Applied Econometrics – Introduction to Time Series

Roman Horváth

Lecture 2

Contents

- Stationarity
  - What it is and what it is for
- Some basic time series models
  - Autoregressive (AR)
  - Moving average (MA)
- Consequences of non-stationarity (spurious regression)
- Testing for (non)-stationarity
  - Dickey-Fuller test
  - Augmented Dickey-Fuller test

(Weak) Stationarity

- **X_t is stationary if:**
  - the series fluctuates around a constant long run mean
  - X_t has finite variance which is not dependent upon time
  - Covariance between two values of X_t depends only on the difference apart in time (e.g. covariance between x_t and x_{t-1} is the same as for x_{t-8} and x_{t-9})

$$\mathbf{E}(X_t) = \mu \quad \text{(mean is constant in } t\text{)}$$
$$\mathbf{Var}(X_t) = \sigma^2 \quad \text{(variance is constant in } t\text{)}$$
$$\mathbf{Cov}(X_t, X_{t+k}) = \chi(k) \quad \text{(covariance is constant in } t\text{)}$$

- ***If data not stationary, spurious regression problem***

Examples of Times Series Models

- AR – autoregressive models
  - $X_t = \beta + \alpha{*}X_{t-1} + u_t$ ……is called AR(1) process
  - $X_t = \beta + \alpha_1{*}X_{t-1} + \alpha_2{*}X_{t-2} +\ldots.+\alpha_k{*}X_{t-k} + u_t$ …is AR(k) process

- MA – moving average models
  - $X_t = \beta + u_t + \alpha_2{*}u_{t-1}$ ……is called MA(1) process
  - $X_t = \beta + u_t + \alpha_2{*}u_{t-1} +\ldots.+\alpha_k{*}u_{t-k}$ ….is called MA(k) process

- If you combine AR and MA process, you get ARMA process
  - E.g. ARMA (1,1) is $X_t = \beta +\alpha{*}X_{t-1}+u_t+ \alpha_2\, u_{t-1}$

Is MA and AR process stationary?

- Compute mean, variance and covariance and check if it depends on time
- For AR process, you may easily derive that the process is stationary if $|\alpha| < 1$
- For MA(1) process, mean is β, variance is $u_t^2 *(1+\alpha_2^2)$ and covariance $\mathrm{cov}(\boldsymbol{X_t}, \boldsymbol{X_{t-k}})$ is either 0 if k>1 or $u_t^2 * \alpha_2$,so it does not depend on time (MA(k) is stationary process)

MA(1) process

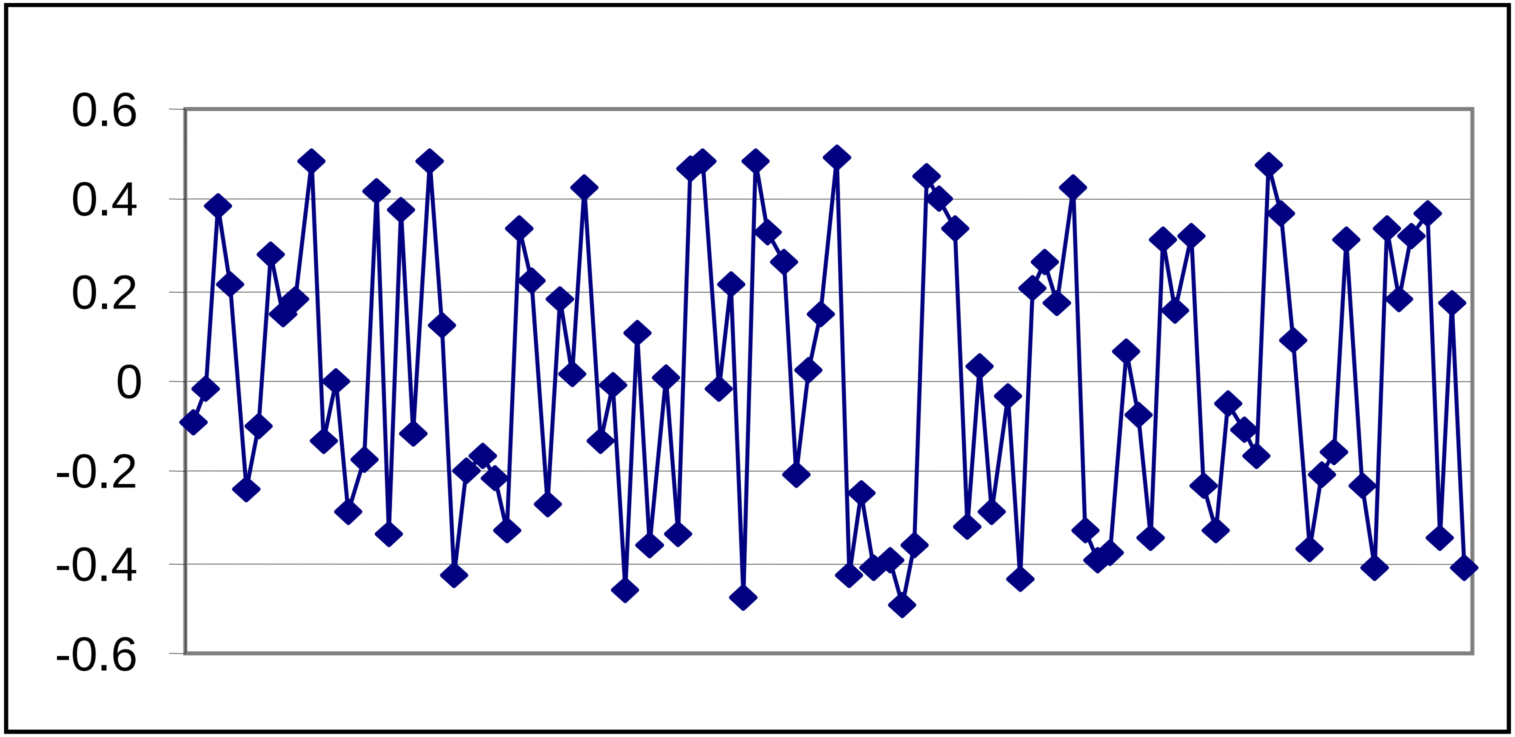
- $\mathbf{X_t = \beta + u_t + \alpha_1 * u_{t-1}}$

