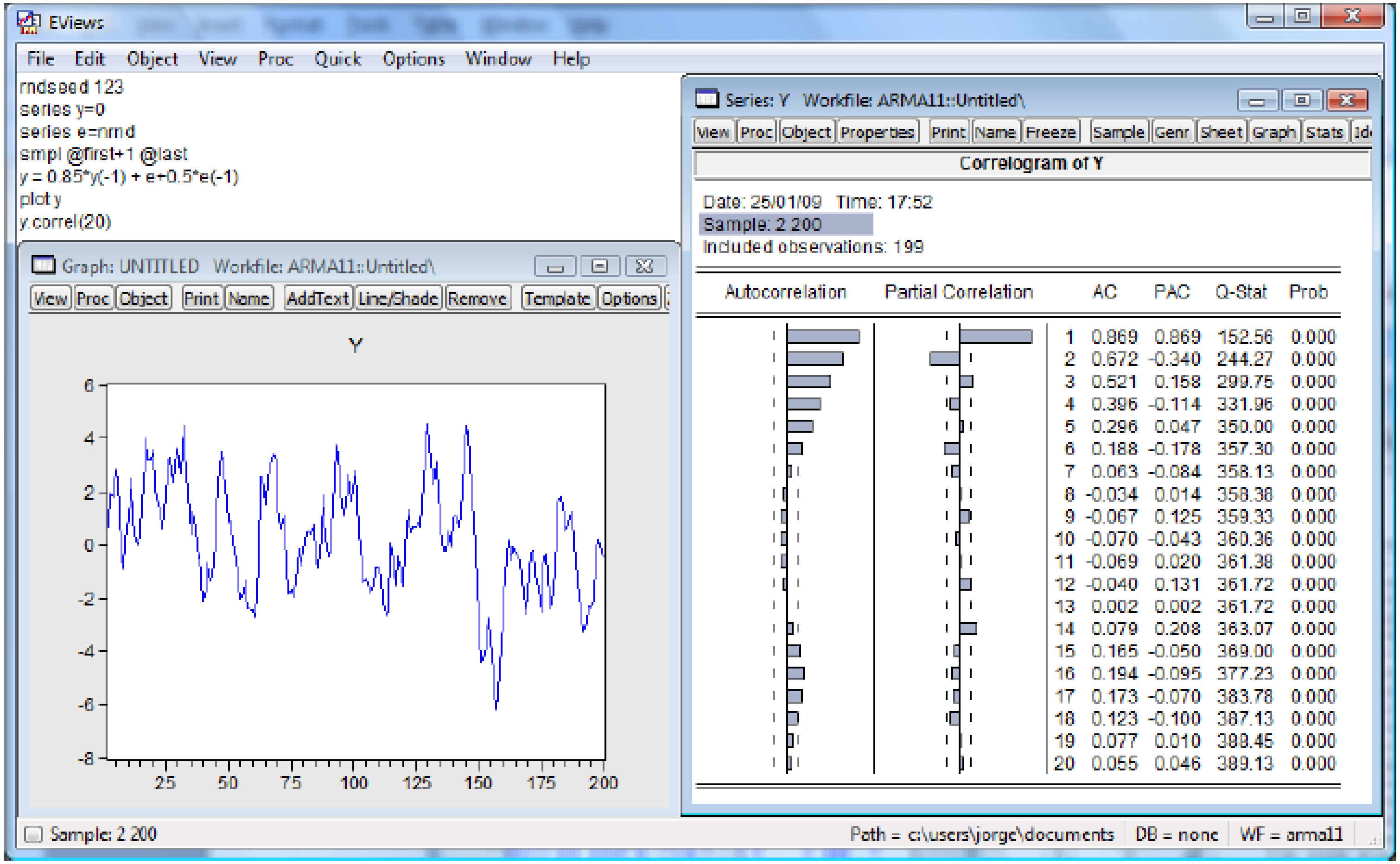
$$Y_t = \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q},$$

or

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

where $\phi_p(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$, $\theta_q(B) = 1 - \theta_1 B - \cdots - \theta_q B^q$ and ε_t is white noise. To be stationary, the roots of $\phi_p(B) = 0$ must lie outside the unit circle. To be invertible, the roots of $\theta_q(B) = 0$ must lie outside the unit circle.

Autoregressive and moving average models

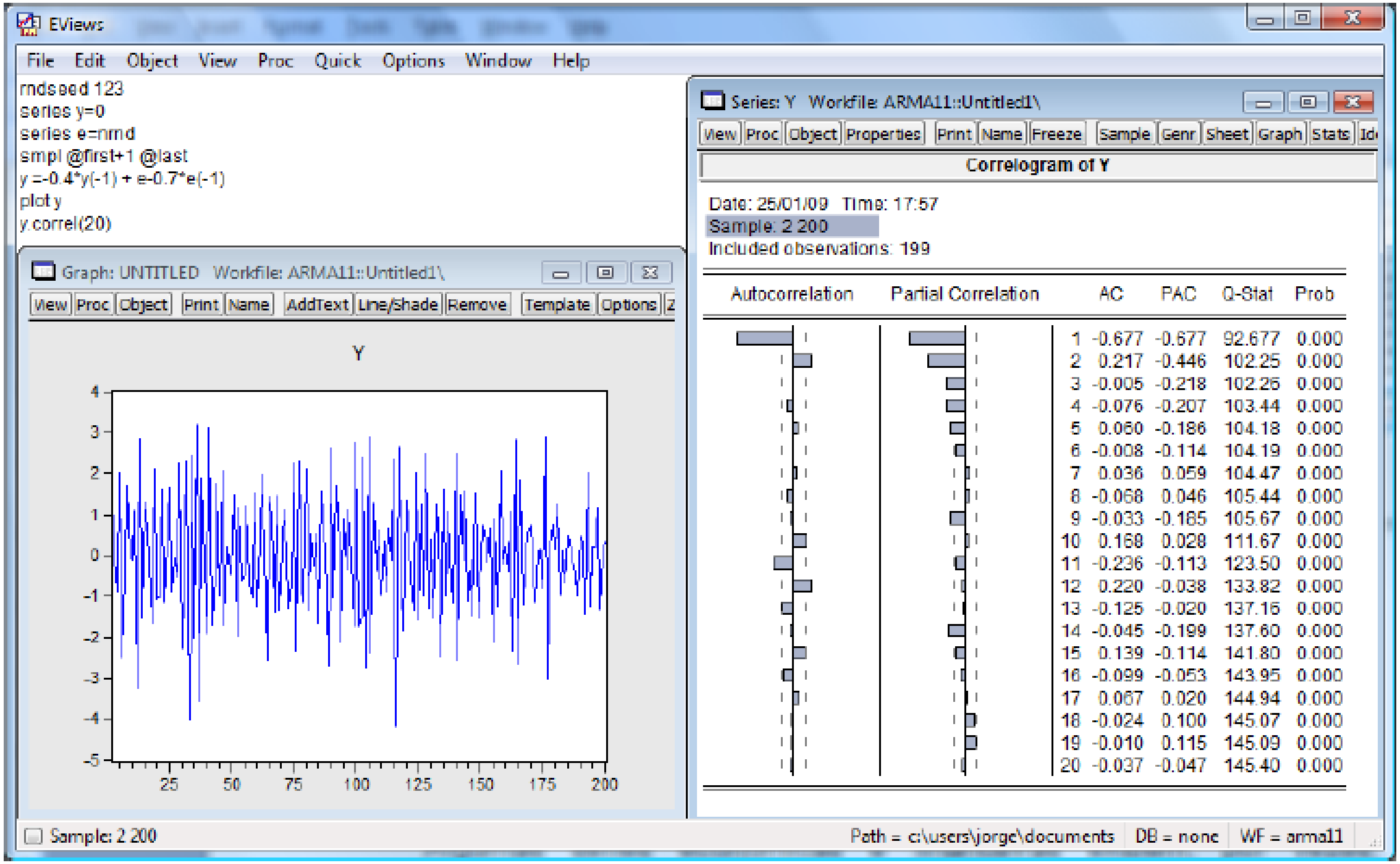


```
rndseed 123
series y=0
series e=nrnd
smpl @first+1 @last
y = 0.85*y(-1) + e+0.5*e(-1)
plot y
y.correl(20)
```

```
Correlogram of Y

Date: 25/01/09   Time: 17:52
Sample: 2 200
Included observations: 199

Autocorrelation   Partial Correlation      AC     PAC    Q-Stat   Prob

                                      1   0.869   0.869  152.56   0.000
                                      2   0.672  -0.340  244.27   0.000
                                      3   0.521   0.158  299.75   0.000
                                      4   0.396  -0.114  331.96   0.000
                                      5   0.296   0.047  350.00   0.000
                                      6   0.188  -0.178  357.30   0.000
                                      7   0.063  -0.084  358.13   0.000
                                      8  -0.034   0.014  358.38   0.000
                                      9  -0.067   0.125  359.33   0.000
                                     10  -0.070  -0.043  360.36   0.000
                                     11  -0.069   0.020  361.38   0.000
                                     12  -0.040   0.131  361.72   0.000
                                     13   0.002   0.002  361.72   0.000
                                     14   0.079   0.208  363.07   0.000
                                     15   0.165  -0.050  369.00   0.000
                                     16   0.194  -0.095  377.23   0.000
                                     17   0.173  -0.070  383.78   0.000
                                     18   0.123  -0.100  387.13   0.000
                                     19   0.077   0.010  388.45   0.000
                                     20   0.055   0.046  389.13   0.000
```

ACF and PACF of the ARMA(1,1) model: $Y_t = 0{,}85Y_{t-1} + \varepsilon_t + 0{,}5\varepsilon_{t-1}$

Autoregressive and moving average models



```
rndseed 123
series y=0
series e=nrnd
smpl @first+1 @last
y =-0.4*y(-1) + e-0.7*e(-1)
plot y
y.correl(20)
```

```
                     Correlogram of Y

Date: 25/01/09   Time: 17:57
Sample: 2 200
Included observations: 199

Autocorrelation  Partial Correlation     AC     PAC    Q-Stat   Prob

                                     1 -0.677 -0.677  92.677  0.000
                                     2  0.217 -0.446  102.25  0.000
                                     3 -0.005 -0.218  102.26  0.000
                                     4 -0.076 -0.207  103.44  0.000
                                     5  0.060 -0.186  104.18  0.000
                                     6 -0.008 -0.114  104.19  0.000
                                     7  0.036  0.059  104.47  0.000
                                     8 -0.068  0.046  105.44  0.000
                                     9 -0.033 -0.185  105.67  0.000
                                    10  0.168  0.028  111.67  0.000
                                    11 -0.236 -0.113  123.50  0.000
                                    12  0.220 -0.038  133.82  0.000
                                    13 -0.125 -0.020  137.16  0.000
                                    14 -0.045 -0.199  137.60  0.000
                                    15  0.139 -0.114  141.80  0.000
                                    16 -0.099 -0.053  143.95  0.000
                                    17  0.067  0.020  144.94  0.000
                                    18 -0.024  0.100  145.07  0.000
                                    19 -0.010  0.115  145.09  0.000
                                    20 -0.037 -0.047  145.40  0.000
```

```
Sample: 2 200          Path = c:\users\jorge\documents   DB = none   WF = arma11
```

