ECOM073: Topics in Financial Econometrics

Queen Mary, University London, 2012-13
Lecturer: Liudas Giraitis, CB301, L.Giraitis@qmul.ac.uk

Exercise 4.

Problem 4.1.

Determine whether the following AR processes are stationary.
(a) $X_t = -10 + 0.3X_{t-1} + 0.5X_{t-2} + \varepsilon_t$.
(b) $X_t = 35 + 0.3X_{t-1} + 0.7X_{t-2} + \varepsilon_t$.

Solution: For AR(2) process

$$X_t = a + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t.$$

to be stationary it is required that coefficients satisfy the the relations:

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

(a) To verify if the process (a) is stationary we check

$\phi_2 + \phi_1 = 0.5 + 0.3 = 0.8 < 1$
$\phi_2 - \phi_1 = 0.5 - 0.3 = 0.2 < 1, \quad -1 < \phi_2 = 0.5 < 1.$
All conditions are satisfied. Therefore process (a) is stationary.

(b) To verify if the process (b) is stationary we check

$\phi_2 + \phi_1 = 0.7 + 0.3 = 1$
$\phi_2 - \phi_1 = 0.7 - 0.3 = 0.4 < 1, \quad -1 < \phi_2 = 0.7 < 1.$
Since condition $\phi_2 + \phi_1 < 1$ is not satisfied, the process (b) is non-stationary.

Problem 4.2. Find the mean of the following stationary AR processes, where ε_t is a white noise with zero mean and variance σ_ε^2.

(a) $X_t = 40 + 0.5X_{t-1} + \varepsilon_t$.

(b) $X_t = 20 + 0.2X_{t-1} + 0.6X_{t-2} + \varepsilon_t$.

Solution. (a) First we find the mean of an AR(1) process $X_t = a + \phi X_{t-1} + \varepsilon_t$ where $|\phi| < 1$ Such time series is stationary.

We take expectation of both sides of AR(1) equation:

$$\begin{aligned} E[X_t] &= E[a + \phi X_{t-1} + \varepsilon_t] \\ &= a + E[\phi X_{t-1}] + E[\varepsilon_t] \\ &= a + \phi E[X_{t-1}] \end{aligned}$$

since $E[\varepsilon_t] = 0$. Since X_t is a stationary process, then $E[X_t] = E[X_{t-1}] = \mu_X$ does not depend on time t. Therefore

$$\mu_X = a + \phi\mu_X, \quad or \quad \mu_X = \frac{a}{1-\phi}.$$

Since $a = 40$ and $\phi = 0.5$, we obtain

$$\mu_X = \frac{40}{1-0.5} = 80.$$

(b) First we find the mean of AR(2) process

$$X_t = a + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t.$$

Again, since X_t is stationary, then $EX_t = \mu_X$ does not depend on t. Taking expectation of both sides, we get

$$\begin{aligned} E[X_t] &= E[a + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t] \\ &= a + E[\phi_1 X_{t-1}] + E[\phi_2 X_{t-2}] + E[\varepsilon_t] \\ &= a + \phi_1 E[X_{t-1}] + \phi_2 E[X_{t-2}]. \end{aligned}$$

So,

$$\mu_X = a + \phi_1\mu_X + \phi_2\mu_X, \quad or \quad \mu_X = \frac{a}{1-\phi_1-\phi_2}.$$

Now, $a = 20$, $\phi_1 = 0.2$, $\phi_2 = 0.6$. We obtain

$$\mu_X = \frac{20}{1-0.2-0.6} = \frac{20}{0.2} = 100.$$

Problem 4.3. Assume that sample size is $N = 400$, and the sample autocorrelation function at lags $1, 2, ..., 9, 11, 12$ is taking values

$$-0.24, 0.32, -0.15, 0.12, 0.001, 0.05, 0.01, 0.011, 0.009, 0.04, 0.002, 0.003.$$

while PACF is taking values

$$0.4, 0.2, -0.05, 0.012, 0.04, 0.015, 0.011, 0.031, 0.0019, 0.024, 0.0022, 0.0033.$$

(i) Test at 5% significance level, that this time series is a white noise.

(ii) What p would you use fitting AR(p) model to this data? Explain your decision.