- $\mathbf{E(X_t)} = \beta$
- $\mathbf{Var(X_t)} = E(\beta + u_t + \alpha_1 * u_{t-1} - \beta)^2 =$
  $= E(u_t^2 + 2\alpha_1 u_t u_{t-1} + \alpha_1^2 u_{t-1}^2) = u_t^2(1 + \alpha_1^2)$
- $\mathbf{Cov(X_t, X_{t-k})} = E[(X_t - E(X_t))(X_{t-k} - E(X_{t-k}))] = u_t^{2*}\alpha_1$,
  if $k = 1$ and 0, if $k > 1$.
- $\mathbf{Corr(X_t, X_{t-k})} = Cov(X_t, X_{t-k}) / Var(X_t)^{1/2} * Var(X_{t-k})^{1/2}$

Example of Stationary Time Series

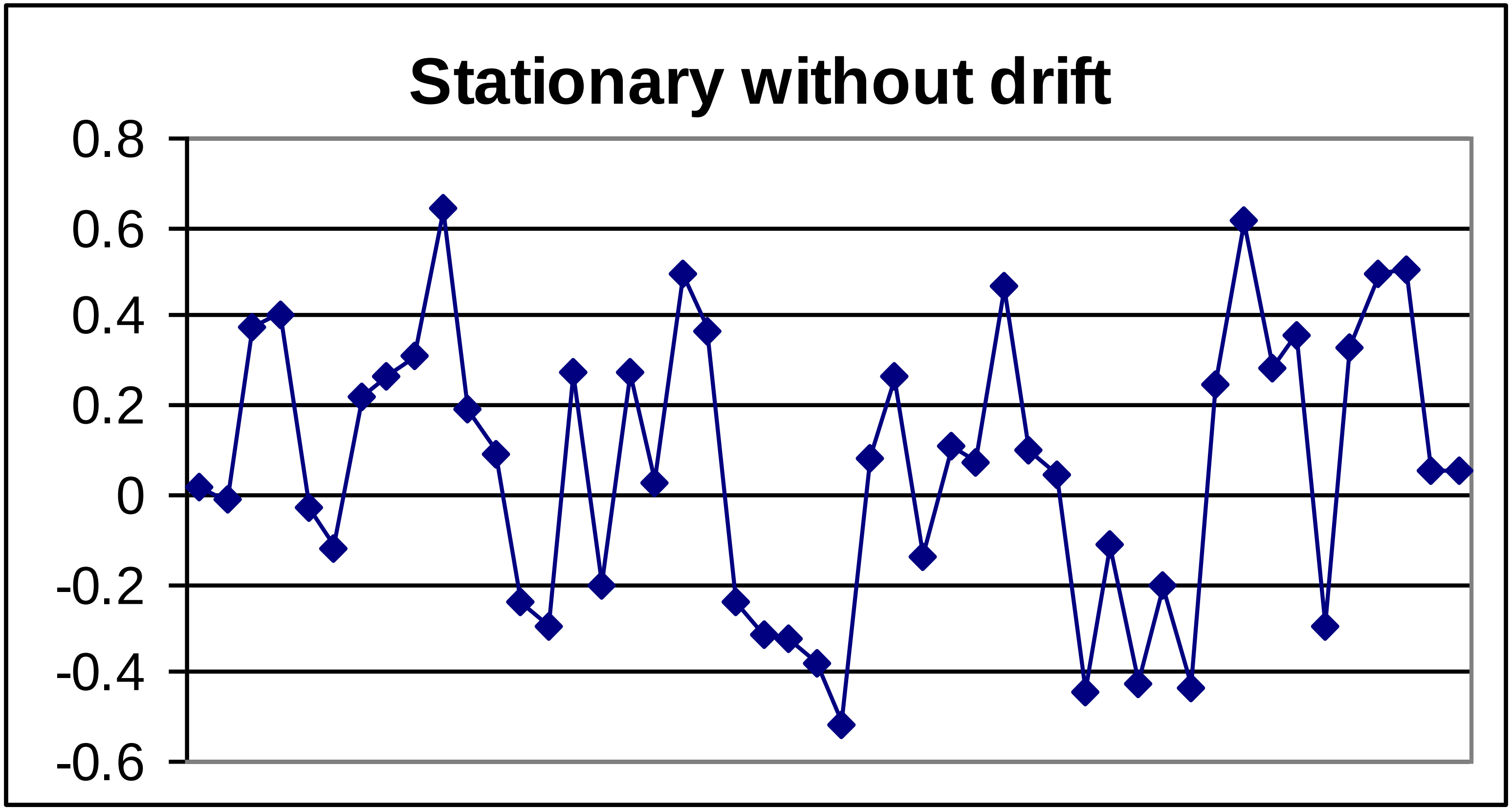
- White noise process:

$$X_t = u_t \qquad u_t \sim IID(0, \sigma^2)$$



Another Example of Stationary Time Series

$$X_t = 0.5 * X_{t-1} + u_t \qquad u_t \sim IID(0, \sigma^2)$$



Example of Non-stationary Time Series

- $\boldsymbol{Y_t = \alpha + \beta^* t + u_t}$ **,** where ***t*** is time trend

