Econometrics & Financial Markets

Time series analysis

Toulouse Business School
MSc BIF

Anna CALAMIA
a.calamia@tbs-education.fr

Time series data and analysis

Times series data

- Series of data points ordered in time
- E.g. series of index values or stock prices or returns over time (Y_t)
- Multivariate and univariate analysis
- Frequency: daily, weekly, monthly, quarterly, annual…

Date	MSFT_ExR	SP500_ExR
01/01/2020	0.07820	-0.00289
01/02/2020	-0.04931	-0.08514
01/03/2020	-0.02391	-0.12514
01/04/2020	0.13625	0.12677
01/05/2020	0.02244	0.04518
01/06/2020	0.11354	0.01828
01/07/2020	0.00730	0.05503
01/08/2020	0.10001	0.06999
01/09/2020	-0.06521	-0.03930
01/10/2020	-0.03744	-0.02773
01/11/2020	0.05723	0.10748
01/12/2020	0.04167	0.03707
01/01/2021	0.04285	-0.01118
01/02/2021	0.00563	0.05937

Linear regression (e.g. CAPM regression) of Y_t on X_t

- Stationarity assumption
- Non stationary TS ➡ possible spurious regression

Time series analysis: Outline

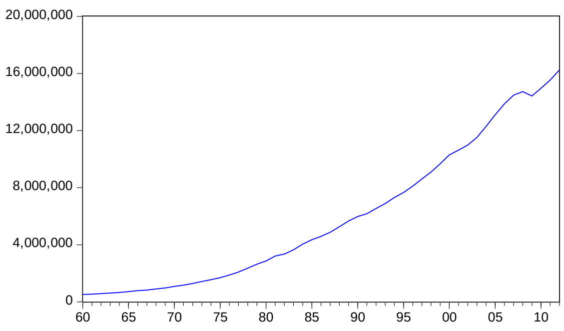
- Spurious Regression
- Stationarity Definition
- From non-stationarity to … Stationarity
- Induce Stationarity
- Modelling and forecasting stationary time series

Spurious Regression

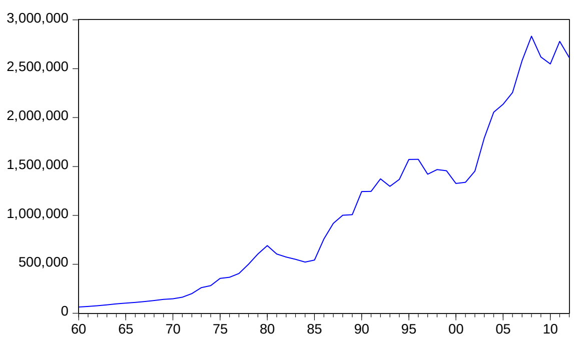
Example of Spurious Regression

- Example : regression of (Yt) to (Xt) and a constant

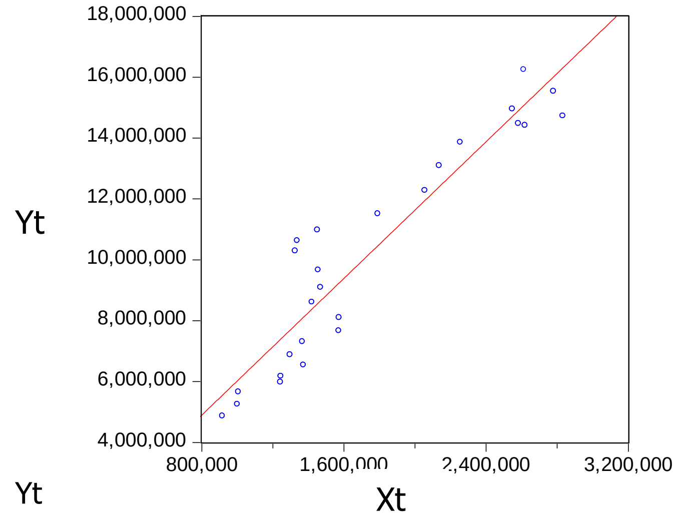
Yt



Xt



Comment???



Dependent Variable: Yt
Method: Least Squares
Date: 04/03/14 Time: 15:55
Sample: 1960 2012
Included observations: 53

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	93292.53	204437.6	0.456338	0.6501
Xt	5.764271	0.155177	37.14646	0.0000

R-squared	0.964357	Mean dependent var	5944537.
Adjusted R-squared	0.963658	S.D. dependent var	4976665.
S.E. of regression	948727.8	Akaike info criterion	30.40064
Sum squared resid	4.59E+13	Schwarz criterion	30.47499
Log likelihood	-803.6169	Hannan-Quinn criter.	30.42923
F-statistic	1379.859	Durbin-Watson stat	0.464692
Prob(F-statistic)	0.000000		

Why does stationarity matter?

If two variables are trending over time, a regression of one on the other could have a high R^2 even if the two are unrelated

- consider 2 series Xt and Yt having a similar trend and regress Xt on Yt. Applying the standard asymptotic properties of estimators, you will find that the conjunctions of Mars and Saturn is a powerful predictor of excess returns on NYSE (Novy-Marx 2014). What is the economic justification ?

If variables in the regression model not stationary:

- usual "*t*-ratios" will not follow a *t*-distribution
- cannot validly undertake hypothesis tests about the regression parameters.
- very high R^2 *and* *t*-statistic, but the results may have no economic meaning
- high level of residuals autocorrelation (DW very low)

➔ **Stationary processes: no <u>spurious regression</u>**

Stationarity Definition

Stationarity

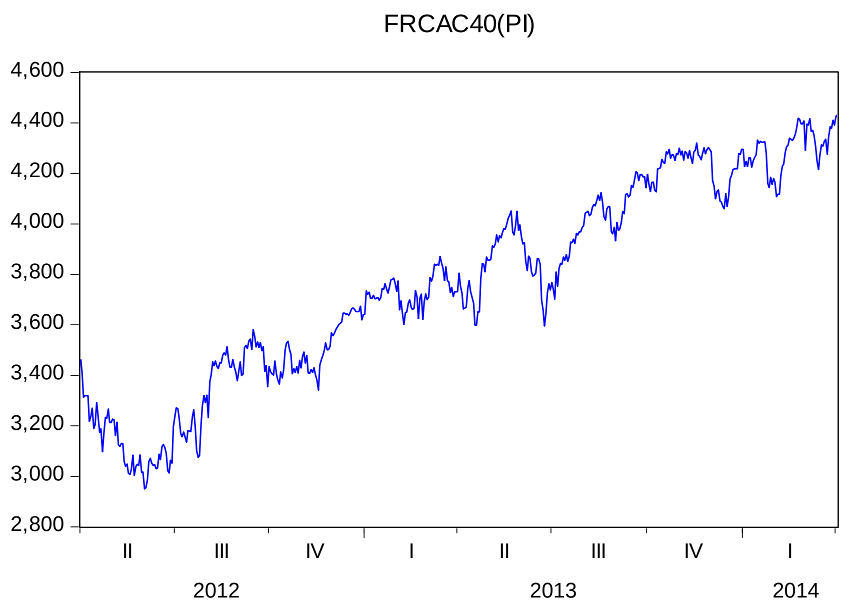
Y_t (with mean μ and variance σ^2) is a Weakly Stationary Process