ACF and PACF of the ARMA(1,1) model: $Y_t = -0{,}4Y_{t-1} + \varepsilon_t - 0{,}7\varepsilon_{t-1}$

Seasonal ARMA models

Seasonal autoregressive and moving average SARMA(P,Q)$_s$ model

The seasonal SARMA(P,Q)$_s$ model is represented by

$$Y_t = \Phi_1 Y_{t-s} + \cdots + \Phi_P Y_{t-Ps} + \varepsilon_t - \Theta_1 \varepsilon_{t-s} - \cdots - \Theta_Q \varepsilon_{t-Qs} ,$$

or

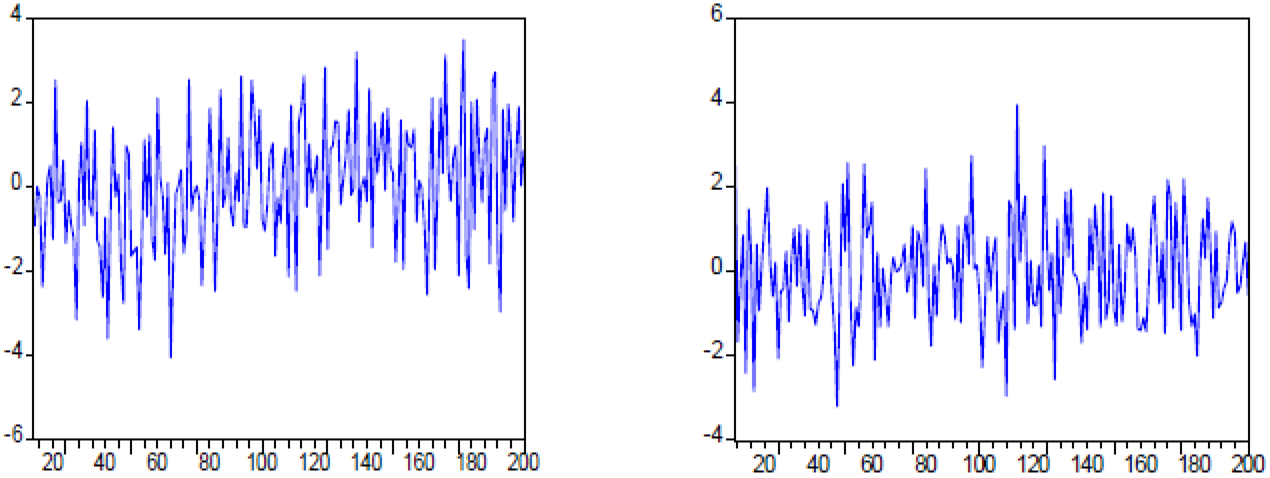
$$\Phi_P(B^s) Y_t = \Theta_Q(B^s) \varepsilon_t ,$$

where $\Phi_P(B^s) = 1 - \Phi_1 B^s - \cdots - \Phi_P B^{Ps}$, $\Theta_Q(B^s) = 1 - \Theta_1 B^s - \cdots - \Theta_Q B^{Qs}$ and ε_t is a zero mean white noise. To be stationary and invertible, the roots of $\Phi_P(B^s) = 0$ e $\Theta_Q(B^s) = 0$ must lie outside of the unit circle, respectively.

Both the ACF and PACF of the SARMA(P,Q)$_s$ model exhibit exponential decays and damped sine waves at the seasonal lags.

Seasonal ARMA models

(i) $(1-0.65B^{12})Y_t = (1+0.25B^{12})\varepsilon_t$ (ii) $(1-0.3B^4)Y_t = (1-0.4B^4+0.15B^8)\varepsilon_t$

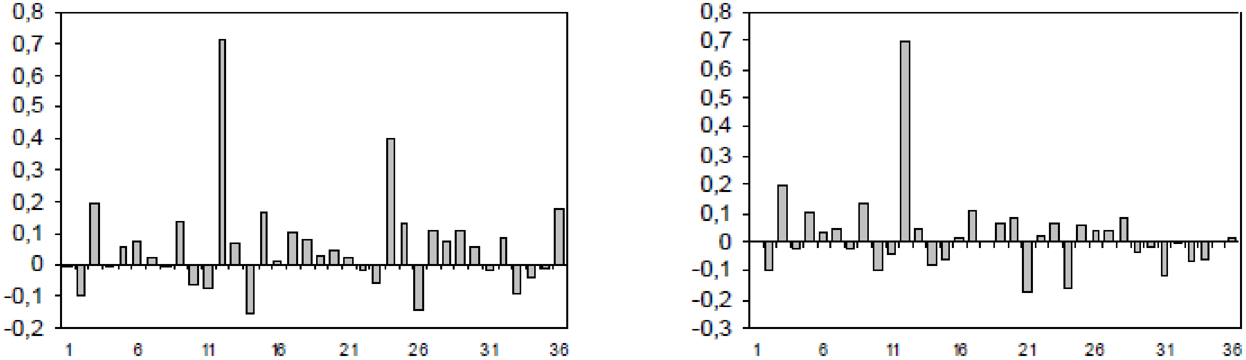


Simulated SARMA(1,1)$_{12}$ and SARMA(1,2)$_4$ models

Seasonal ARMA models

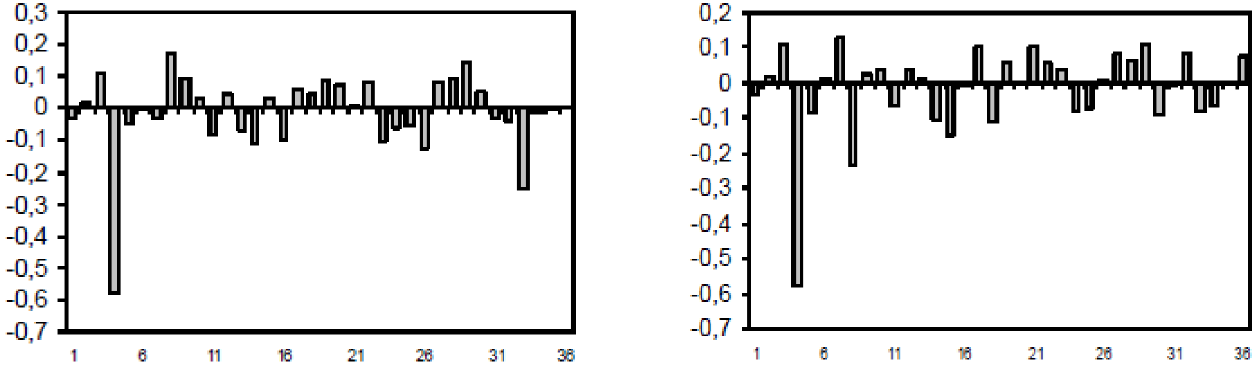
ACF of SARMA(1,1)$_{12}$ PACF of SARMA(1,1)$_{12}$

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



ACF of SARMA(1,2)$_4$ PACF of SARMA(1,2)$_4$

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



ACF and PACF of simulated SARMA(1,1)$_{12}$ and SARMA(1,2)$_4$ models

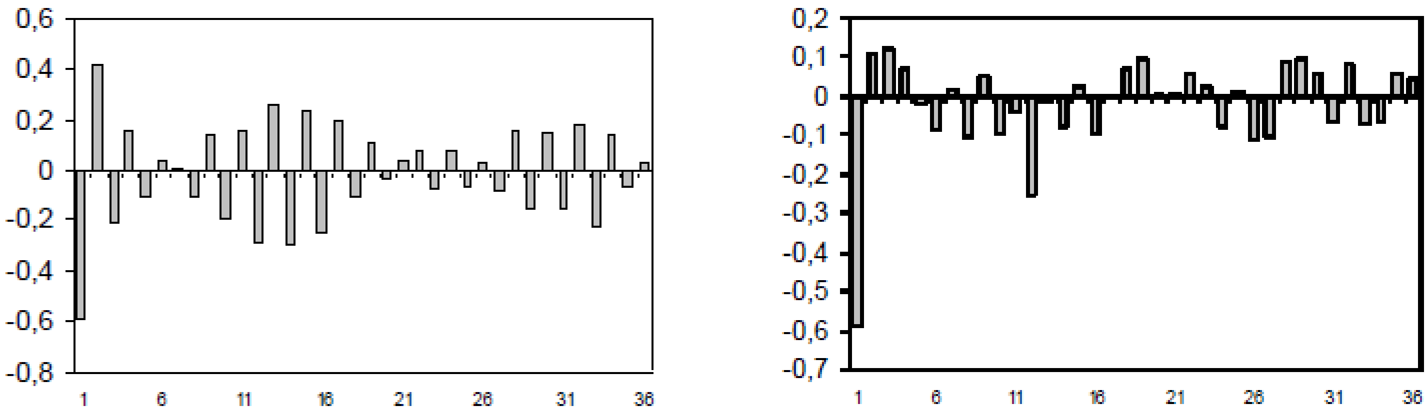
General multiplicative ARMA models

If we combine non-seasonal ARMA(p,q) and seasonal SARMA(P,Q)$_s$ models, we obtain a general multiplicative model of order $(p,q)\times(P,Q)_s$

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

or

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



ACF and PACF of a simulated SARMA(1,0)(1,0)$_{12}$ model: $(1-0.7B)(1+0.25B^{12})Y_t=\varepsilon_t$

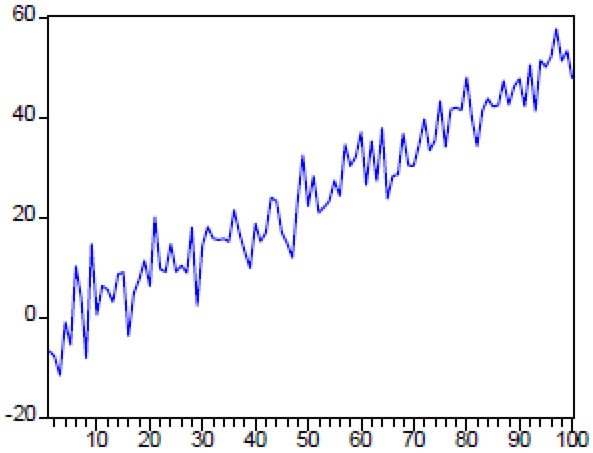
Linear nonstationary time series models

Nonstationary model in the mean

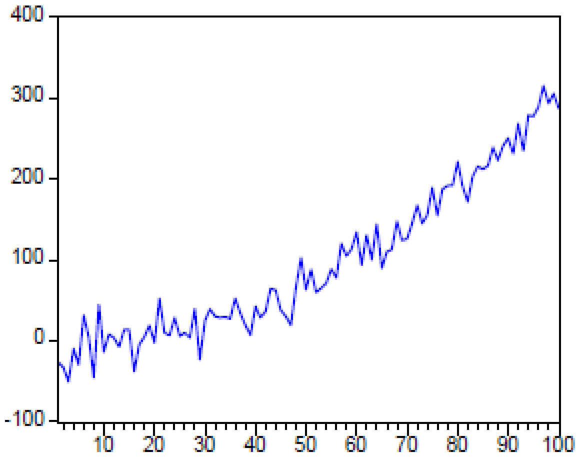
The mean function of a nonstationary model can be represented essentially by two models: deterministic trend models and stochastic trend models.