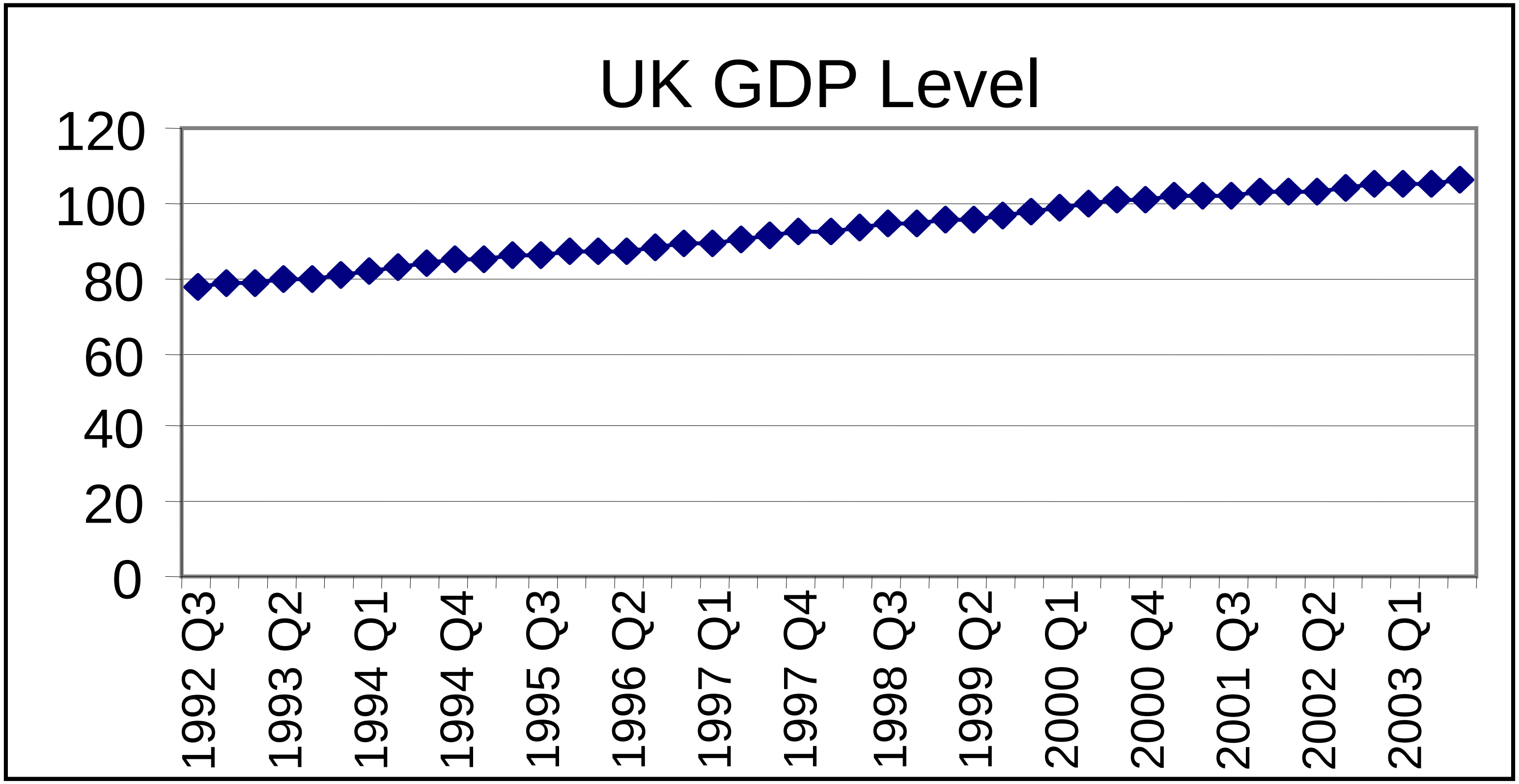
- Take the expected value $E(Y_t) = \alpha + \beta^* t$, clearly the mean depends on time and the series is non-stationary

Non-stationary time series

In contrast a non-stationary time series has at least one of the following characteristics:

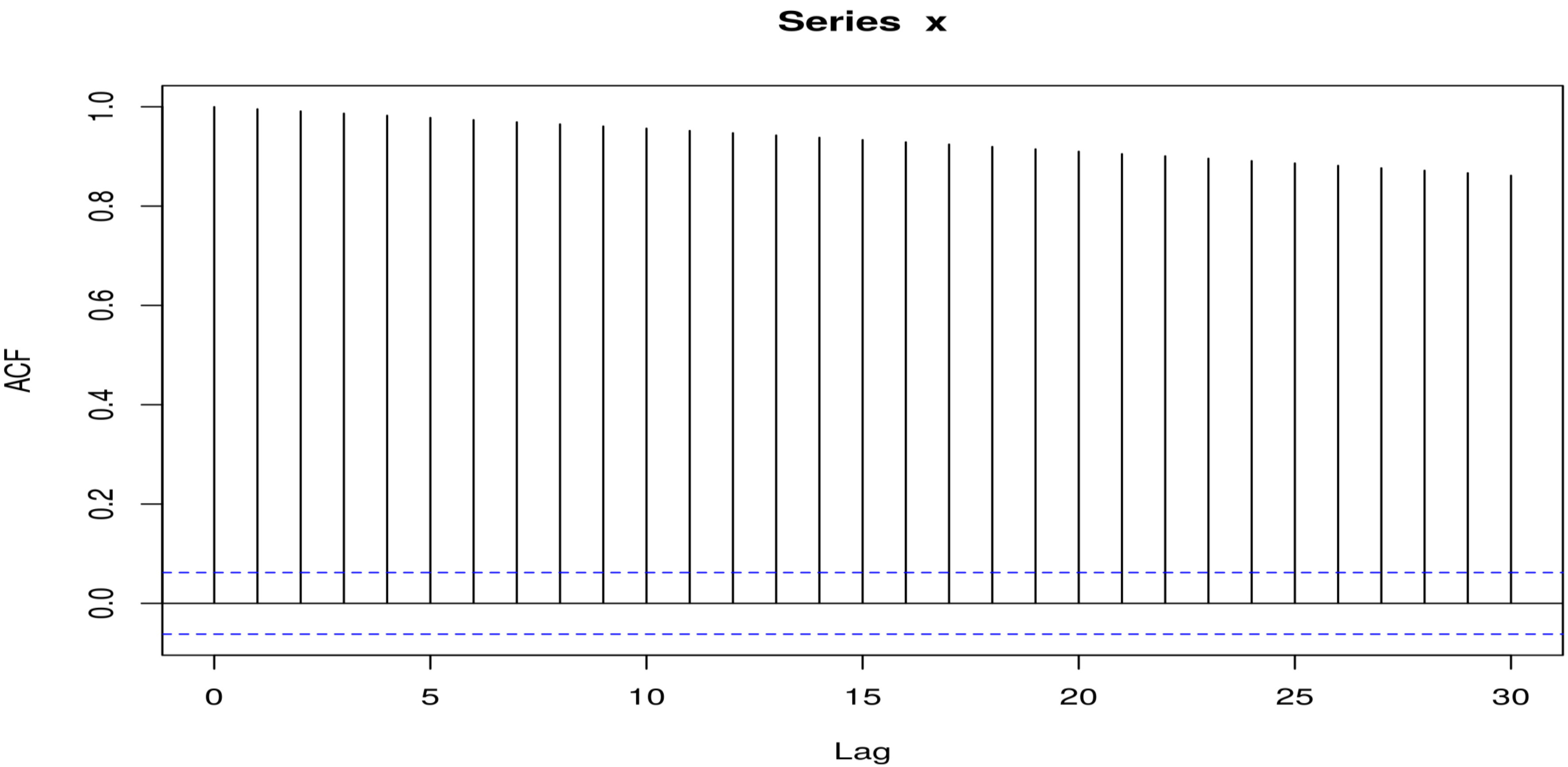
- Does not have a long run mean which the series returns
- Variance is dependent upon time and goes to infinity as the sample period approaches infinity
- Correlogram does not die out - long memory