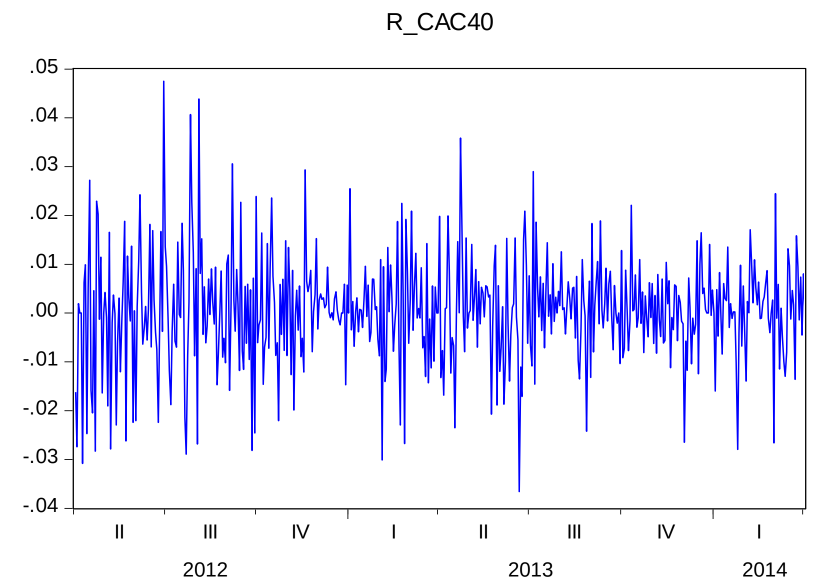
1. $E(Y_t) = \mu$,for all *t* → *Mean is constant over time*
2. $Var(Y_t) = \sigma^2$ → *Variance is constant over time*
3. $Cov(Y_t, Y_s) = \gamma(t-s), t \neq s$ → *Autocovariances do not depend on time, but only on the difference (t-s)*

Notation : $\gamma(0) = \sigma^2$

Weakly stationary (or covariance stationary) processes have no trend in mean, and no trend in variance, but it does not mean that they have a stable graphic …

Stationarity





Question 13: Which process seems stationary?
A- The series of values of CAC40 because it is mean stationary
B- The series of returns on CAC40 because it is mean stationary
C- The series of values of CAC40 because there is a positive trend
D- The series of returns on CAC40 because it's time trending

Autocorrelation Function

➔ use the **autocorrelations** $\tau(s)$: $\boxed{\tau(s)=\dfrac{\gamma(s)}{\gamma(0)}}$ s=0,1,2,… and $\gamma(s)$=cov(Yt,Ts)

- Autocorrelation of time series at various lags: Plot $\tau(s)$ against s=0,1,2,…

➔**autocorrelation function (autocorrelogram)**

$$\hat{\boldsymbol{\tau}}(\boldsymbol{s})=\frac{\sum_{t=s+1}^{T}(Y_t-\bar{Y})(Y_{t-s}-\bar{Y})}{\sum_{t=1}^{T}(Y_t-\bar{Y})^2},\ \mathbf{0}\le \boldsymbol{s}\le \boldsymbol{T}-\mathbf{1}$$

- ***Partial Autocorrelation Function***, *$\rho(k)$ is the coefficient of Yt-k in the regression of Yt on Yt-1, Yt-2, …, Yt-k*

  ➔ *measures the correlation between Yt and Yt-k after removing the effects of Yt-k+1 , Yt-k+2 , …, Yt-1 (conditional correlations)*

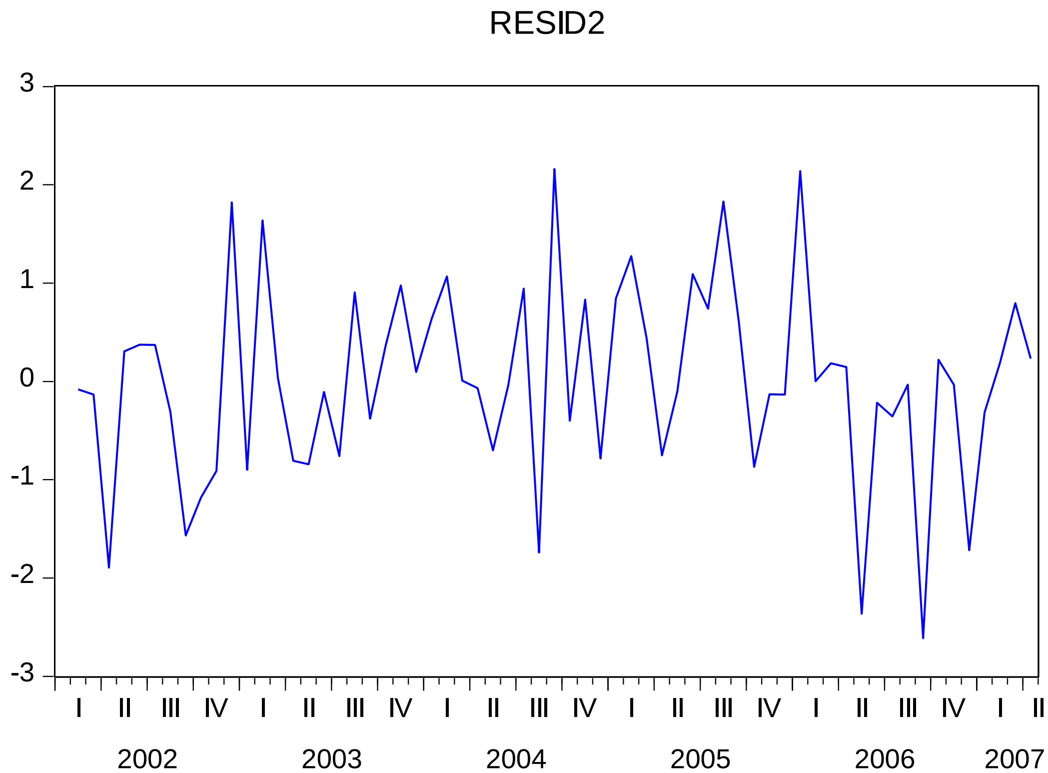
A particular stationary variable : White Noise Process

no discernible structure

$$E(Y_t) = 0$$

$$Var(Y_t) = \sigma^2$$ for all t

$$Cov(Y_t, Y_r) = 0$$ for t$\neq$ r



Significance tests for the autocorrelation coefficients

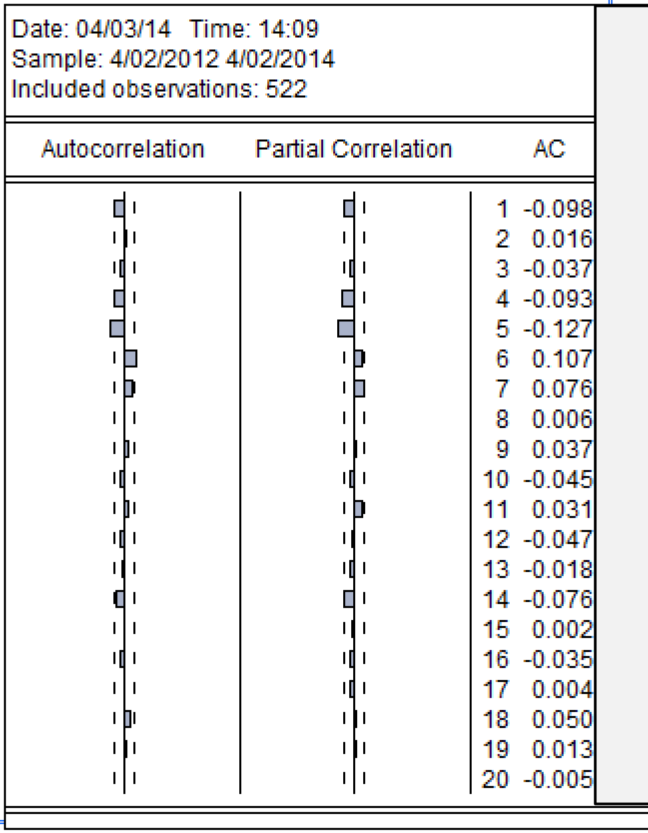
Significance tests for the autocorrelation coefficients at lag s, $\tau(s)$:

1. Compute the t-test of the null hypothesis, H0: $\tau(s)=0$
   - Under HO: $\hat{\tau}(s) \sim$ approximately $N(0,1/T)$
   - ➔ t-test: $\tau(s)/(1/\sqrt{T})$

     Where $1/\sqrt{T}$ is the standard error of $\tau(s)$ and

     T is the number of observations in the time series
   - Reject if t-test larger >1.96, in absolute value (5% critical value)

2. Equivalently, compute the 95% confidence interval as: $\boxed{0 \pm 1.96 \times \frac{1}{\sqrt{T}}}$ (reject if outside the interval)



Date: 04/03/14 Time: 14:09
Sample: 4/02/2012 4/02/2014
Included observations: 522

Autocorrelation	Partial Correlation		AC
		1	-0.098
		2	0.016
		3	-0.037
		4	-0.093
		5	-0.127
		6	0.107
		7	0.076
		8	0.006
		9	0.037
		10	-0.045
		11	0.031
		12	-0.047
		13	-0.018
		14	-0.076
		15	0.002
		16	-0.035
		17	0.004
		18	0.050
		19	0.013
		20	-0.005

❖ Question 14: is $\tau(1)$ significantly different from 0?