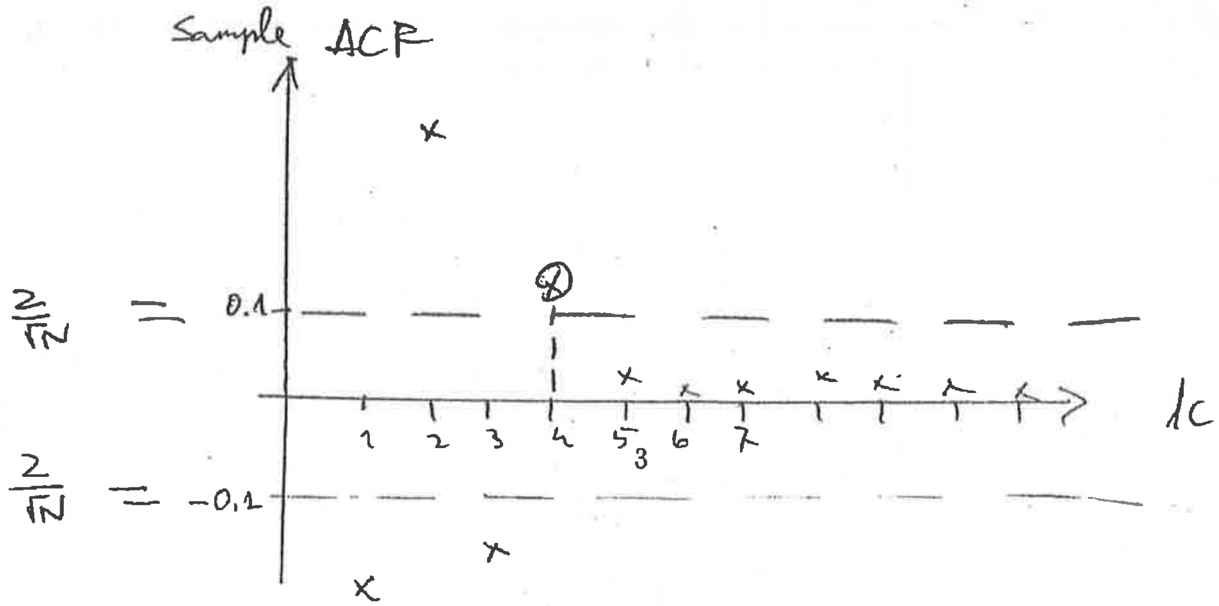
(iii) What q would you use fitting MA(q) model to this data?

(iv) Which model, AR or MA, would you apply?

Solution. (i) To answer this question, we test for significance of correlation ρ_k at lags $k \geq 1$ at significance level 5%.

- Correlation at lag k is significant, if $|\hat{\rho}_k| > 2/\sqrt{N}$, where N is the number of observations.
- If $|\hat{\rho}_k| \leq 2/\sqrt{N}$, then correlation at lag k is not significantly different from 0.

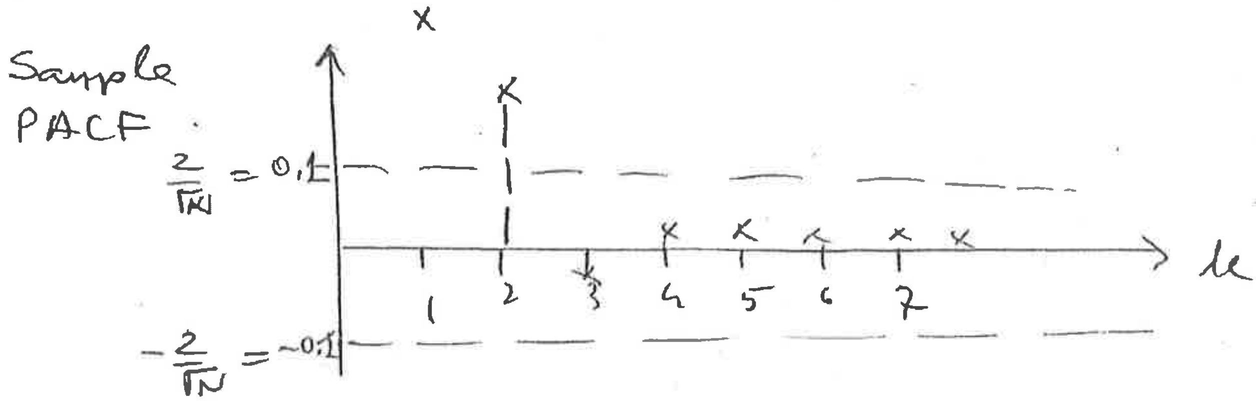
We have $2/\sqrt{N} = 2 = 2/\sqrt{400} = 0.1$. The following graph of ACF shows significant correlation at lags 1,2,3,4. So the time series is not a white noise.