Example of Non-stationary Time Series



- The level of GDP is not constant; the mean increases over time.

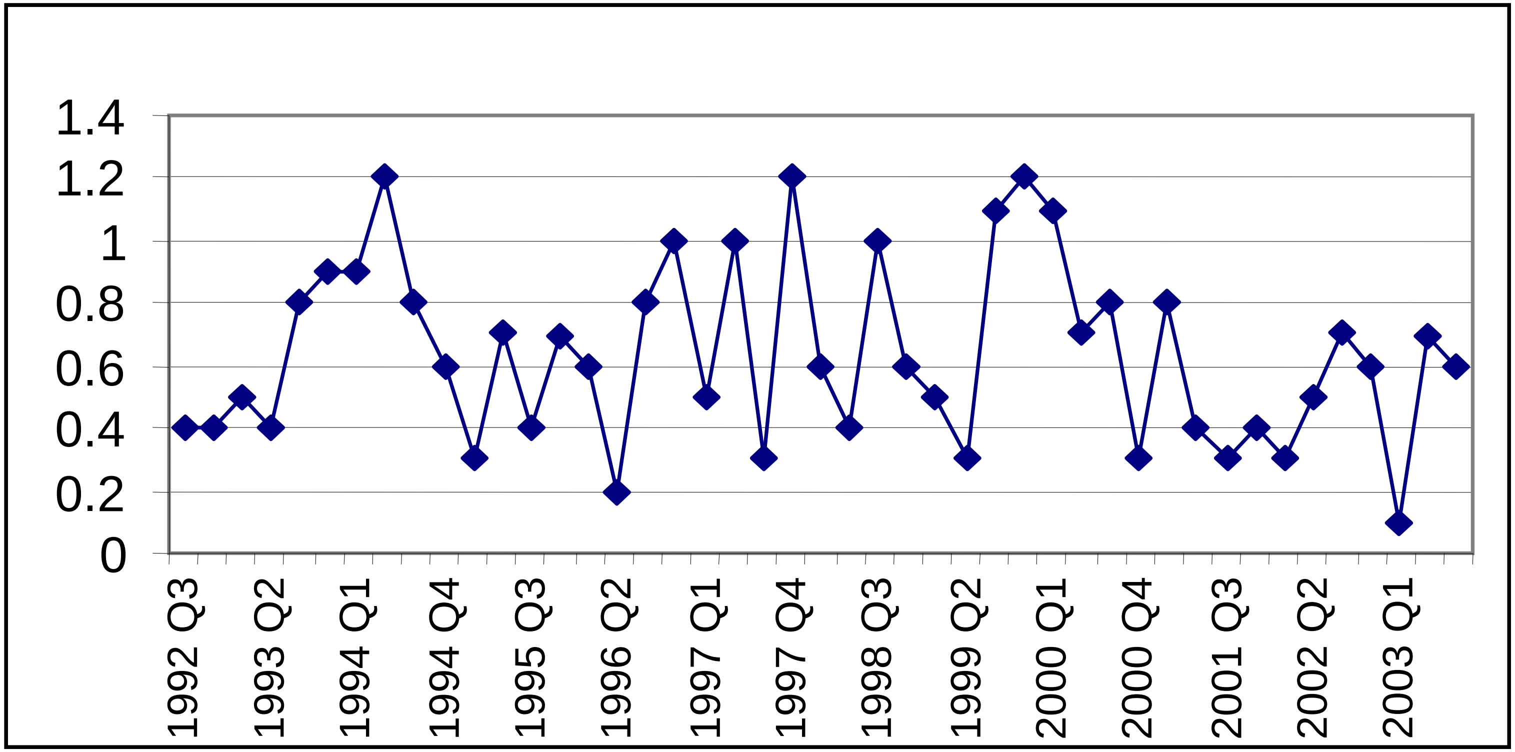
Non-stationary time series – correlogram



- For non-stationary series the Autocorrelation Function (ACF) declines towards zero at a very slow rate as *k* increases (or does not decline at all).