BOX-PIERCE / LJUNG-BOX Q test

➔ **BOX-PIERCE / LJUNG-BOX Q** tests the **joint hypothesis that all correlation coefficients are simultaneously equal to zero.**

Reject H0 if pvalue <5%: one coefficient is significantly different from zero (the process cannot be approximated by a white noise)

Date: 04/19/17 Time: 21:37
Sample: 1960Q1 2002Q1
Included observations: 168

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.018	0.018	0.0534	0.817
		2	-0.058	-0.058	0.6299	0.730
		3	0.000	0.002	0.6299	0.890
		4	0.146	0.143	4.3464	0.361
		5	-0.073	-0.080	5.2851	0.382
		6	0.125	0.148	8.0178	0.237
		7	-0.033	-0.053	8.2059	0.315
		8	-0.118	-0.126	10.707	0.219
		9	-0.137	-0.116	14.054	0.120
		10	0.065	0.012	14.819	0.139
		11	-0.069	-0.058	15.687	0.153
		12	-0.063	-0.044	16.416	0.173
		13	0.072	0.109	17.378	0.183
		14	0.036	0.027	17.622	0.225
		15	-0.018	0.043	17.683	0.280
		16	0.011	-0.013	17.707	0.341
		17	0.038	-0.005	17.983	0.390
		18	-0.002	-0.003	17.983	0.457
		19	0.023	-0.009	18.082	0.517
		20	0.106	0.085	20.256	0.442
		21	0.026	0.030	20.390	0.497
		22	0.021	0.073	20.473	0.553
		23	-0.006	-0.017	20.479	0.613
		24	0.120	0.123	23.328	0.501
		25	-0.086	-0.095	24.816	0.473
		26	0.079	0.076	26.082	0.459
		27	-0.023	-0.041	26.187	0.508

Question 15: All the pvalues >5%, which implies that ….
A- none of the coefficients are significant
B- all the coefficients are significant
C- the process is non stationary
D- The process is non gaussian

From non-stationarity to … Stationarity

3 types of Non-Stationarity

Various illustrations of non-stationarity :

- (1) the **random walk model with drift**
  (Difference Stationary with drift):

  $$Y_t = \mu + Y_{t-1} + U_t \quad , U_t \text{ WN} \quad (1)$$

- (2) the **random walk model without drift**
  ($\mu = 0$, DS without drift) :

  $$Y_t = Y_{t-1} + U_t \quad , U_t \text{ WN} \quad (2)$$

Variable at date t depends on the value of the previous period : shock at a period will have permanent rather than transitory effects (shocks persist in the system)

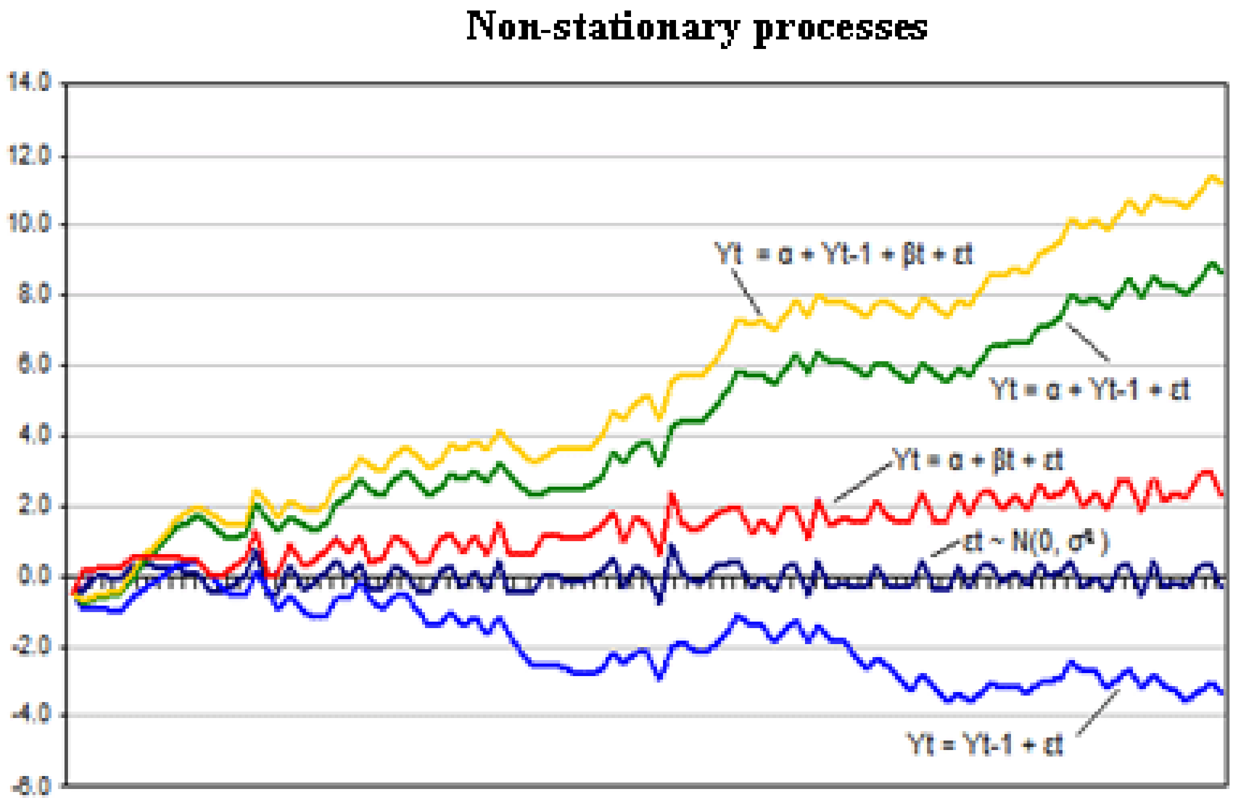
STOCHASTIC NON STATIONARITY

- (3) the **deterministic trend process** (Trend stationary):

  $$Y_t = \alpha + \beta t + U_t \quad , U_t \text{ WN} \quad (3)$$

- The series moves linearly in time
- ut is stationary : shocks have no impact on the later evolution, the series always returns to its long term trend

Non-Stationarity



Non-stationarity: study of shocks

- Generalization: consider the process defined by

  $Y_t = \phi Y_{t-1} + U_t$, with ϕ a generic parameter
- By T successive substitutions (of Yt-1…) we get:

$$Y_T = U_T + \phi U_{T-1} + \phi^2 U_{T-2} + \phi^3 U_{T-3} + \ldots + \phi^T y_0$$