For a deterministic trend model, one can use the linear trend model, $Y_t = a + bt + \varepsilon_t$ or the quadratic trend model, $Y_t = a + bt + ct^2 + \varepsilon_t$.

Linear trend model



Quadratic trend model



Linear nonstationary time series models

Differencing and stochastic trend model

The *d*th differenced series, for some integer $d \geq 1$, is given by

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

For $d = 1$, we have first differences

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

For seasonal time series, we can use a *s*th seasonal differencing

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

Finally, a *s*th seasonal differencing of order *D*, for some integer $D \geq 1$ is given by

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

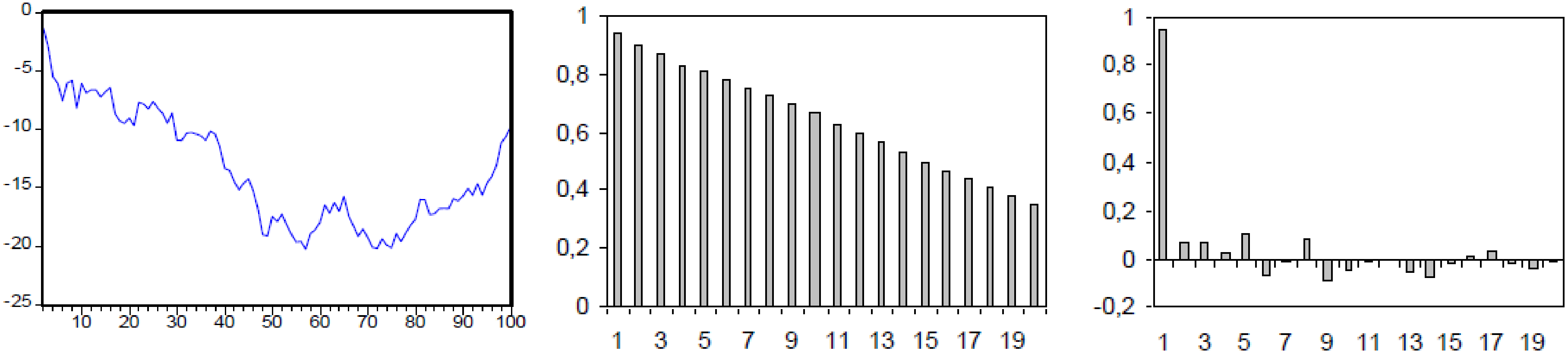
Usually *D* = 1,2 is sufficient to obtain seasonal stationarity.

Linear nonstationary time series models

A special case of the nonstationary models is the **stochastic trend model**,

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

where ε_t is white noise. This is the so-called "*random walk*" model.



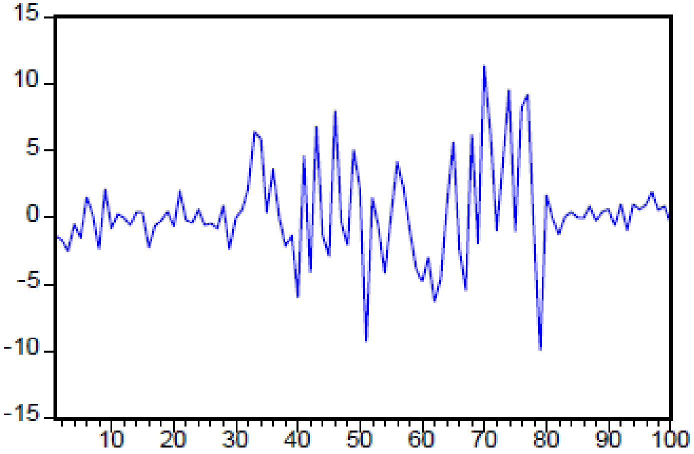
ACF and PACF of a simulated random walk model

Linear nonstationary time series models

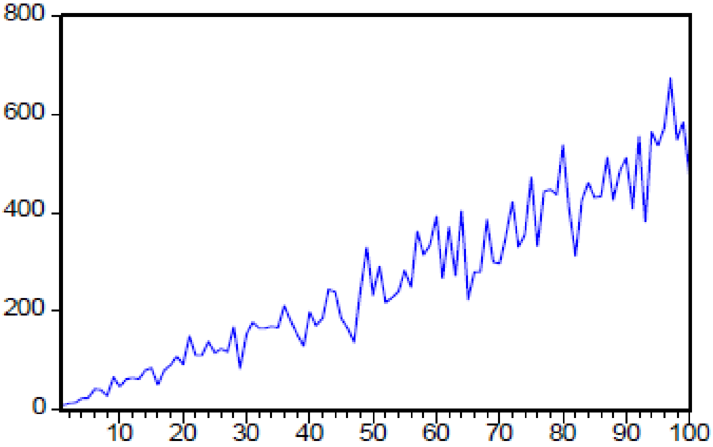
Nonstationarity in the variance

Many time series are stationary in the mean but are nonstationary in the variance. To reduce this type of nonstationarity, we need variance stabilizing transformations such as the power transformation of Box-Cox (1964),

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



A simulated time series nonstationary in the variance but stationary in the mean



A simulated time series nonstationarity in both the mean and variance

Linear nonstationary time series models

In practice, we fit the model to $Y_t^{(\lambda)} = \frac{Y_t^{\lambda} - 1}{\lambda \tilde{Y}^{\lambda-1}}$, for various values of $\lambda \neq 0$, where $\tilde{Y}$ is the geometric mean of the series Y_t, and choose the value of λ that results in the smallest residual sum of squares. For $\lambda = 0$, we have $Y_t^{(0)} = \tilde{Y} \log(Y_t)$.

Autoregressive integrated moving average (ARIMA) models

A general model for representing nonstationary nonseasonal time series is given by the autoregressive integrated moving average ARIMA(p,d,q) model

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

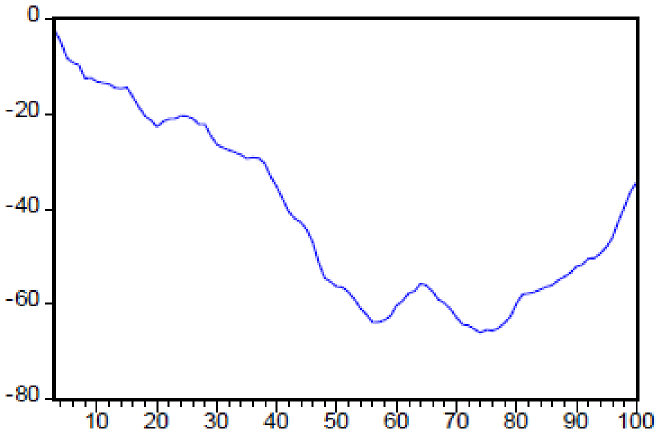
or

$$\phi_p(B)(1 - B)^d Y_t = \theta_q(B)\varepsilon_t \, ,$$

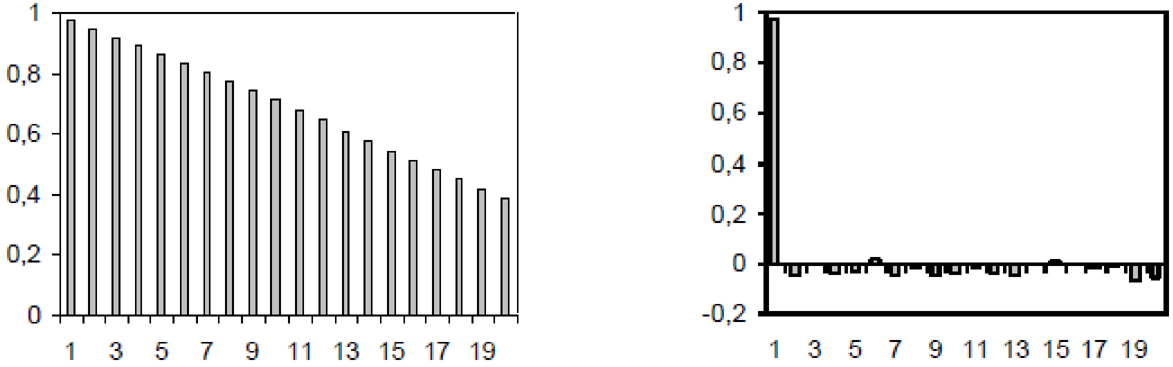
where $(1 - B)^d$ is the differencing operator of order d, for $d \geq 1$, $\phi_p(B)$ is a stationary autoregressive (AR) operator, $\theta_q(B)$ is an invertible moving average (MA) operator and ε_t is a zero mean white noise.

Some important special cases of the ARIMA model are ARIMA(0,1,0), ARIMA(1,1,0), ARIMA(0,1,1) and ARIMA (1,1,1) models.

Linear nonstationary time series models



A simulated series from ARIMA(1,1,0) model: $(1-0{,}75B)(1-B)Y_t=\varepsilon_t$



ACF and PACF of the ARIMA(1,1,0) model: $(1-0{,}75B)(1-B)Y_t=\varepsilon_t$

Linear nonstationary time series models

Multiplicative autoregressive integrated moving average models

The multiplicative seasonal ARIMA model is an extension of the nonseasonal ARIMA model, by adding seasonal autoregressive and moving average factors. The model, often denoted as SARIMA$(p,d,q)(P,D,Q)_S$, is represented by

$$\phi_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D Y_t = \theta_q(B)\Theta_Q(B^s)\varepsilon_t\,,$$

where $\phi_p(B)$ and $\theta_q(B)$ are regular (nonseasonal) autoregressive and moving average factors, respectively, $\Phi_P(B^s)$ and $\Theta_Q(B^s)$ are seasonal autoregressive and moving average factors, respectively, and s is the seasonal period.

For example, consider the SARIMA$(0,1,1)(0,1,1)_{12}$ model

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

Model identification

Steps for model identification

- Plot the time series and examine whether the series contains a trend, seasonality, ouliers, nonconstant variances and other nonstationary phenomena. Choose proper variance-stabilizing (Box-Cox's power transformation) and differencing transformations.
- Compute the sample ACF and the sample PACF of the original series and identify the degree of differencing d and D necessary to achieve stationarity. In practice, d and D are either 0, 1, or 2.
- Compute the sample ACF and the sample PACF of the transformed and differenced and identify the orders p and q for the regular autoregressive and moving average operators and the orders P and Q for the seasonal autoregressive and moving average operators, respectively. Usually, the needed orders of integers p, q, P and Q are less or equal to 3.