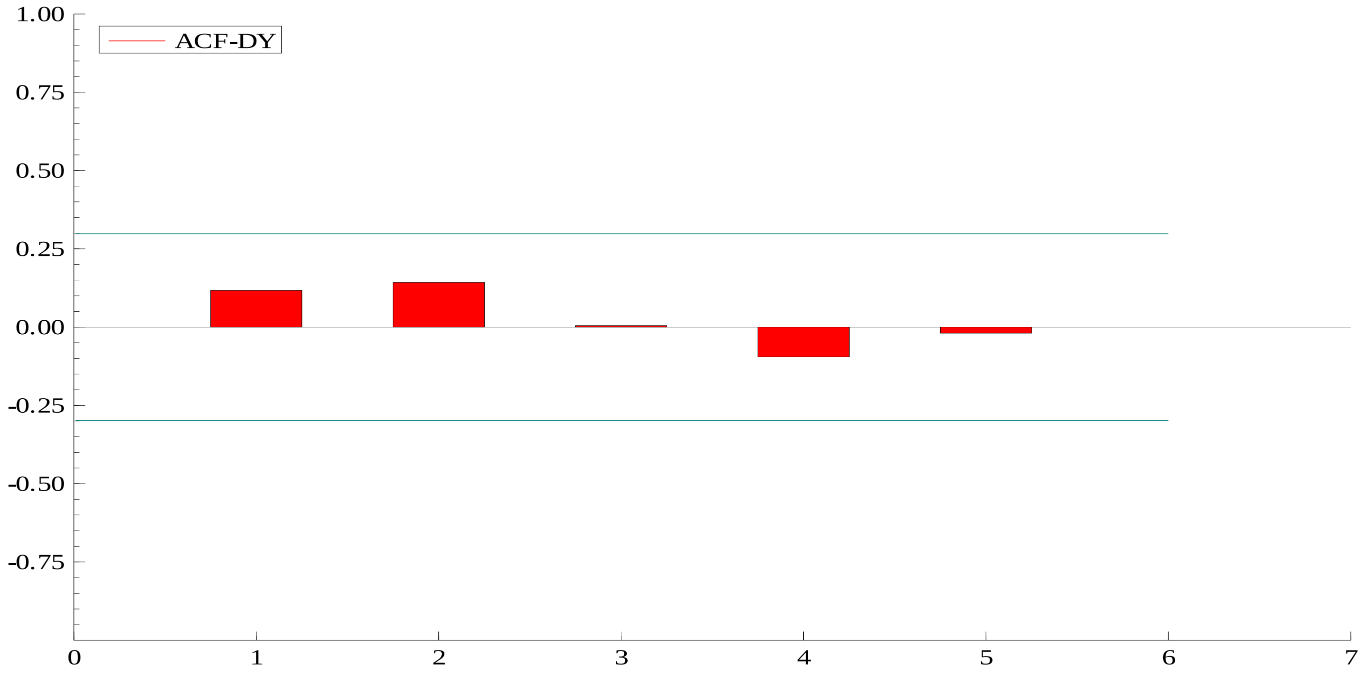
Possible solutions of non-stationarity

- Some transformation = first differenece, logarithm, second difference ...
- First difference of UK GDP ($\Delta Y_t = (Y_t - Y_{t-1})/Y_{t-1}$) is stationary:
  - growth rate is reasonably constant through time
  - variance is also reasonably constant through time



Stationary time series - correlogram

UK GDP Growth (ΔY_t)



- ACF decline towards zero as k increases
- Decline of ACF is rapid for stationary series

Non-stationary Time Series Continued – Random Walk

- $\boldsymbol{X_t = X_{t-1} + u_t}$**, where** $\boldsymbol{u_t \sim IID(0, \sigma^2)}$
- Mean is constant in t: $E(X_t) = E(X_{t-1})$