1. $\phi < 1 \Rightarrow \phi^T \to 0$ as $T \to \infty$

   shocks to the system gradually die away➔ **stationarity**

2. $\phi = 1 \Rightarrow \phi^T = 1 \,\forall\, T$

   shocks persist in the system ➔ **random walk, non stationarity**

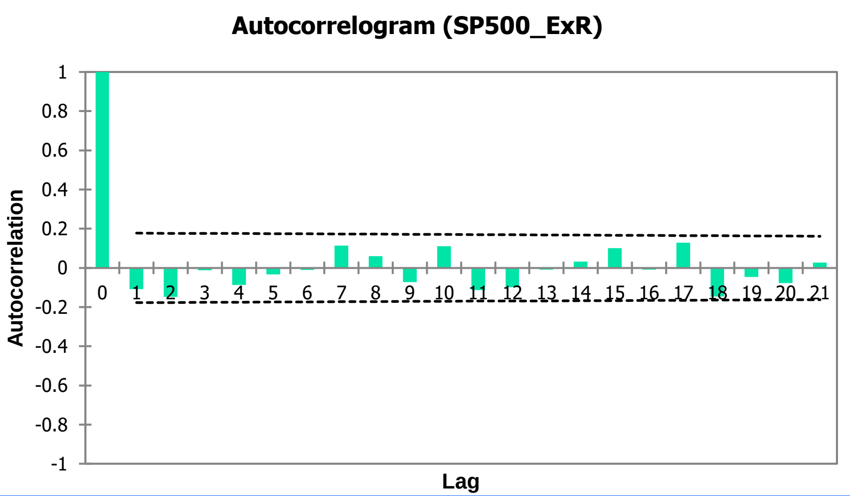
3. $\phi > 1$ shocks propagate and become <u>more</u> influential as time goes on➔ **explosive case, non stationarity**

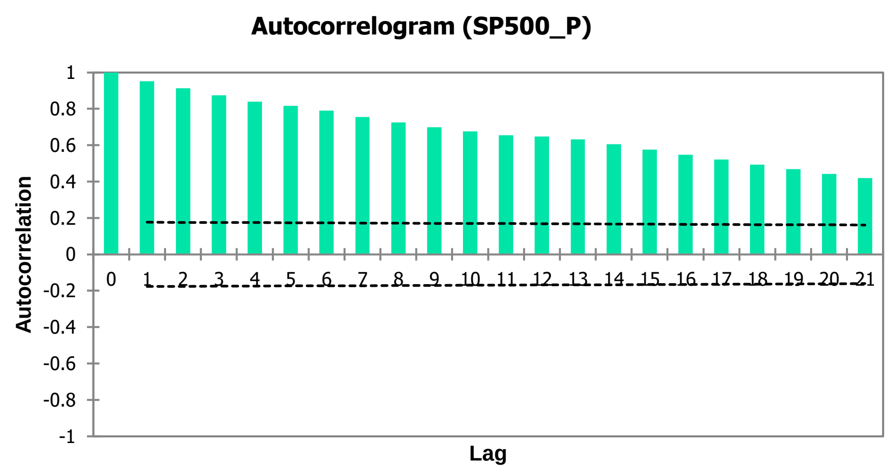
➔ Explosive case does not describe many data series in economics and finance.

we use ϕ = 1 to characterise the non-stationarity

Test for non stationarity

- Autocorrelogram :
  - ➔ For a stationary time series, either autocorrelations at all lags are statistically indistinguishable from zero, or the autocorrelations drop off rapidly to zero as the number of lags becomes large.
  - ➔ The autocorrelation function of a nonstationary process decreases very slowly even at very high lags.





Test for non stationarity

- Unit root test: If $\phi = 1 \Rightarrow$ the series has a unit root, it is a random walk and is not covariance stationary.

  - Dickey Fuller test based on a transformed version of the model:

$$Y_t = \mu + \phi Y_{t-1} + U_t$$

$$Y_t - Y_{t-1} = \mu + (\phi - 1)Y_{t-1} + U_t$$

  - The null hypothesis of the Dickey-Fuller test is H0: $\phi - 1 = 0$ and the alternative hypothesis is Ha: $\phi - 1 < 0$ (stationary).
  - Specific (larger) critical values (Dickey Fuller statistical tables)
  - Possible to add an intercept and a deterministic trend
  - If Yt serially correlated, may include lags of Yt (Augmented DF test)

Induce Stationarity

Induce Stationarity for Random Walk Process : Difference-Stationary series

Random walk with (or without) drift:

$$Y_t = \mu + Y_{t-1} + U_{t}, \qquad U_t \quad \text{WN} \qquad (1)$$

If we take (1) and subtract Y_{t-1} from both sides:

$$Y_t - Y_{t-1} = \mu + U_t$$

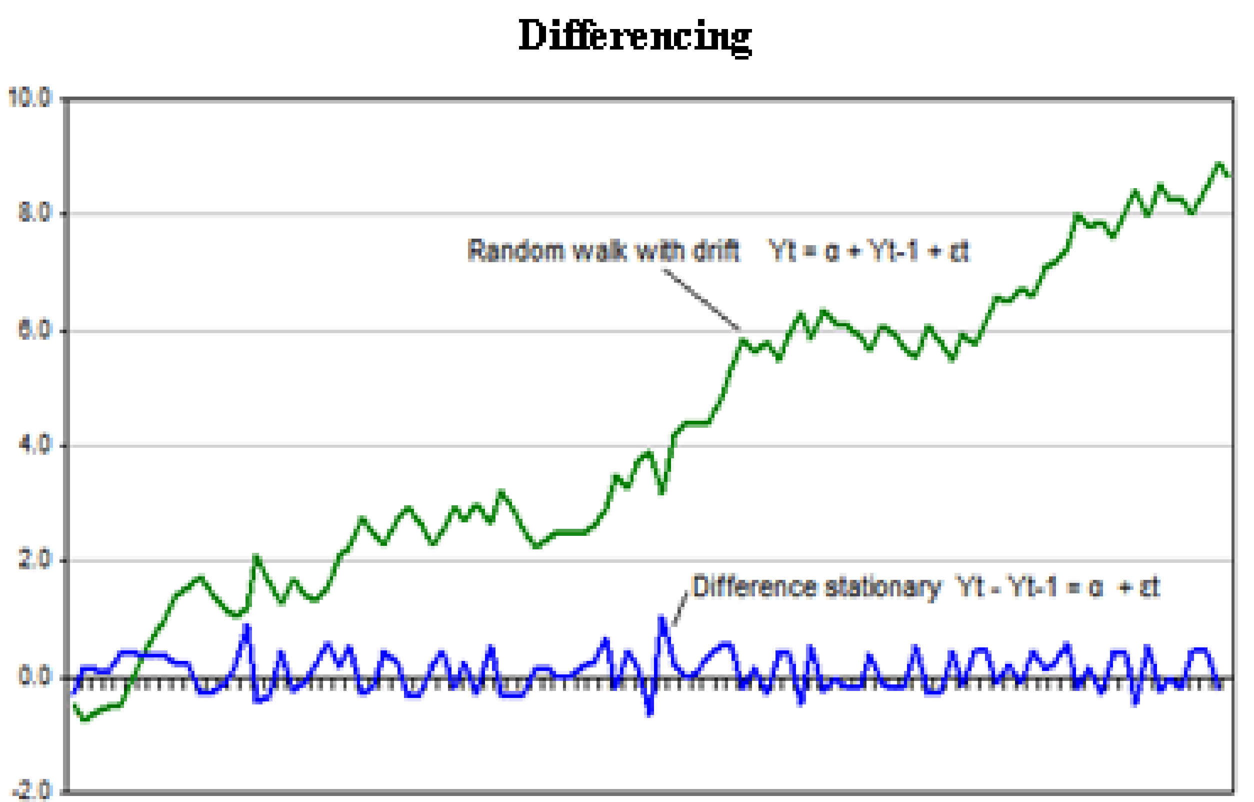
$$\Delta Y_t = \mu + U_t$$

We say that we have induced stationarity by "differencing once".