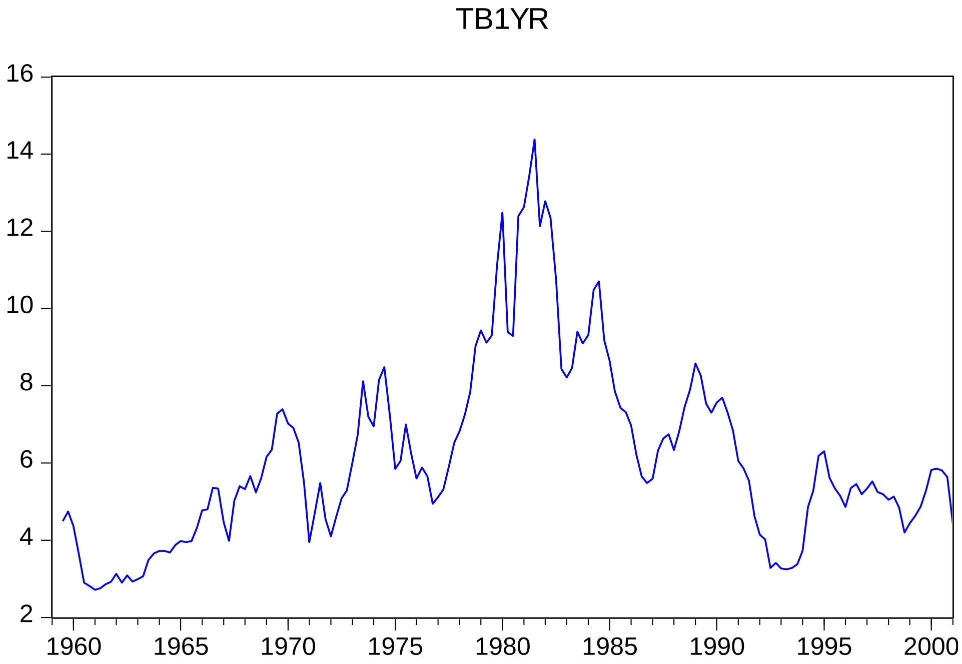
Model identification

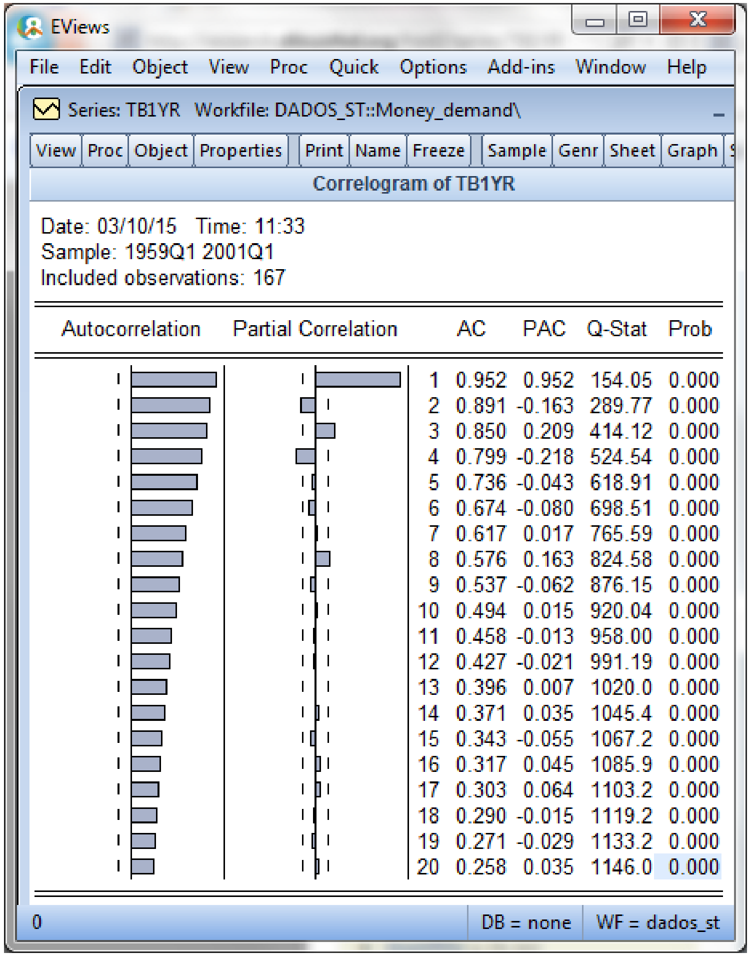
Theoretical ACF and PACF patterns for ARMA models

Model	**ACF**	**PACF**
AR(p)	Tails off as exponential decay or damped sine wave	Cuts off after lag p
MA(q)	Cuts off after lag q	Tails off as exponential decay or damped sine wave
ARMA(p,q)	Tails off after lag (q-p)	Tails off after lag (q-p)
SAR(P)	Tails off as exponential decay or damped sine wave at the seasonal lags s, 2s, ...	Cuts off after lag $P \times s$
SMA(Q)	Cuts off after lag $Q \times s$	Tails off as exponential decay or damped sine wave at the seasonal lags s, 2s, ...
SARMA(P,Q)	Tails off as exponential decay or damped sine wave at the seasonal lags s, 2s, ...	Tails off as exponential decay or damped sine wave at the seasonal lags s, 2s, ...
SARMA(p,q)(P,Q)$_s$	Tails off as exponential decay or damped sine wave at the seasonal and nonseasonal lags	Tails off as exponential decay or damped sine wave at the seasonal and nonseasonal lags

Model identification

Example: 1-Year US Treasury Bill: Secondary Market Rate

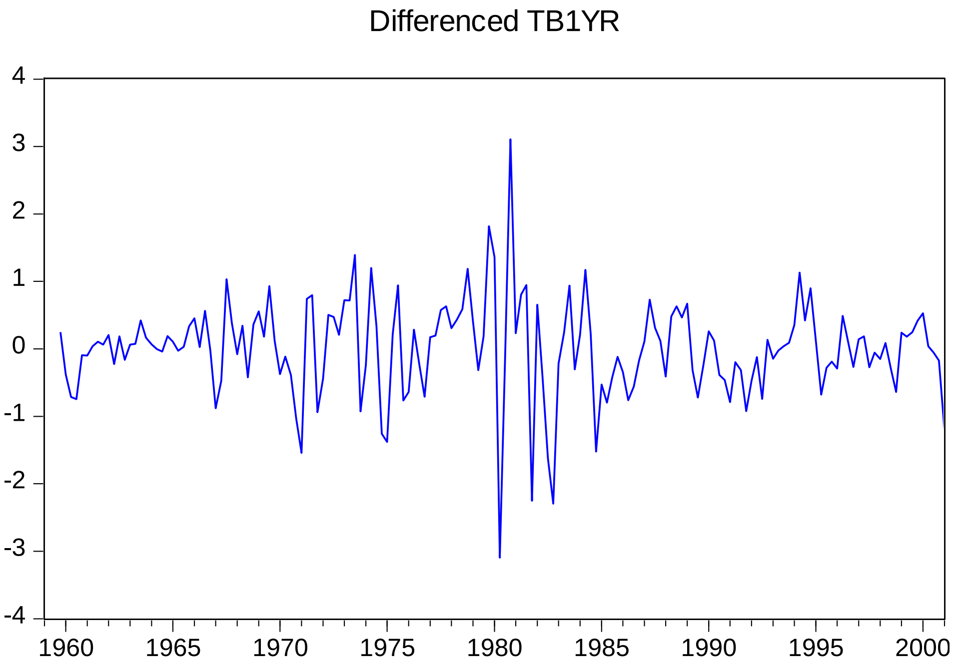


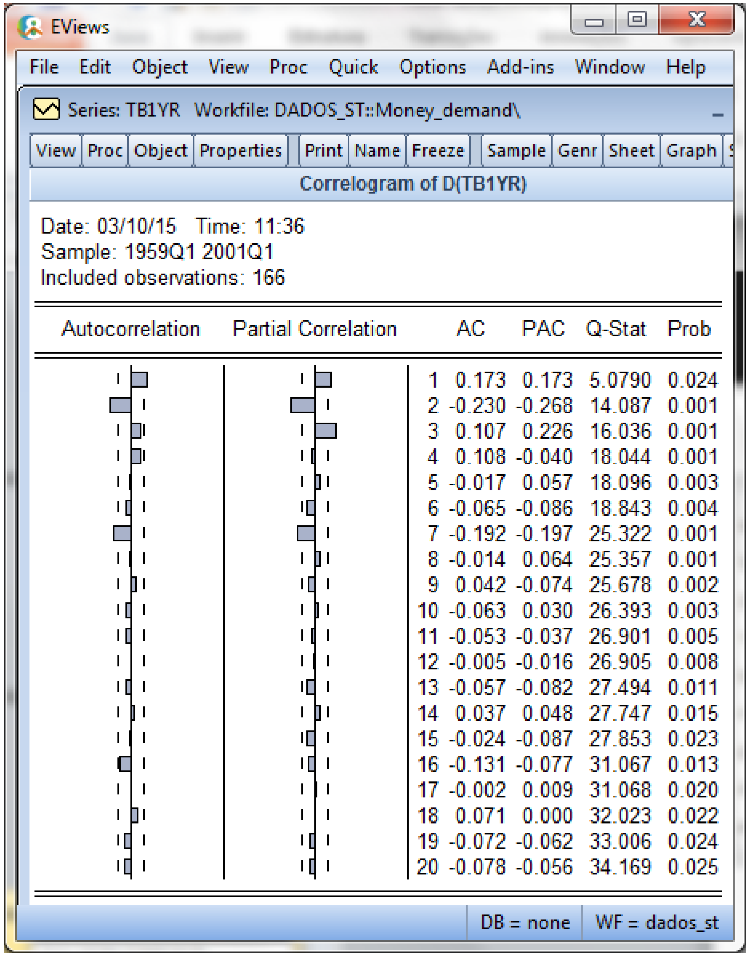


```
Date: 03/10/15   Time: 11:33
Sample: 1959Q1 2001Q1
Included observations: 167

Autocorrelation   Partial Correlation      AC     PAC    Q-Stat  Prob

                                       1  0.952   0.952  154.05  0.000
                                       2  0.891  -0.163  289.77  0.000
                                       3  0.850   0.209  414.12  0.000
                                       4  0.799  -0.218  524.54  0.000
                                       5  0.736  -0.043  618.91  0.000
                                       6  0.674  -0.080  698.51  0.000
                                       7  0.617   0.017  765.59  0.000
                                       8  0.576   0.163  824.58  0.000
                                       9  0.537  -0.062  876.15  0.000
                                      10  0.494   0.015  920.04  0.000
                                      11  0.458  -0.013  958.00  0.000
                                      12  0.427  -0.021  991.19  0.000
                                      13  0.396   0.007  1020.0  0.000
                                      14  0.371   0.035  1045.4  0.000
                                      15  0.343  -0.055  1067.2  0.000
                                      16  0.317   0.045  1085.9  0.000
                                      17  0.303   0.064  1103.2  0.000
                                      18  0.290  -0.015  1119.2  0.000
                                      19  0.271  -0.029  1133.2  0.000
                                      20  0.258   0.035  1146.0  0.000
```