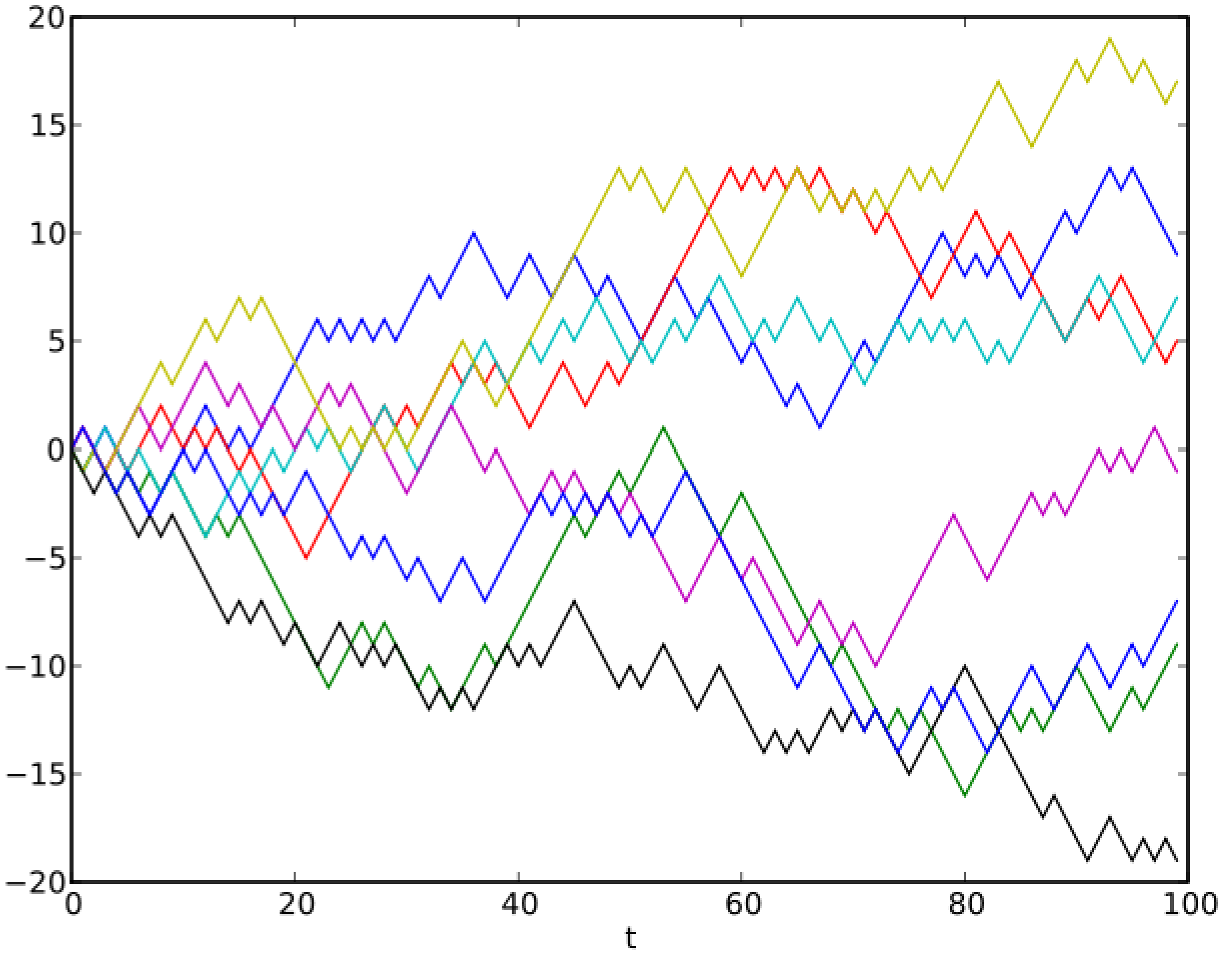
$$X_1 = X_0 + u_1 \qquad \text{(take initial value } X_0\text{)}$$
$$X_2 = X_1 + u_2 = (X_0 + u_1) + u_2$$
$$\ldots$$
$$X_t = X_0 + u_1 + u_2 + \ldots + u_t \text{ (take expectations)}$$
$$E(X_t) = E(X_0 + u_1 + u_2 + \ldots + u_t) = E(X_0) = \text{constant}$$

- Variance is not constant in t:

$$\text{Var}(X_t) = \text{Var}(X_0) + Var(u_1) + \ldots + Var(u_t)$$
$$= 0 + \sigma^2 + \ldots + \sigma^2 = t\,\sigma^2$$

Random walk

- $X_t = X_{t-1} + u_t \qquad\qquad u_t \sim IID(0, \sigma^2)$



Relationship between stationary and non-stationary process

- **AR(1) process**: $X_t = \beta + \alpha X_{t-1} + u_t$;$(u_t \sim IID(0, \sigma^2))$

  $|\alpha| < 1$stationary process - "process forgets its past"

  otherwise non-stationary process - "process does not forget its past"

  $\beta = 0$ without drift (constant)

  $\beta \neq 0$ with drift

  AR(k) analogous to AR(1), sum of α's instead of α

- **MA process** is always stationary

Summary on basic time series processes

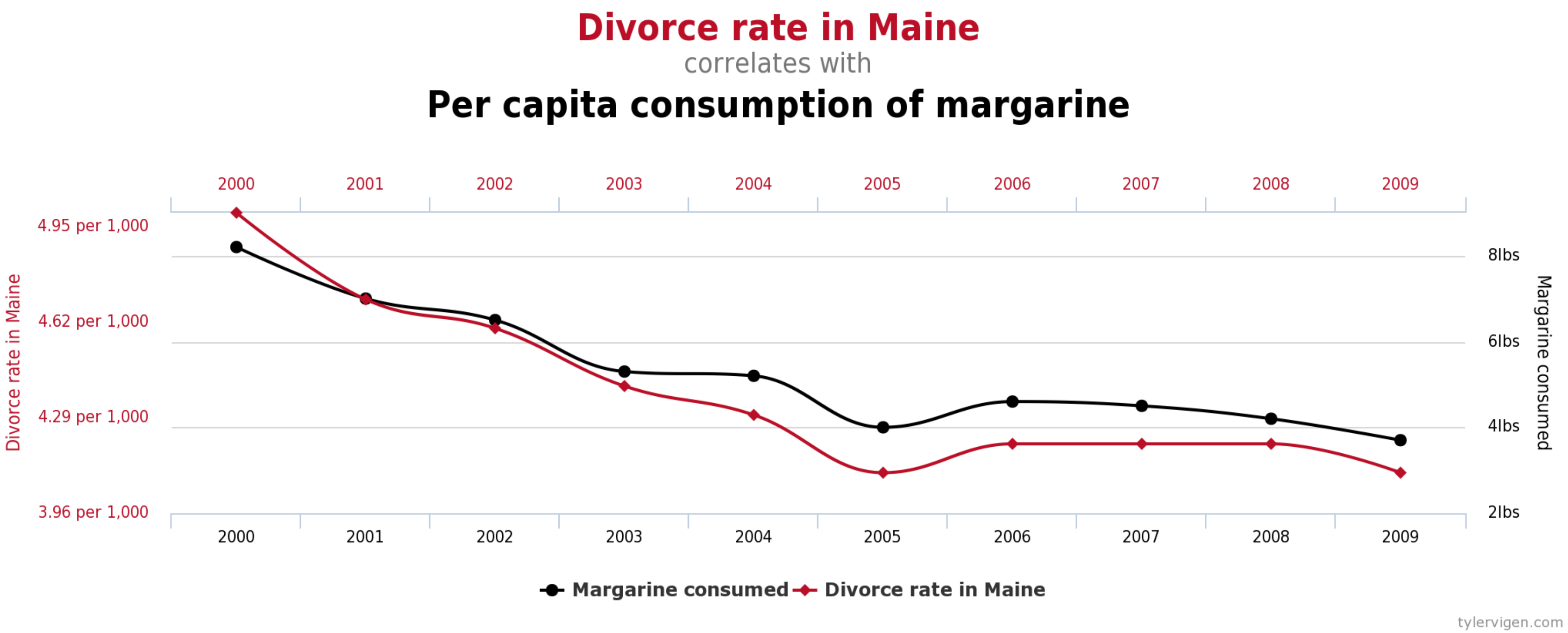
- AR (k) process
- MA(k) process
- ARMA (p,l) process
- If you k-th difference the data, then you have ARIMA (p,k,l) – estimation of ARIMA models is a subject of next lecture