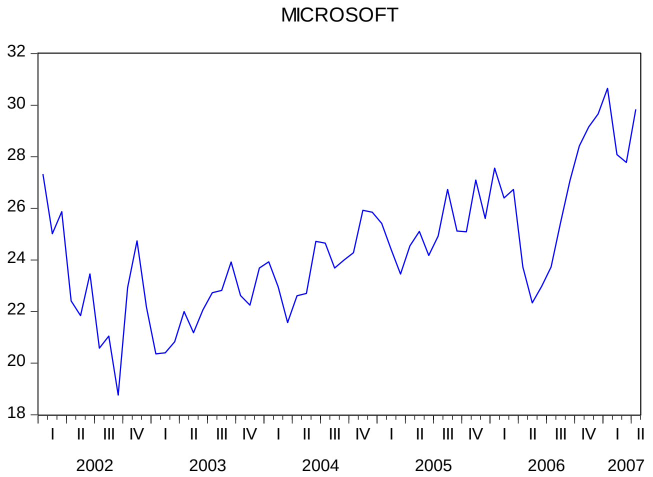
- A series is integrated of order 1 ($Y_t \sim I(1)$) if **Y_t is non-stationary but ΔY_t is stationary** (The series contains a unit-root)
- A series is integrated of order d ($Y_t \sim I(d)$) if **Y_t is non-stationary but $\Delta^d Y_t$ is stationary** (The series contains d unit-root)

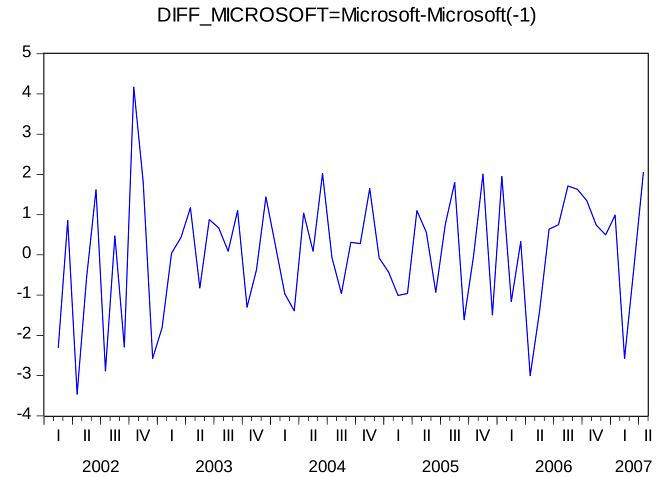
Most economic and financial series contain a single unit root

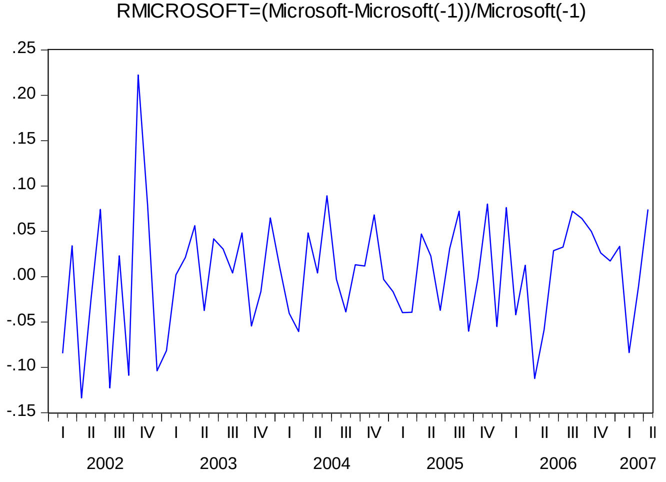
Induce Stationarity for Random Walk Process : Difference-Stationary series

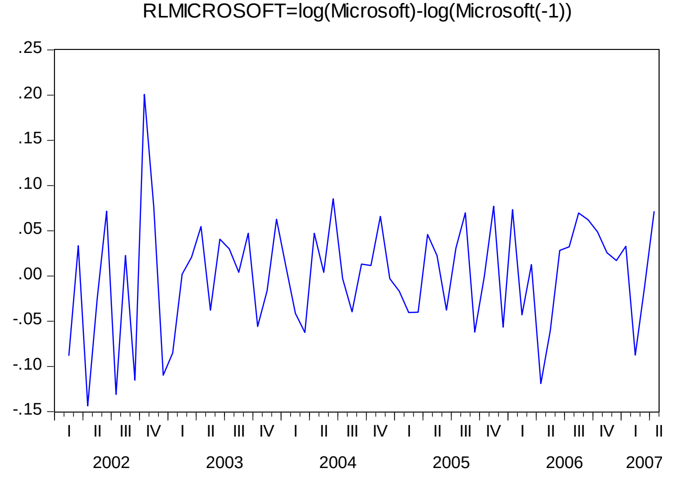


Example Price/Return









Induce Stationarity for a Deterministic Trend Process: Detrending

The trend-stationary process

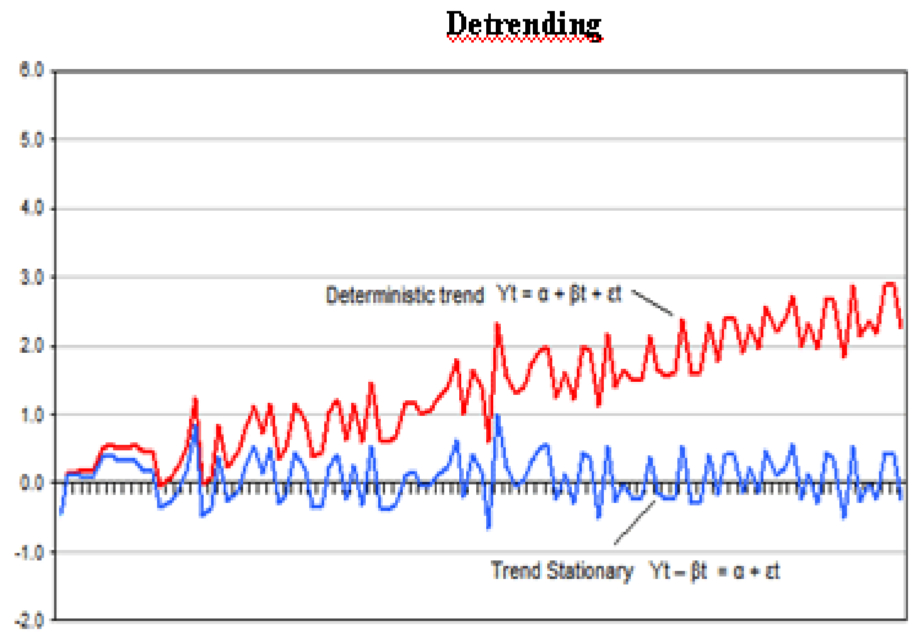
$$Y_t = \alpha + \beta t + U_t$$

➔ **deterministic non-stationarity**

- Subtracting the trend βt:

=> $Y_t - \beta t = \alpha + U_t$, is stationary

Detrending : run a regression of the form $Yt=\alpha+Bt+Ut$ and fit a model on the residuals from $Yt-Bt=\alpha+Ut$ (from which the linear trend has been removed)



Seasonality: regular patterns of movement within the year => include seasonal lags in an autoregressive model. Then, Y_t is non stationary but Y_{t-s} is stationary (quarterly (s=4) or annual(s=12) seasonality)

TUTORIAL XLSTAT

6. Time series: Test for stationarity

Tutorial

- XLSTAT – Time series analysis
- Plot the series as a function of time
  - ➔ Prices and returns
- Autocorrelation function and partial autocorrelation function
- Are the return series stationary? What about the price series?