Model identification

Example: 1-Year US Treasury Bill: Secondary Market Rate





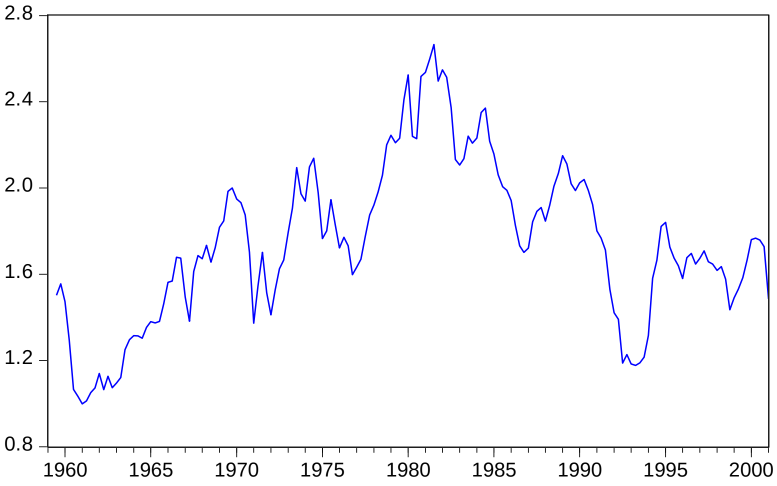
```
Correlogram of D(TB1YR)

Date: 03/10/15   Time: 11:36
Sample: 1959Q1 2001Q1
Included observations: 166

Autocorrelation   Partial Correlation      AC     PAC    Q-Stat   Prob

                                       1   0.173   0.173  5.0790  0.024
                                       2  -0.230  -0.268  14.087  0.001
                                       3   0.107   0.226  16.036  0.001
                                       4   0.108  -0.040  18.044  0.001
                                       5  -0.017   0.057  18.096  0.003
                                       6  -0.065  -0.086  18.843  0.004
                                       7  -0.192  -0.197  25.322  0.001
                                       8  -0.014   0.064  25.357  0.001
                                       9   0.042  -0.074  25.678  0.002
                                      10  -0.063   0.030  26.393  0.003
                                      11  -0.053  -0.037  26.901  0.005
                                      12  -0.005  -0.016  26.905  0.008
                                      13  -0.057  -0.082  27.494  0.011
                                      14   0.037   0.048  27.747  0.015
                                      15  -0.024  -0.087  27.853  0.023
                                      16  -0.131  -0.077  31.067  0.013
                                      17  -0.002   0.009  31.068  0.020
                                      18   0.071   0.000  32.023  0.022
                                      19  -0.072  -0.062  33.006  0.024
                                      20  -0.078  -0.056  34.169  0.025
```

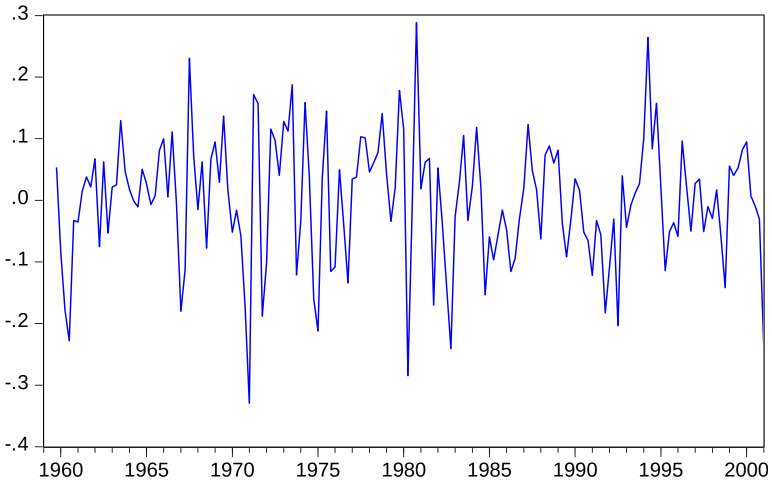
Model identification

Example: 1-Year US Treasury Bill - ARIMA(0,1,2) model

Box-Cox's power transformation on the 1-year Treasury Bill data

λ	Residual sum of squares
1	74.49
0.5	56.24
0	50.04
-0.5	52.04
-1	61.10


Log TB1YR


Log Differenced TB1YR

Model estimation

After identifying a tentative model, we need to estimate the parameters of the model.

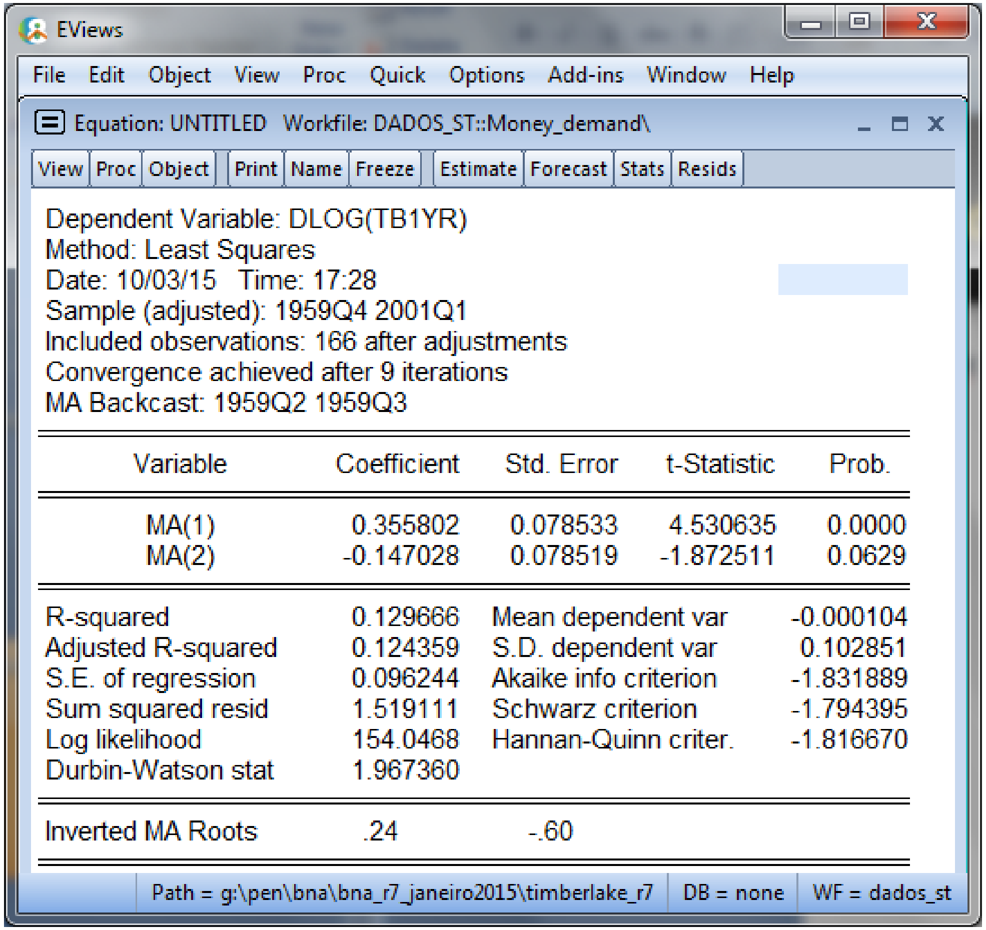
We discuss two widely used estimation procedures:

- **Maximum likelihood estimators (MLE) method**
  The parameter values of the ARIMA model are obtained by minimizing the conditional log-likelihood function
  $$\ln L_*(\phi,\theta,\sigma_\varepsilon) = -\frac{n}{2}\ln 2\pi\sigma_\varepsilon^2 - \frac{S_*(\phi,\theta)}{2\sigma_\varepsilon^2}$$
  where $S_*(\phi,\theta) = \sum_{t=p+1}^{n} \sigma_\varepsilon^2(\phi,\theta \mid \mathbf{Y})$ is the conditional sum of squares function.
- **Ordinary Least Squares (OLS) method**
  OLS is the most commonly used regression method in data analysis.
  However, for ARMA(p,q) models, the OLS estimator will be inconsistent unless we have q=0. For more details, see Wei (2006).

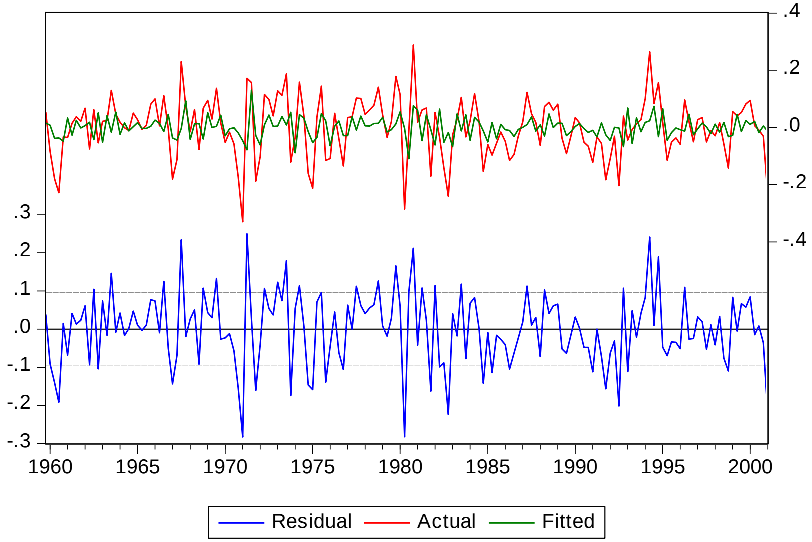
Different software will give different estimates. We use the EViews software.