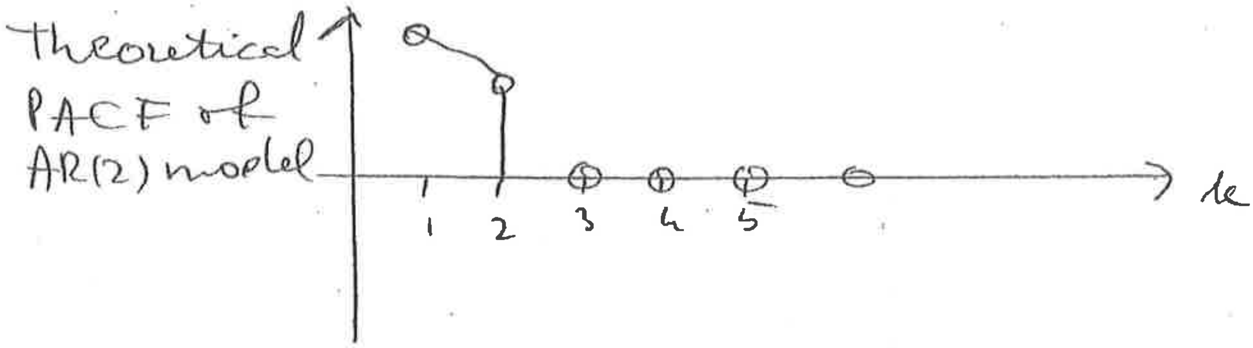


(ii) Fittting AR(p) model to the data, we select the p as the largest lag at which the PACF is significant. The rule to determine if PACF is significant at lag $k \geq 1$ is the same as for ACF. Such rule can be used because the PAFC of the AR(p) model becomes 0 for lags greater than p.

The graph of PACF below shows that the last significant PACF is at lag 2. So we fit to the data AR(2) model:

$$X_t = a + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t.$$

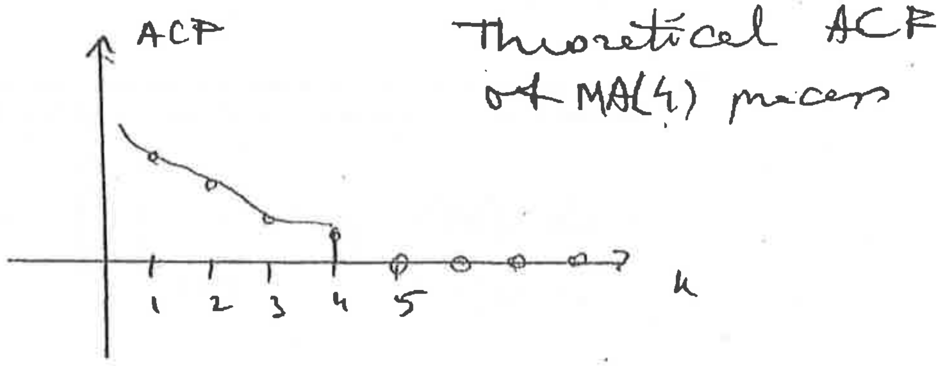




(iii) Fittting MA(q) model to the data, we select the q as the largest lag at which the ACF is significant. This rule can be used because the AFC of the MA(q) model becomes 0 for lags greater than q.

We above graph shows that the last significant ACF is at lag 4. So we fit to the data MA(4) model:

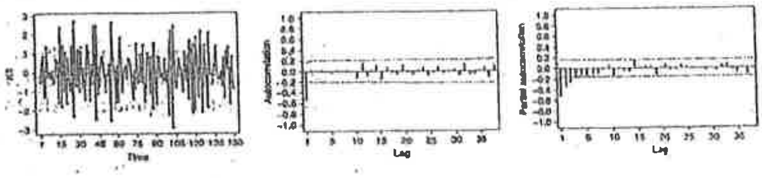
$$X_t = \theta_0 + \varepsilon_t - \theta_1\varepsilon_{t-1} - \cdots - \theta_4\varepsilon_{t-4}.$$



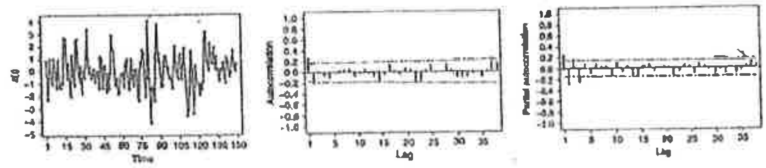
(iv) We can fit to the data either AR(2) model or MA(4) model. We select AR(2) model which has smaller number of parameters.

Problem 4.4.

3.9 For the time series plot and corresponding ACF and PACF plots below, determine the orders p and q of a tentative ARMA(p, q) model that can be used for this data



3.10 For the time series plot and corresponding ACF and PACF plots below, determine the orders p and q of a tentative ARMA(p, q) model that can be used for this data



Problem 4.5. Consider the monthly log returns of CRSP equally-weighted index 1962 to 1999 for 456 observations

(i) Build an AR model for the series and check for the fitted model.

(ii) Build an MA model for the series and check for the fitted model.

(iii) Compare the fitted AR and MA models.

Solution (i) PACF suggest fitting AR(1) model.