Modelling stationary time series

Univariate Time Series Modelling and Forecasting

- Time-series models: to explain the past and to predict the future of a time series, Yt (stock price, return)
- Predicting or forecasting the future behaviour of financial variables :
  - Times series models use the information in past values of the same variable
  - Alternatively, regression models based on hypothesized causal relationships with other variables
- Objective:
  - We observe values of Yt, for t=1, … T.
  - We want to model this series and forecast futures values (T+1, T+2…) given its past values
    - One period ahead forecast
    - Multi period forecast (chain-rule)

Classical models

- Different models to represent univariate time series
- Classical TS models:

  Assumptions:
  - (Yt)t is a (weakly) stationary process
  - Ut is a white noise WN(0,σ^2)

- Autoregressive model: Yt is explained by its own past values
- Moving average: Yt explained by a moving average of current and past WN errors
- ARMA(p,q): Generalization of the first two models
- ARIMA Model: non stationary processes
- SARIMA(p,d,q): processes with seasonality s
- ARCH, GARCH: Autoregressive conditional heteroscedasticity

Classical models

- *Autoregressive model of order p, AR(p)*

$Y_t = \mu + \phi_1 Y_{t-1} + U_t$ AR(1)

$Y_t = \mu + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + U_t$ AR(2)

$Y_t = \mu + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + U_t$ AR(p)

- *Moving average of order q, MA(q)*

$Y_t = \mu + U_t + \theta_1 U_{t-1}$ MA(1)

$Y_t = \mu + U_t + \theta_1 U_{t-1} + \theta_2 U_{t-2}$ MA(2)

$Y_t = \mu + U_t + \theta_1 U_{t-1} + \theta_2 U_{t-2} + \ldots + \theta_q U_{t-q}$ MA(q)

- *ARMA(p,q)*

$$Y_t = \mu + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + U_t + \theta_1 U_{t-1} + \theta_2 U_{t-2} + \ldots + \theta_q U_{t-q}$$

- **ARIMA(p,d,q) and SARIMA(pdq)**

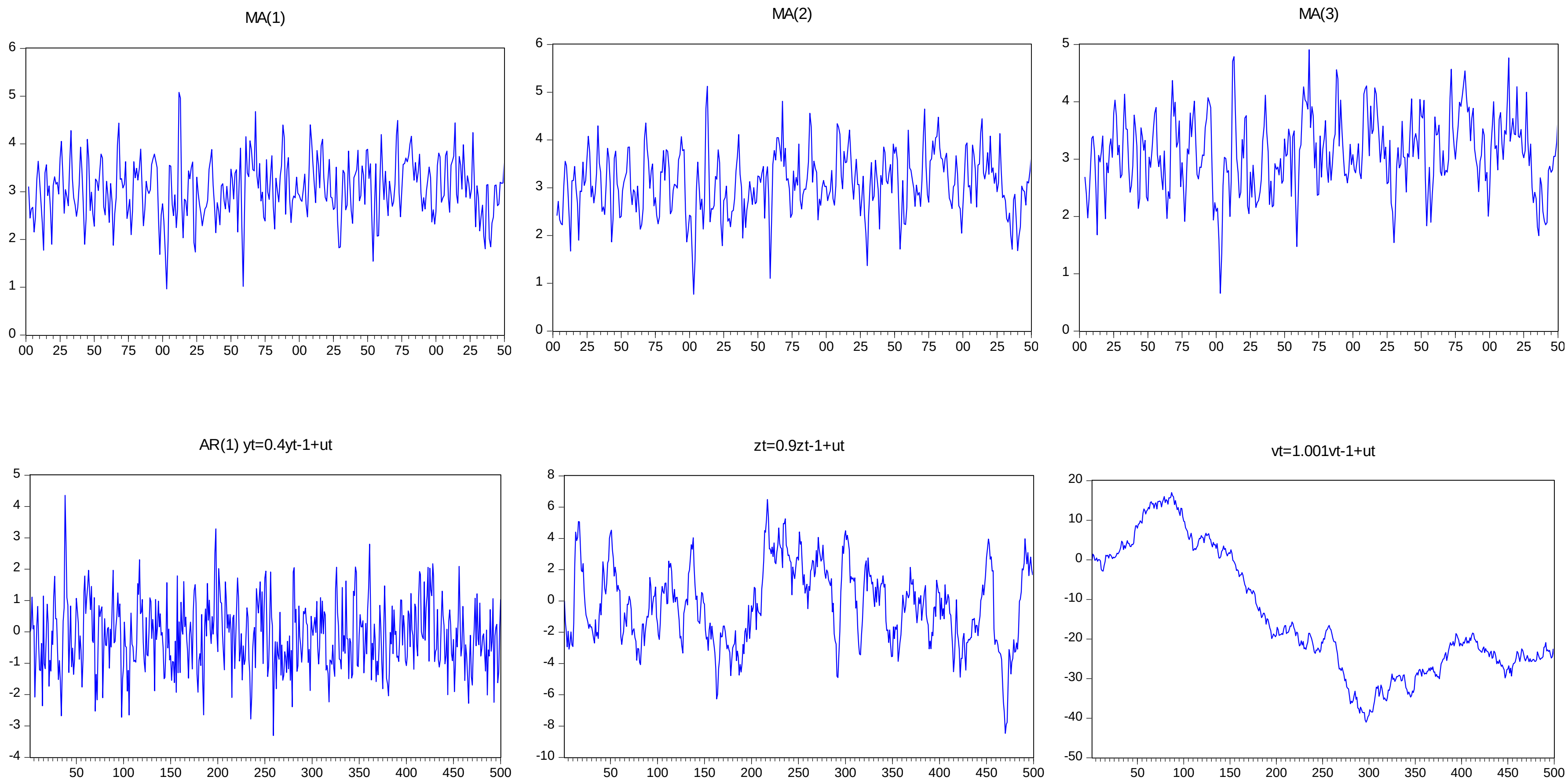
Y_t is I(1) if $\Delta Y_t = Y_t - Y_{t-1}$ is stationary

→ Y_t is an ARIMA(p,1,q) if ΔY_t is an ARMA(p,q) process

→ Y_t is an ARIMA(p,d,q) if $\Delta^d Y_t$ is an ARMA(p,q) process

Y_t is a SARIMA(p,d,q) process with seasonality s if Y_t is nonstationary but $Y_t - Y_{t-s}$ is stationary

Example of MA and AR Processes



Classical models

How to choose among these models?

How to select the parameters p, q?

Use the properties of the autocorrelation function to identify AR and MA processes and the order p,q.

ACF and PACF

- *If Yt is stationary AR(p), The autocorrelation function decays exponentially to zero (autocorrelations start large and decline gradually) while the PACF drops to zero after p:*
  - ➔ $\tau(k)$ ➔ *0* at exponential decay
  - ➔ $\rho(k)=0$ *for* $k>p$
  - ➔ $\rho(p) \neq 0$

- *If Yt is MA(q), it is weekly stationary and the autocorrelations drop to zero after first q autocorrelations, while The PACF decays to zero at an exponential decay:*
  - ➔ $\tau(k)=0$ *for* $k>q$
  - ➔ $\tau(q)\neq 0$
  - ➔ $\rho(k)$ ➔ *0* at exponential decay

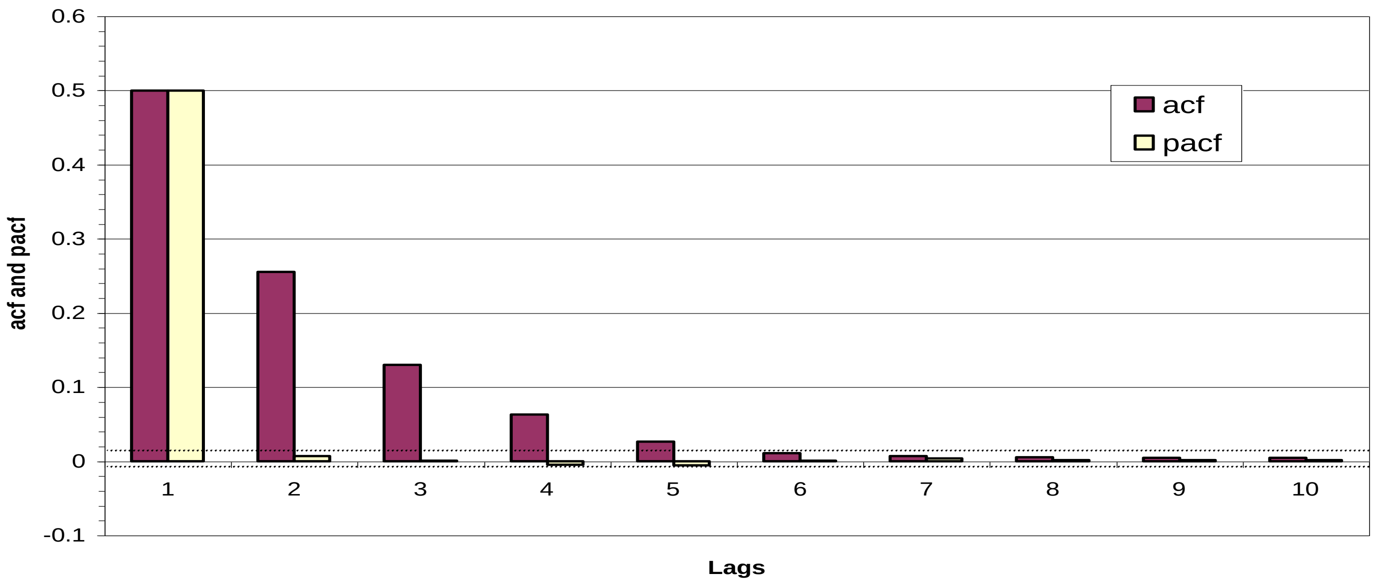
- *For a stationary (and invertible) ARMA process :*
  - ➔ *both acf and pacf are geometrically decaying*