Model estimation

Example: 1-Year US Treasury Bill - ARIMA(0,1,2) model



```
Dependent Variable: DLOG(TB1YR)
Method: Least Squares
Date: 10/03/15   Time: 17:28
Sample (adjusted): 1959Q4 2001Q1
Included observations: 166 after adjustments
Convergence achieved after 9 iterations
MA Backcast: 1959Q2 1959Q3

    Variable        Coefficient    Std. Error    t-Statistic    Prob.

     MA(1)           0.355802      0.078533      4.530635      0.0000
     MA(2)          -0.147028      0.078519     -1.872511      0.0629

R-squared            0.129666   Mean dependent var     -0.000104
Adjusted R-squared   0.124359   S.D. dependent var      0.102851
S.E. of regression   0.096244   Akaike info criterion  -1.831889
Sum squared resid    1.519111   Schwarz criterion      -1.794395
Log likelihood       154.0468   Hannan-Quinn criter.   -1.816670
Durbin-Watson stat   1.967360

Inverted MA Roots        .24             -.60
```



Coefficient covariance matrix

	MA(1)	MA(2)
MA(1)	0.006167	0.002640
MA(2)	0.002640	0.006165

Model diagnostic checking

Check on whether a particular model is adequate or not. This involves:

- **Analysis of the quality of parameter estimates.** Inspecting the statistical significance of individual parameter estimates provides some insight into the potential relative goodness of fit of the ARIMA model. To test the null hypothesis $H_0 : \beta_i = 0$, we use the test statistic:

$$|t| = \left| \frac{\hat{\beta}_i}{\sigma_{\hat{\beta}_i}} \right| > t_{(n-m)} \Rightarrow \text{Reject } H_0 : \beta_i = 0 .$$

- **Check whether the residuals are approximately white noise.** Compute the sample ACF and sample PACF of the residuals to check whether they are uncorrelated. Box and Pierce (1970) introduced a 'portmanteu' test to check the null hypothesis $H_0 : \rho_1 = \rho_2 = \cdots = \rho_k = 0$, with the test statistic

$$Q = n \sum_{j=1}^{k} \hat{\rho}_j^2 ,$$

which is asymptotically distributed as χ^2 with $k - m$ degrees of freedom, with m the number of estimated parameters.

Model diagnostic checking

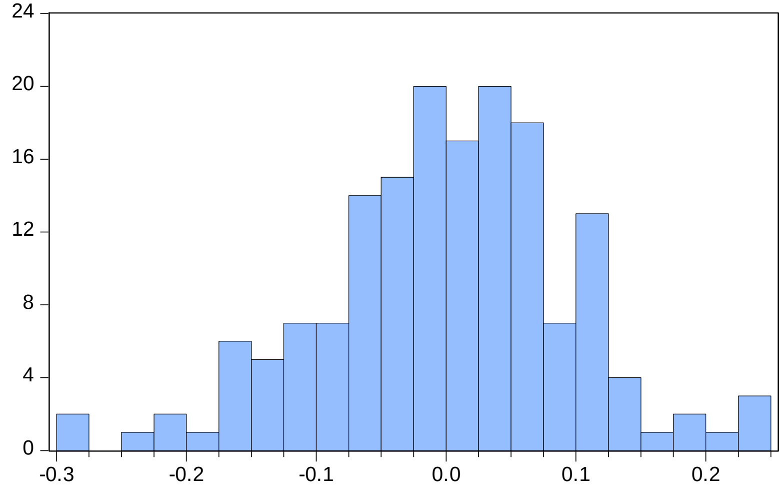
Ljung e Box (1978) proposed a modified version of the statistic Q,

$$Q^{*} = n(n+2)\sum_{j=1}^{k}\frac{\hat{\rho}_j^2}{n-j}.$$

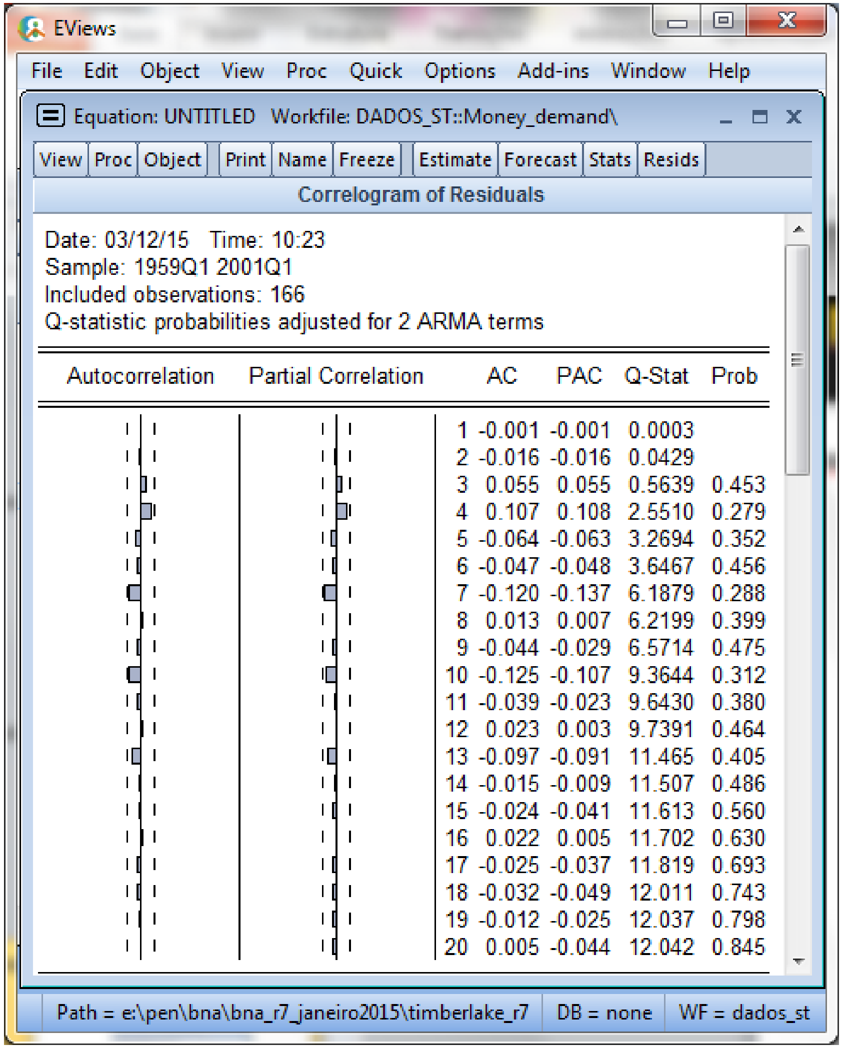
This modified form of the 'portmanteu' test statistic is much closer to the $\chi^2(k-m)$ distribution for typical sample sizes n. Thus, if the calculated Q^{*} statistic exceeds the value $\chi^2(k-m)$ then the adequacy of the fitted ARMA model would be questionated.

Model diagnostic checking

Example: 1-Year US Treasury Bill ARIMA(0,1,2) model



```
Series: Residuals
Sample 1959Q4 2001Q1
Observations 166

Mean         -0.000345
Median        0.004791
Maximum       0.249859
Minimum      -0.282943
Std. Dev.     0.095951
Skewness     -0.229303
Kurtosis      3.464358

Jarque-Bera   2.946140
Probability   0.229221
```



```
Correlogram of Residuals

Date: 03/12/15   Time: 10:23
Sample: 1959Q1 2001Q1
Included observations: 166
Q-statistic probabilities adjusted for 2 ARMA terms

 Autocorrelation   Partial Correlation      AC     PAC    Q-Stat  Prob

                                        1 -0.001 -0.001  0.0003
                                        2 -0.016 -0.016  0.0429
                                        3  0.055  0.055  0.5639  0.453
                                        4  0.107  0.108  2.5510  0.279
                                        5 -0.064 -0.063  3.2694  0.352
                                        6 -0.047 -0.048  3.6467  0.456
                                        7 -0.120 -0.137  6.1879  0.288
                                        8  0.013  0.007  6.2199  0.399
                                        9 -0.044 -0.029  6.5714  0.475
                                       10 -0.125 -0.107  9.3644  0.312
                                       11 -0.039 -0.023  9.6430  0.380
                                       12  0.023  0.003  9.7391  0.464
                                       13 -0.097 -0.091  11.465  0.405
                                       14 -0.015 -0.009  11.507  0.486
                                       15 -0.024 -0.041  11.613  0.560
                                       16  0.022  0.005  11.702  0.630
                                       17 -0.025 -0.037  11.819  0.693
                                       18 -0.032 -0.049  12.011  0.743
                                       19 -0.012 -0.025  12.037  0.798
                                       20  0.005 -0.044  12.042  0.845

Path = e:\pen\bna\bna_r7_janeiro2015\timberlake_r7   DB = none   WF = dados_st
```

Model selection criteria

Selection criteria are based on summary statistics from residuals, computed from a fitted model (or on forecast errors calculated from out-of-sample forecasts).

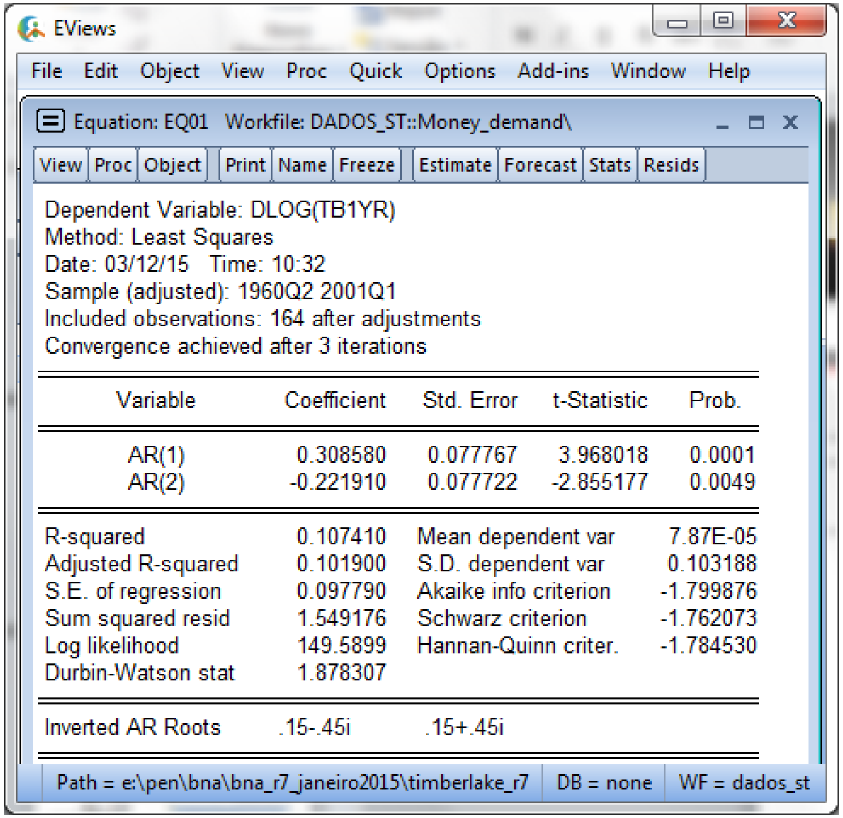
- **Akaike Information Criteria (AIC)**
  Assume that a statistical model of m parameters is fitted to a given time series. Akaike (1974) introduced an information criterion defined as
  $$AIC = -2\ln L + 2m,$$
  where L is the maximum likelihood and n is the effective number of observations (or number of computed residuals from the series). The EViews software computes the AIC value as
  $$AIC = n\ln\hat{\sigma}_{\varepsilon}^{2} + 2m,$$
  where $\hat{\sigma}_{\varepsilon}^{2}$ is the residual variance for the fitted model.
- **Schwartz Bayesian criterion (SBC).** Schwartz (1978) introduced the following Bayesian criterion of model selection:
  $$SBC = n\ln\hat{\sigma}_{\varepsilon}^{2} + m\ln n,$$
  where $\hat{\sigma}_{\varepsilon}^{2}$ is the residual variance for the fitted model, m is the number of parameters and n is the effective number of observations.