EViews - [Series: SERIES01 Workfile: M-EW6299::Untitled\]

File Edit Object View Proc Quick Options Add-ins Window Help

View Proc Object Properties Print Name Freeze Sample Genr Sheet Graph Stats Ident

Date: 02/07/12 Time: 11:35
Sample: 1 456
Included observations: 456

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.226	0.226	23.518	0.000
		2	-0.010	-0.065	23.564	0.000
		3	-0.038	-0.022	24.234	0.000
		4	-0.016	-0.002	24.349	0.000
		5	0.009	0.012	24.388	0.000
		6	-0.009	-0.016	24.424	0.000

Fitting AR(1) model : we get

$$X_t = 1.06 + 0.227 X_{t-1} + \varepsilon_t, \qquad \hat{\sigma}_\varepsilon = 5.4655$$

Estimation

EViews - [Equation: UNTITLED Workfile: M-EW6299::Untitled\]

File Edit Object View Proc Quick Options Add-ins Window Help

View Proc Object Print Name Freeze Estimate Forecast Stats Resids

Dependent Variable: SERIES01
Method: Least Squares
Date: 02/07/12 Time: 16:32
Sample (adjusted): 2 456
Included observations: 455 after adjustments
Convergence achieved after 3 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.069163	0.331516	3.225074	0.0014
AR(1)	0.227095	0.045829	4.955257	0.0000

R-squared	0.051417	Mean dependent var	1.063580
Adjusted R-squared	0.049323	S.D. dependent var	5.605513
S.E. of regression	5.465524	Akaike info criterion	6.239183
Sum squared resid	13531.99	Schwarz criterion	6.257294
Log likelihood	-1417.414	Hannan-Quinn criter.	6.246318
F-statistic	24.55457	Durbin-Watson stat	1.968754
Prob(F-statistic)	0.000001		

Inverted AR Roots	.23

Residuals Identification

EViews - [Equation: UNTITLED Workfile: M-EW6299::Untitled\]

File Edit Object View Proc Quick Options Add-ins Window Help

View Proc Object Print Name Freeze Estimate Forecast Stats Resids

Correlogram of Residuals

Date: 02/07/12 Time: 16:34
Sample: 2 456
Included observations: 455
Q-statistic probabilities adjusted for 1 ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.015	0.015	0.0980	
		2	-0.055	-0.056	1.5119	0.219
		3	-0.037	-0.036	2.1501	0.341
		4	-0.012	-0.014	2.2169	0.529
		5	0.015	0.011	2.3150	0.678
		6	0.002	-0.001	2.3172	0.804

Fitting AR(2) model: since PACF at lag 2 is rather large, we try also to fit AR(2) model:

$$X_t = 1.06 + 0.241X_{t-1} - 0.064\varepsilon_t, \qquad \sigma_\varepsilon = 5.466.$$

The coefficient −0.064 is not significant. Hence AR(1) is preferable over AR(2).

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.063964	0.311843	3.411856	0.0007
AR(1)	0.241579	0.047036	5.136077	0.0000
AR(2)	-0.064260	0.047121	-1.363735	0.1733

R-squared	0.055353	Mean dependent var	1.062548
Adjusted R-squared	0.051164	S.D. dependent var	5.611654
S.E. of regression	5.466211	Akaike info criterion	6.241635
Sum squared resid	13475.64	Schwarz criterion	6.268847
Log likelihood	-1413.851	Hannan-Quinn criter.	6.252356
F-statistic	13.21356	Durbin-Watson stat	2.000597
Prob(F-statistic)	0.000003		

Inverted AR Roots	.12-.22i	.12+.22i

Fitting MA(1) model: we get

$$X_t = 1.06 + \varepsilon_t + 0.2390\varepsilon_{t-1}, \qquad \sigma_\varepsilon = 5.45$$

MA Estimation

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.061394	0.316180	3.356929	0.0009
MA(1)	0.239034	0.045587	5.243483	0.0000

R-squared	0.054539	Mean dependent var	1.059511
Adjusted R-squared	0.052457	S.D. dependent var	5.600024
S.E. of regression	5.451166	Akaike info criterion	6.233913
Sum squared resid	13490.71	Schwarz criterion	6.251994
Log likelihood	-1419.332	Hannan-Quinn criter.	6.241035
F-statistic	26.18910	Durbin-Watson stat	1.995328
Prob(F-statistic)	0.000000		

Inverted MA Roots	-.24

Residual identification

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.001	0.001	0.0008	
		2	-0.002	-0.002	0.0020	0.964
		3	-0.035	-0.035	0.5830	0.747
		4	-0.011	-0.011	0.6370	0.888
		5	0.012	0.012	0.7063	0.951
		6	-0.005	-0.006	0.7165	0.982

(iii): Models AR(1) and MA(1) seem to fit equally well. We may select $\overline{\text{AR}}(1)$ which is more intuitive.