Exemples of AR(1) Process

Question 16: Is $Y_t = 0.5Y_{t-1} + U_t$ stationary?
A- Yes because $\phi = 0.5$ is strictly less than 1
B- No because Y_t is an AR(1) which is always non stationary
C-Yes because Y_t is an AR(1) which is always stationary
D- No because Yt has a unit root

Question 17: Is $Y_t = 1.0Y_{t-1} + U_t$ stationary?
A- Yes because $\phi = 1$ is strictly larger than 1
B- No because Y_t is an AR(1) which is always non stationary
C-Yes because Y_t is an AR(1) which is always stationary
D- No because Yt is a Random Walk

Sample acf and pacf plots for standard processes



Question 18: Which model could fit the best?

A- an AR(5)

B- an AR(1)

C- a white moise

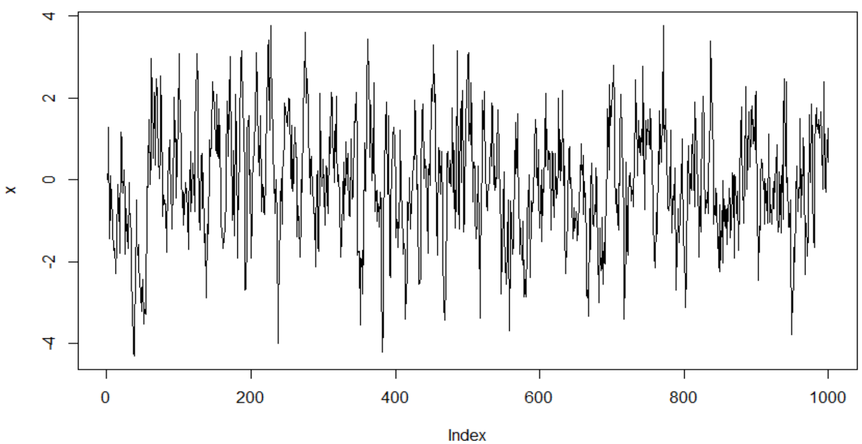
D- This series is non stationary

Model selection

- Check for stationarity of the series (trend, ACF and PACF, unit root)
- Transform if necessary (difference, detrend…)
- ACF, PACF to **choose the model AR or MA** and determine the order
  - ➔ choose the **parameter q** such that the autocorrelation values are not significant for any lag greater than q
  - ➔ choose the **parameter p** such that the partial autocorrelation values are not significant for any lag greater than p
- Estimate the model parameters and the residuals
- Model check:
  - ➔ Keep lag only if coefficient is significant
  - ➔ Residual diagnostics (residuals autocorrelation,…)
  - ➔ Goodness of fit $\overline{R}^2$ and information criteria for model selection:
    - AIC, SBC: based on RSS + correction for the number of parameters. The smaller the AIC or SBC the better the fit of the model.
  - ➔ The model should be parsimonious and plausible

Example: AR(1) model

$$Y_t = 0.7\,Y_{t-1} + U_t$$

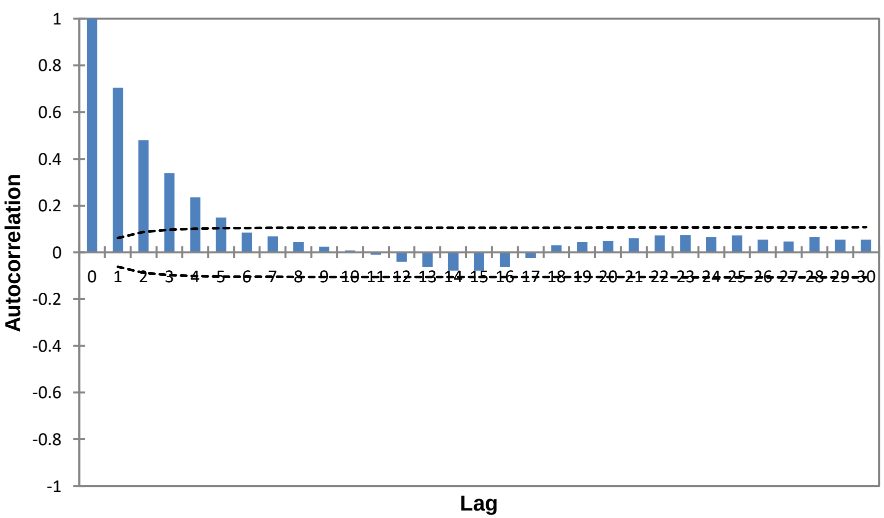


Dickey-Fuller test (x):

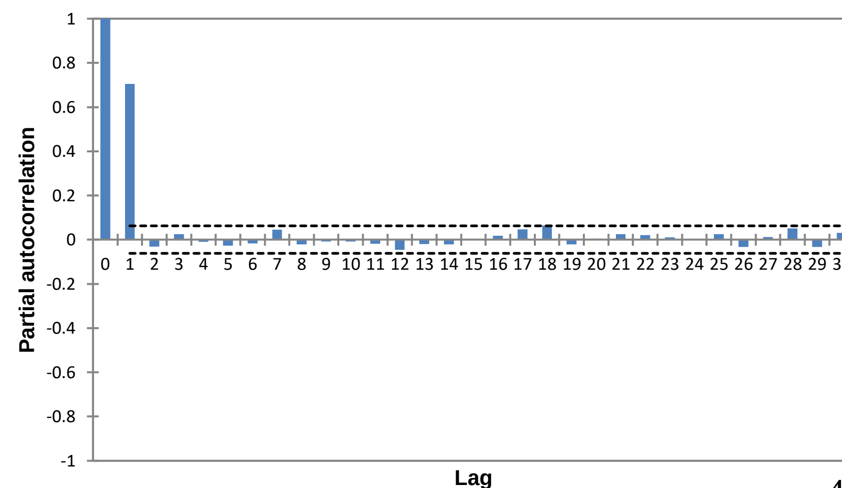
Tau (Observed value)	-8,428
Tau (Critical value)	-3,393
p-value (one-tailed)	< 0,0001
alpha	0,05

Statistic	DF	Value	p-value
Jarque-Bera	2	1,089	0,580
Box-Pierce	6	927,599	< 0,0001
Ljung-Box	6	931,142	< 0,0001
Box-Pierce	12	936,692	< 0,0001
Ljung-Box	12	940,329	< 0,0001