Model selection criteria

Example: 1-Year US Treasury Bill

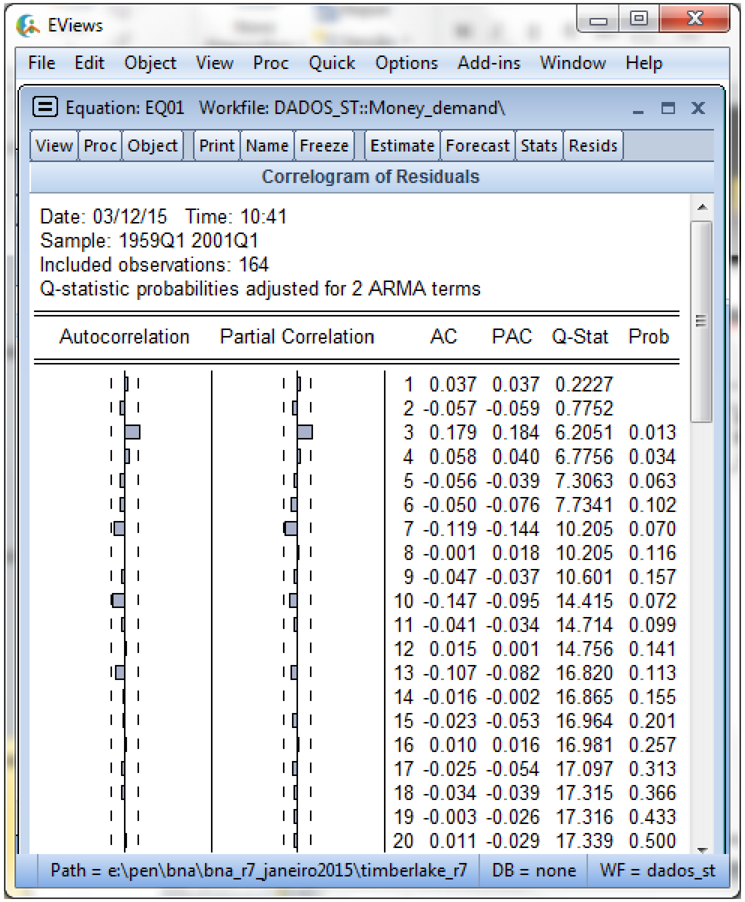
AIC, BIC and HQ values for TB1YR models

	ARIMA(0,1,2)	ARIMA(2,1,0)
AIC	-1.832	-1.800
BIC	-1.794	-1.762
HQ	-1.817	-1.785



```
Dependent Variable: DLOG(TB1YR)
Method: Least Squares
Date: 03/12/15   Time: 10:32
Sample (adjusted): 1960Q2 2001Q1
Included observations: 164 after adjustments
Convergence achieved after 3 iterations

        Variable        Coefficient    Std. Error    t-Statistic    Prob.

          AR(1)          0.308580      0.077767      3.968018      0.0001
          AR(2)         -0.221910      0.077722     -2.855177      0.0049

R-squared                0.107410    Mean dependent var       7.87E-05
Adjusted R-squared       0.101900    S.D. dependent var       0.103188
S.E. of regression       0.097790    Akaike info criterion   -1.799876
Sum squared resid        1.549176    Schwarz criterion       -1.762073
Log likelihood           149.5899    Hannan-Quinn criter.    -1.784530
Durbin-Watson stat       1.878307

Inverted AR Roots         .15-.45i      .15+.45i
```

Path = e:\pen\bna\bna_r7_janeiro2015\timberlake_r7   DB = none   WF = dados_st



Correlogram of Residuals

```
Date: 03/12/15   Time: 10:41
Sample: 1959Q1 2001Q1
Included observations: 164
Q-statistic probabilities adjusted for 2 ARMA terms

Autocorrelation   Partial Correlation      AC     PAC    Q-Stat   Prob

                                      1   0.037   0.037  0.2227
                                      2  -0.057  -0.059  0.7752
                                      3   0.179   0.184  6.2051   0.013
                                      4   0.058   0.040  6.7756   0.034
                                      5  -0.056  -0.039  7.3063   0.063
                                      6  -0.050  -0.076  7.7341   0.102
                                      7  -0.119  -0.144  10.205   0.070
                                      8  -0.001   0.018  10.205   0.116
                                      9  -0.047  -0.037  10.601   0.157
                                     10  -0.147  -0.095  14.415   0.072
                                     11  -0.041  -0.034  14.714   0.099
                                     12   0.015   0.001  14.756   0.141
                                     13  -0.107  -0.082  16.820   0.113
                                     14  -0.016  -0.002  16.865   0.155
                                     15  -0.023  -0.053  16.964   0.201
                                     16   0.010   0.016  16.981   0.257
                                     17  -0.025  -0.054  17.097   0.313
                                     18  -0.034  -0.039  17.315   0.366
                                     19  -0.003  -0.026  17.316   0.433
                                     20   0.011  -0.029  17.339   0.500
```

Path = e:\pen\bna\bna_r7_janeiro2015\timberlake_r7   DB = none   WF = dados_st

Forecasting

Suppose that at time $t = T$ we have the observations Y_T, Y_{T-1}, Y_{T-2}, ...

The minimum mean square error forecast of future value Y_{T+m} is defined in terms of the conditional expectation as a linear function of current and previous observations Y_T, Y_{T-1}, Y_{T-2}, …

$$\hat{Y}_T(m) = E_T(Y_{T+m}) = E(Y_{T+m} \mid Y_T, Y_{T-1}, Y_{T-2}, \ldots),$$

where $\hat{Y}_T(m)$ is the *m*-step ahead forecast of Y_{T+m}, *T* is the forecast *origin* and *m* is the *lead time* (or forecast horizon).

Forecasts for ARMA models

Consider the general stationary ARMA(*p*,*q*) model:

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

Because the model is stationary, it has an equivalent moving average representation

$$Y_t = \varepsilon_t + \psi_1\varepsilon_{t-1} + \psi_2\varepsilon_{t-2} + \cdots = \sum_{j=0}^{\infty}\psi_j\varepsilon_{t-j} = \psi(B)\varepsilon_t,$$

where $\psi_0 = 1$, ε_t is white noise and $\psi(B) = \sum_{j=0}^{\infty}\psi_j B^j = \dfrac{\theta(B)}{\phi(B)}$.

Forecasting

Suppose we have the observations $Y_1, Y_2, ..., Y_T$. For $t = T + m$, we have

$$Y_{T+m} = \sum_{j=0}^{\infty} \psi_j \varepsilon_{T+m-j} \,.$$

Standing at origin T, the forecast $\hat{Y}_T(m)$ of Y_{T+m} is defined as a linear combination of current and previous shocks $\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, ...$

$$\hat{Y}_T(m) = \psi_m^* \varepsilon_T + \psi_{m+1}^* \varepsilon_{T-1} + \psi_{m+2}^* \varepsilon_{T-2} + \cdots$$

where the weights $\psi_m^*, \psi_{m+1}^*, \psi_{m+2}^*, ...$ are to be determined. Then, the mean square error of the forecast is

$$E\left[Y_{T+m} - \hat{Y}_T(m)\right]^2 = \sigma_\varepsilon^2 \sum_{j=0}^{m-1} \psi_j^2 + \sigma_\varepsilon^2 \sum_{j=0}^{\infty} \left[\psi_{m+j} - \psi_{m+j}^*\right]^2 ,$$

which is minimized when $\psi_{m+j} = \psi_{m+j}^*$. Hence,

$$\hat{Y}_T(m) = \psi_m \varepsilon_T + \psi_{m+1} \varepsilon_{T-1} + \psi_{m+2} \varepsilon_{T-2} + \cdots$$

Since $E(\varepsilon_{T+m} \mid Y_T, Y_{T-1}, Y_{T-2}, \ldots) = 0, j > 0$, then the minimum mean square error forecast of Y_{T+m} is the conditional expectation. That is

$$\hat{Y}_T(m) = \psi_m \varepsilon_T + \psi_{m+1} \varepsilon_{T-1} + \psi_{m+2} \varepsilon_{T-2} + \cdots = E_T(Y_{T+m})$$

Forecasting

The forecast error for lead time m is

$$e_T(m) = Y_{T+m} - \hat{Y}_T(m) = \sum_{j=0}^{m-1} \psi_j \varepsilon_{T+m-j} \ .$$

Since $E_T[e_T(m)] = 0$, the variance of the forecast error is

$$Var[e_T(m)] = \sigma_\varepsilon^2 \sum_{j=0}^{m-1} \psi_j^2 \ .$$

Assuming the normality of ε's, the forecast limits are

$$Y_T(m) \pm z_{\alpha/2} \left[1 + \sum_{j=0}^{m-1} \psi_j^2 \right]^{1/2} \sigma_\varepsilon \ ,$$

where $z_{\alpha/2}$ is the standard normal deviate such that $P(Z > z_{\alpha/2}) = \alpha/2$.

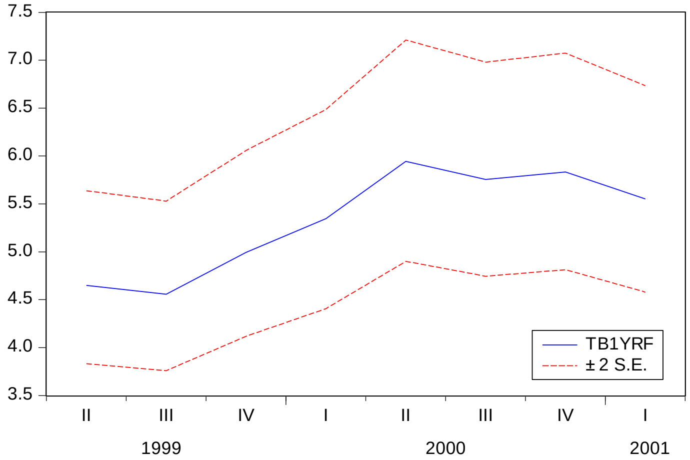
Find the m-step ahead forecast $\hat{Y}_T(m)$, the forecast error and the variance of the forecast error for AR(1), MA(1) and ARMA(1,1) models