Spurious Regression (≡ spurious correlation)

- Problem that time-series data usually includes trend
- Result:
  - Spurious correlation (variables with similar trends are correlated)
  - Spurious regression (independent variable with similar trend looks as dependent = strong statistical relationship)

    ⇒ coefficient significant (high adjusted-R^2, large t-statistics) ... even if unrelated in economic terms

Spurious Correlation: Example



How to avoid spurious regression: 3 approaches to non-stationarity

1. Include a **time trend** as an independent variable (old-fashioned) $y_t = c + \beta_1 x_t + \beta_2 t + u_t$ .....(t = 1,2, ..., T)

2. 1st **difference** the data if variables I(1); 2nd difference if I(2)

   = converts non-stationary variables into stationary variables

   Problems:

   – theory often about levels
   – detrending $\Rightarrow$ loss of information

3. Cointegration + ECM

= Long-run relationship **+** short-run adjustment

How do we identify non-stationary processes?

(A) Informal methods:

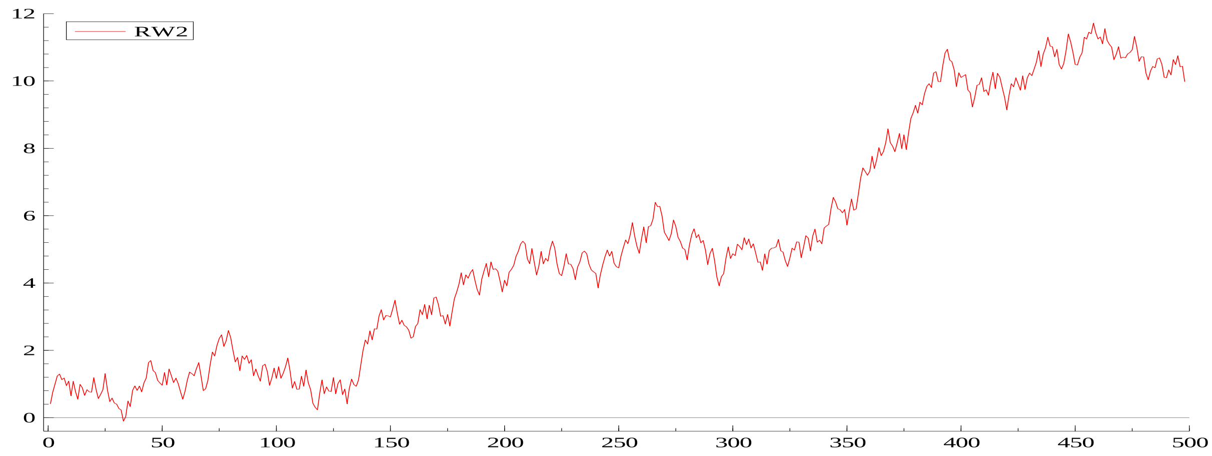
- Plot time series
- Correlogram

(B) Formal methods:

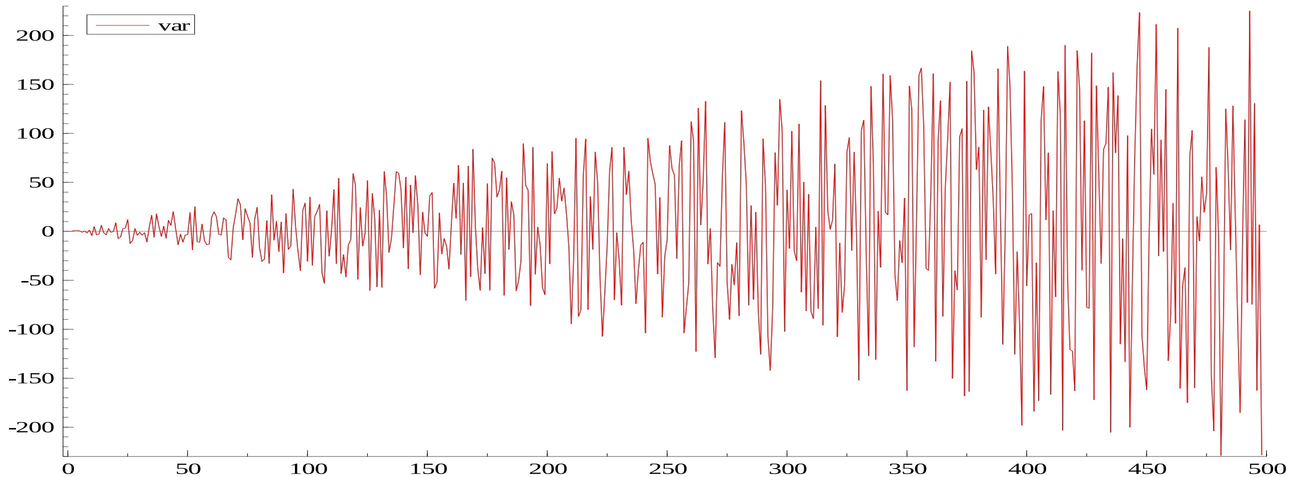
- Statistical test for stationarity
- Dickey-Fuller tests

Informal Procedures to identify non-stationary processes

(a) Constant mean?