Autocorrelogramx



Partial autocorrelogramx



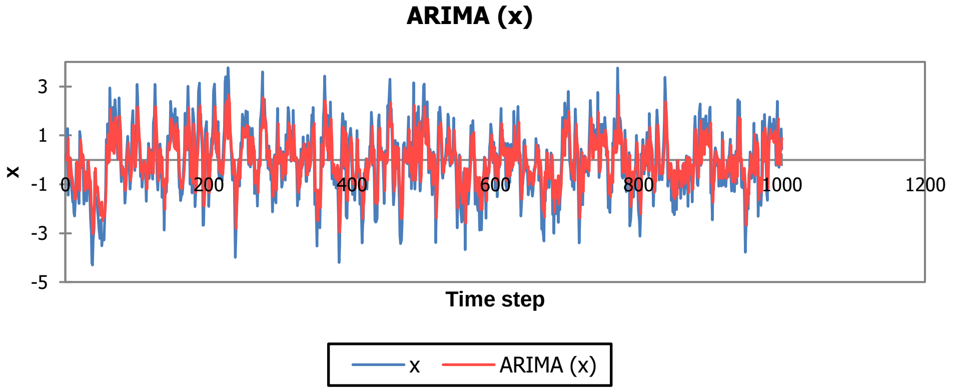
Example: AR(1) model

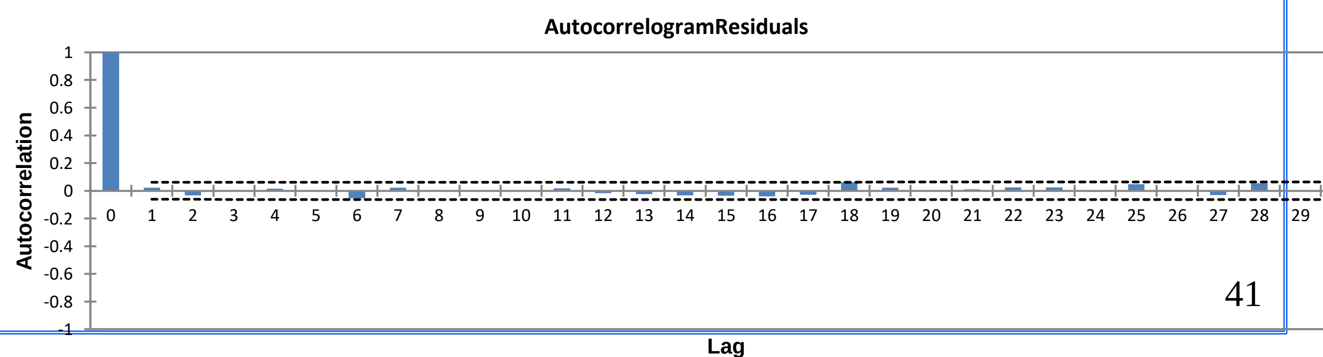
On XLSTAT (or R…) fit an Arima model with autoregressive order 1, 0 degrees of differencing, and an MA order of 0.

Goodness of fit statistics:	
Observations	1000
DF	998
SSE	992,8727718
MSE	0,992872772
RMSE	0,996430014
WN Variance	0,992872772
MAPE(Diff)	244,6768992
MAPE	244,6768992
-2Log(Like.)	2831,411794
FPE	0,994860505
AIC	2835,411794
AICC	2835,42383
SBC	2845,227304
Iterations	8

Model parameters:

Parameter	Value	standard error	Lower bound (95%)	Upper bound (95%)
Constant	0,000	0,107	-0,209	0,209
AR(1)	0,705	0,022	0,661	0,749





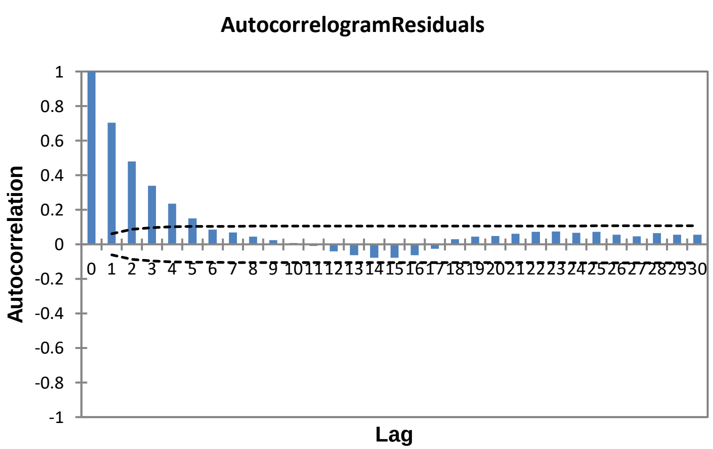
Example: AR(1) model

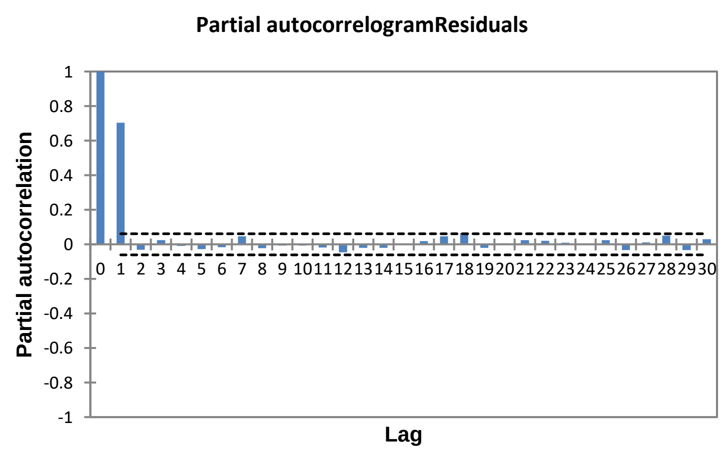
Compare with an ARIMA(2,0,0)

Model parameters:				
Parameter	Value	standard error	Lower bound (95%)	Upper bound (95%)
Constant	0,000	0,103	-0,203	0,203
Parameter	Value	standard error	Lower bound (95%)	Upper bound (95%)
AR(1)	0,727	0,032	0,665	0,789
AR(2)	-0,030	0,032	-0,092	0,031

Goodness of fit statistics:	
Observations	1000
DF	997
SSE	991,9509
MSE	0,991951
RMSE	0,995967
WN Variance	0,991951
MAPE(Diff)	240,7977
MAPE	240,7977
-2Log(Like.)	2830,485
FPE	0,995927
AIC	2836,485
AICC	2836,509
SBC	2851,208
Iterations	57

Compare with an ARIMA(0,0,0)





Forecasting with ARMA Models

- We have estimated an AR(2)
- We are at time t and we want to forecast 1,2,..., s steps ahead.
- We know Y_t , Y_{t-1}, ..., and U_t , U_{t-1}, ...

$$Y_t = \mu + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + U_t$$
$$Y_{t+1} = \mu + \phi_1 Y_t + \phi_2 Y_{t-1} + U_{t+1}$$
$$Y_{t+2} = \mu + \phi_1 Y_{t+1} + \phi_2 Y_t + U_{t+2}$$
$$Y_{t+3} = \mu + \phi_1 Y_{t+2} + \phi_2 Y_{t+1} + U_{t+3}$$

$$f_{t,1} = \mathrm{E}(Y_{t+1\,|\,t}) = \mathrm{E}_t(\mu + \phi_1 Y_t + \phi_2 Y_{t-1} + U_{t+1}) = \mu + \phi_1 Y_t + \phi_2 Y_{t-1}$$
$$f_{t,2} = \mathrm{E}(Y_{t+2\,|\,t}) = \mathrm{E}_t(\mu + \phi_1 Y_{t+1} + \phi_2 Y_t + U_{t+2}) = \mu + \phi_1 f_{t,1} + \phi_2 Y_t$$
$$\ldots$$
$$f_{t,s} = \mu + \phi_1 f_{t,s-1} + \phi_2 f_{t,s-2}$$

- Similarly, we can generate forecasts for a MA(q) and for ARMA(p,q)

In-Sample and Out-of-Sample

- In-sample forecasts: predicted values from the estimated time-series model (generated for the same set of data used to estimate the model's parameters).
- Out-of-sample forecasts: forecasts made from the estimated time-series model for a time period different from the one for which the model was estimated.
- Holdout sample: last observations of the sample used to construct out-of-sample forecasts and test the model performance.

→ Ability of the forecast:

- Measures of out of sample forecast accuracy: RMSE (root mean square error) measures of the difference between values predicted by a model and the actual values:

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n} (Y_{obs,t} - \hat{Y}_{model,t})^2}{n}}$$

- Other measures:

  MAE (Mean Absolute Error), MAPE (Mean Absolute Percentage Error)

- The model with the smallest values for RMSE provides the most accurate forecasts