Forecasting

LISBOA SCHOOL OF ECONOMICS & MANAGEMENT

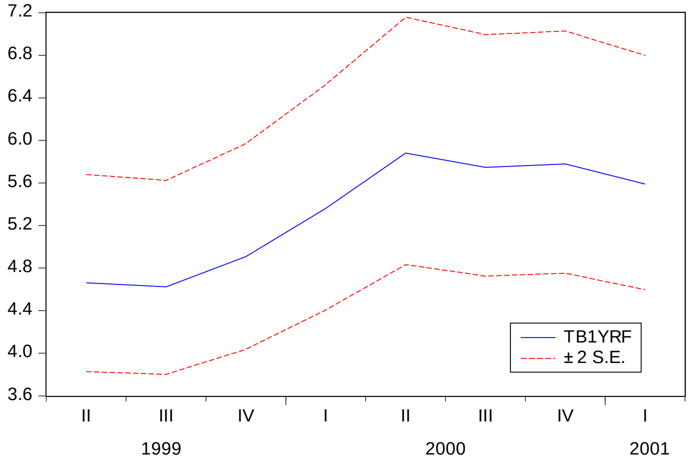
Example: 1-Year US Treasury Bill – **Static Forecasting**

ARIMA(0,1,2) model
log(TB1YR) series



```
Forecast: TB1YRF
Actual: TB1YR
Forecast sample: 1999Q2 2001Q1
Included observations: 8
Root Mean Squared Error      0.467643
Mean Absolute Error          0.322565
Mean Abs. Percent Error      6.540790
Theil Inequality Coefficient 0.043831
     Bias Proportion         0.007367
     Variance Proportion     0.006927
     Covariance Proportion   0.985705
```

ARIMA(2,1,0) model
log(TB1YR) series

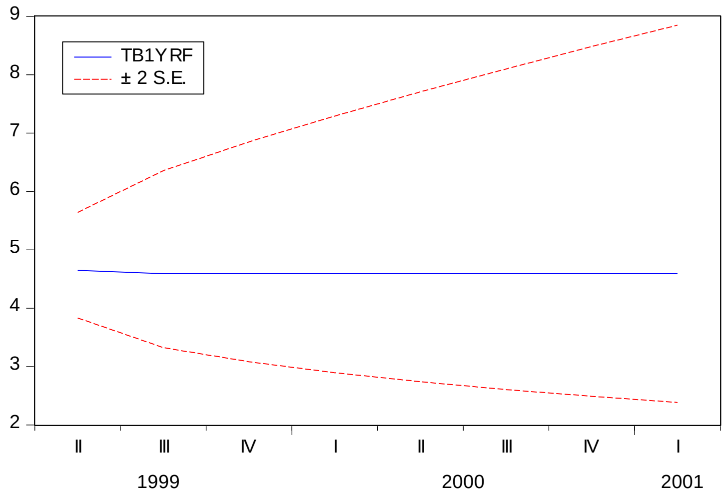


```
Forecast: TB1YRF
Actual: TB1YR
Forecast sample: 1999Q2 2001Q1
Included observations: 8
Root Mean Squared Error      0.476301
Mean Absolute Error          0.315519
Mean Abs. Percent Error      6.439636
Theil Inequality Coefficient 0.044692
     Bias Proportion         0.004036
     Variance Proportion     0.015674
     Covariance Proportion   0.980290
```

Forecasting

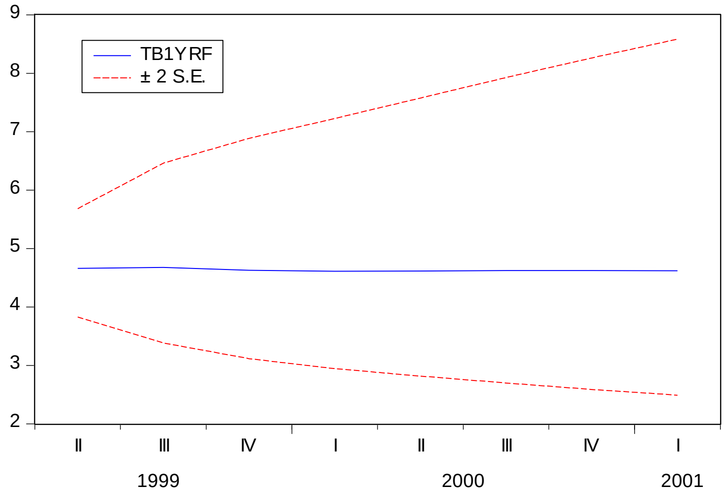
Example: 1-Year US Treasury Bill – **Dynamic Forecasting**

ARIMA(0,1,2) model
log(TB1YR) series



```
Forecast: TB1YRF
Actual: TB1YR
Forecast sample: 1999Q2 2001Q1
Included observations: 8
Root Mean Squared Error        0.883607
Mean Absolute Error            0.740036
Mean Abs. Percent Error        13.18370
Theil Inequality Coefficient   0.089116
     Bias Proportion           0.609898
     Variance Proportion       0.351946
     Covariance Proportion     0.038155
```

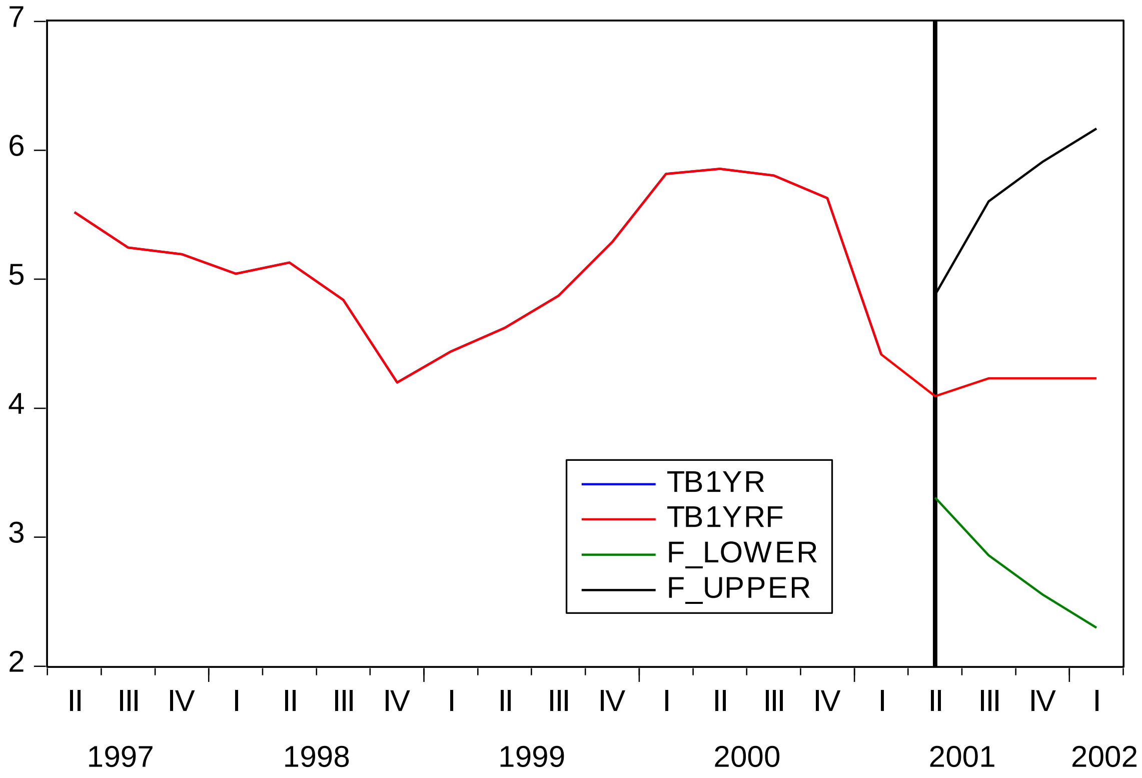
ARIMA(2,1,0) model
log(TB1YR) series



```
Forecast: TB1YRF
Actual: TB1YR
Forecast sample: 1999Q2 2001Q1
Included observations: 8
Root Mean Squared Error        0.861252
Mean Absolute Error            0.717839
Mean Abs. Percent Error        12.78505
Theil Inequality Coefficient   0.086577
     Bias Proportion           0.582794
     Variance Proportion       0.366424
     Covariance Proportion     0.050782
```

Forecasting

Example: 1-Year US Treasury Bill
Forecasts for h=1,2,3 and 4 steps ahead and 95% forecast limits
ARIMA(0,1,2) model



Exercises

1. Consider the quarterly unemployment rate (URATE) in U.S. between 1960:Q1 and 2008:Q1 (193 obs.) given in the EViews file "data_financial_econ.wk1" (Sheet 'Quarterly_US').

a) Describe the time plot. Do the data need transformation?

b) Identify a couple of ARIMA models that might be appropriate for the series.

c) Fit your best ARIMA model and carry out diagnostic checking on the residuals.

d) Produce forecasts for the next 4 periods using your preferred model.

e) Find the 95% forecast limits for forecasts in (d).

2. Consider the model

$$(1-B^4)(1-B)Y_t = (1-0.2B)(1-0.6B^4)\varepsilon_t$$

where ε_t is white noise. Find the eventual forecast function that generates the forecasts.

Exercises

3. Let Y_t be a stationary zero-mean process. Consider the models

$$X_t = (1-0.4B)Y_t \quad \text{and} \quad W_t = (1-2.5B)Y_t$$

a) Find the autocovariance generating functions of X_t and W_t.

b) Show that ACF of the above processes are identical.

4. Consider the ARIMA(0,2,3) model.

a) Write the model in terms of the backshift operator and without using the backshift operator.

b) Find the eventual forecast function.

5. Consider the ARIMA(0,1,1) model. Show that

$$Var\left[e_t(m)\right] = \sigma_\varepsilon^2\left[1+(m-1)(1-\theta)^2\right]$$

Exercises

6. Consider the model

$$(1-0.2B)(1-B)Y_t=(1-0.8B)\varepsilon_t$$

where $\sigma_\varepsilon^2=4$. Suppose we have the observations $Y_{49}=30$, $Y_{48}=25$ and $\varepsilon_{49}=-2$. Compute the forecast $\hat{Y}_{49}(m)$, for m = 50, 51, 52 and 53.

7. Consider the AR(2) model

$$(1-0.3B-0.6B^2)Y_t=\varepsilon_t$$

a) Find the MA representation of this model.

b) Find the PACF.

8. Consider the model

$$Y_t=2+1.3Y_{t-1}-0.4Y_{t-2}+\varepsilon_t+\varepsilon_{t-1}$$

a) Find the mean of Y_t.

b) Is the model invertible?

Exercises

9. Consider the model:

$$Y_t = 2 + \varepsilon_t - 0.6\varepsilon_{t-1}, \text{ with } \sigma_\varepsilon = 0.1$$

a) Find the eventual forecast function.

b) Find the variance of the forecast error.

10. Consider the ARMA(1,1) model. Show that

$$Var[e_t(m)] = \sigma_\varepsilon^2 \left(1 + \sum_{j=1}^{m-1} \phi^{2(j-1)} (\phi - \theta)^2 \right)$$

11. Consider the SARIMA(0,1,1)(0,1,1)$_{12}$ model

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

a) Write the model without using the backshift operator.

b) Suppose that $\theta_1 = 0.33$ e $\Theta_1 = 0.82$. Find the eventual forecast function.