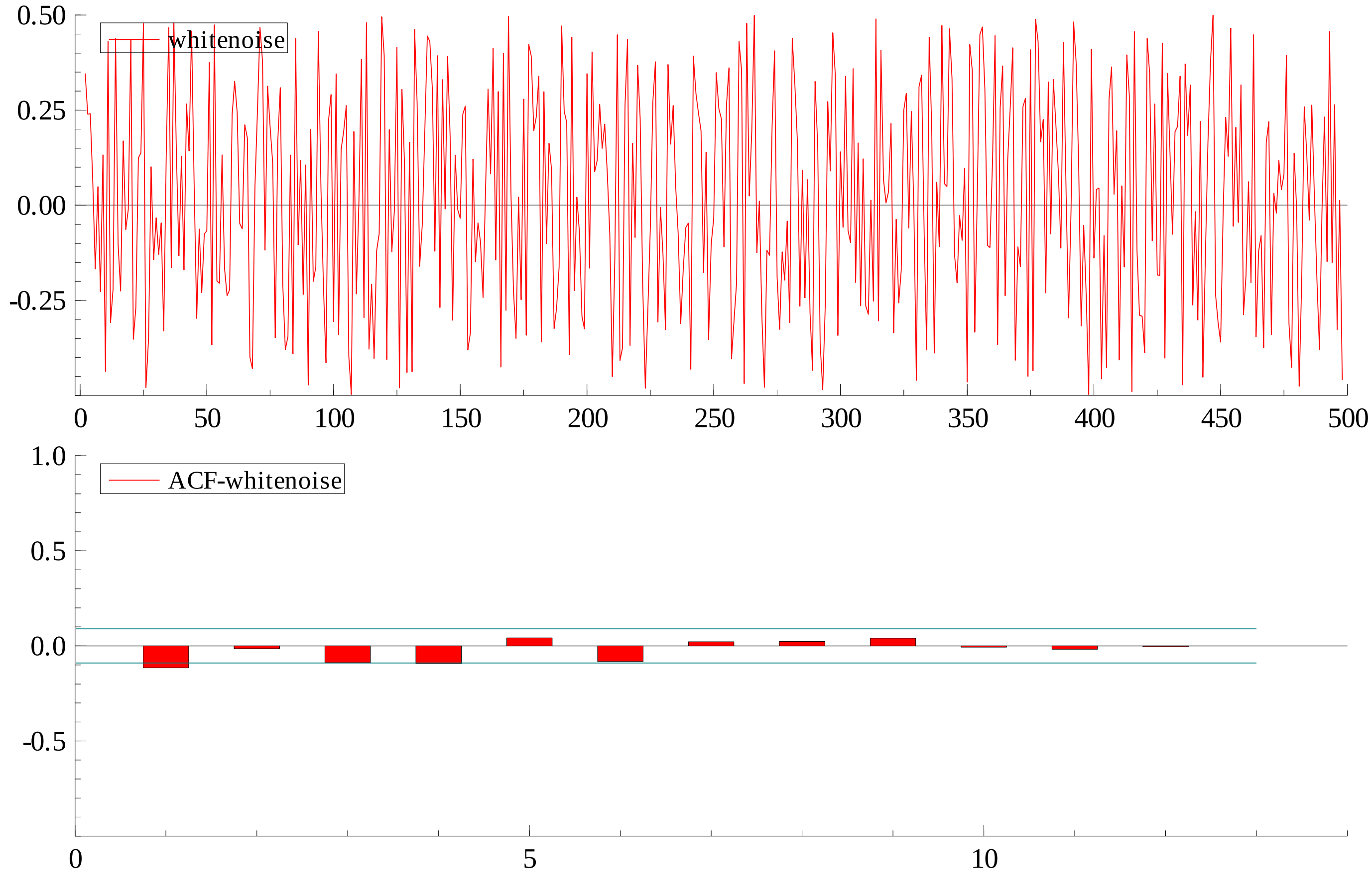


(b) Constant variance?



Informal Procedures to identify non-stationary processes

- Diagnostic test – Correlogram for stationary process (dies out rapidly, series has no memory)



Informal Procedures to identify non-stationary processes

- Diagnostic test – Correlogram for a random walk (does not die out, high autocorrelation for large values of k)

Dickey-Fuller Test

- Test based on $Y_t = \alpha Y_{t-1} + u_t$
  - DF test to determine whether $\alpha=1$
    - Yes $\Rightarrow$ unit root $\Rightarrow$ non-stationary
    - No $\Rightarrow$ no unit root

- Dynamic model:
  – Subtract Y_{t-1} … $Y_t - Y_{t-1} = (\alpha-1)Y_{t-1} + u_t$
  – Reparameterise: $\Delta Y_t = \beta Y_{t-1} + u_t$
    where $\beta = (\alpha-1)$
  – Test $\beta=0$ equivalent to test $\alpha=1$

Augmented Dickey-Fuller Test

- Augment „dynamic model“ $\Delta Y_t = \beta Y_{t-1} + u_t$:

  1) Constant or “drift” term (α_0)

  $$\Delta Y_t = \alpha_0 + \beta Y_{t-1} + u_t$$

  2) Time trend (T)

  $$\Delta Y_t = \alpha_0 + \gamma T + \beta Y_{t-1} + u_t$$

  3) Lagged values of the dependent variable

  $$\Delta Y_t = \alpha_0 + \gamma T + \beta Y_{t-1} + \delta_1 \Delta Y_{t-1} + \delta_2 \Delta Y_{t-2} + ... + u_t$$

- Find the right specification, trade-off parsimony vs. white noise in residual

Critical values

- DF/ADF use t-statistics but critical values are not standard
- Problems:
    - distributions of these statistics are non-standard
    - special tables of critical values (derived from numerical simulations)
    - Usual t- and F-tests not valid in presence of unit roots
- The null hypothesis and its alternative „reversed“