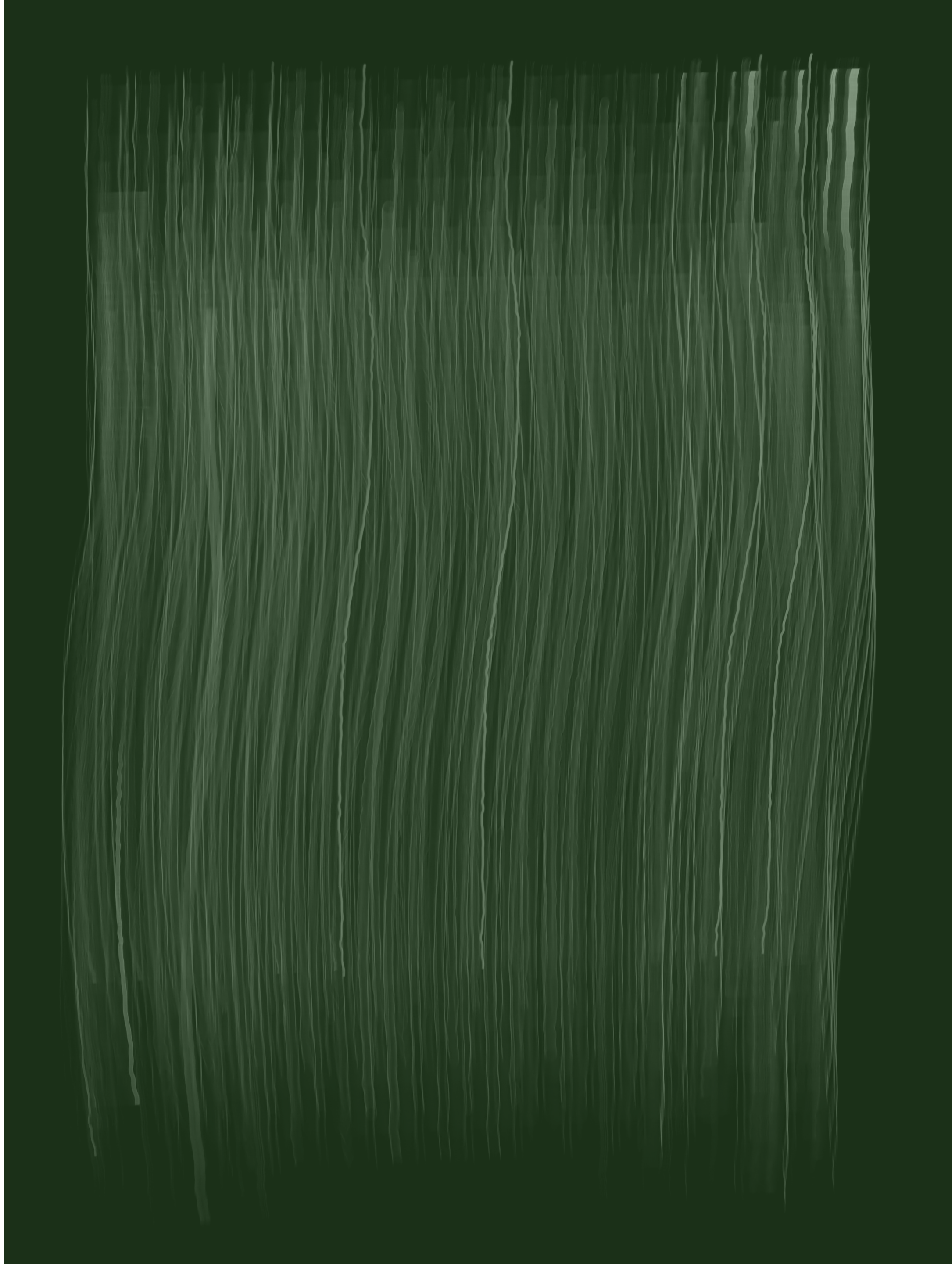
Econometrics I

Workshop X

May 9, 2023

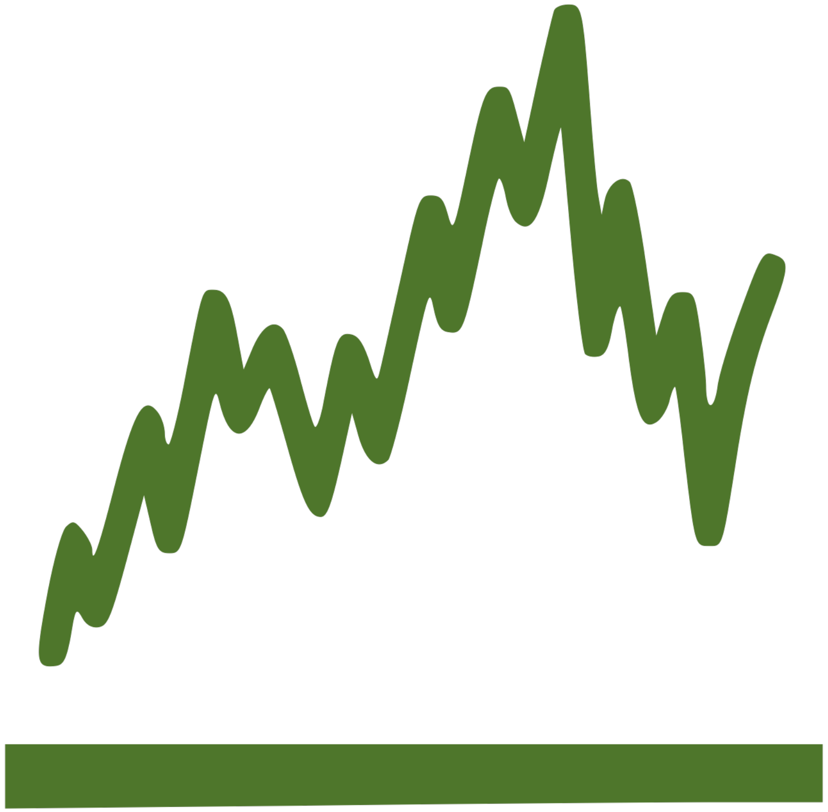
Time Series



What is a Time Series?

A set of measurements of a certain phenomenon or experiment recorded sequentially over time.

These can be: continuous (an electrocardiogram) or discrete (the GDP).



Time Series Applications

- Mathematic techniques
  - Statistics
    Artificial Intelligence
    Microareas
    - Regression
      - Simple regression
        - Linear
        - Non- Linear
      - Multiple regression
        - Linear
        - Non- Linear
      - Econometric (Multivariate linear regression using economic variables)
    - Time Series
      - Decomposition
        AR, MA, ARMA
      - Exponential Smoothing
        - Simple
        - Holt
        - Winter
    - Probability Distribution

Stochastic or random processes

Stochastic process

STATIONARY stochastic processes

NON-STATIONARY stochastic processes

Stochastic or random processes

Stochastic process

A collection of random variables ordered in time

STATIONARY stochastic processes

A stochastic process is stationary if its mean and variance are constant over time.

A time series is stationary if its mean, variance, and autocovariance (at different lags) remain the same regardless of when they are measured; that is, they are time-invariant.

NON-STATIONARY stochastic processes

It will have a mean that varies over time or a variance that changes with time, or both.

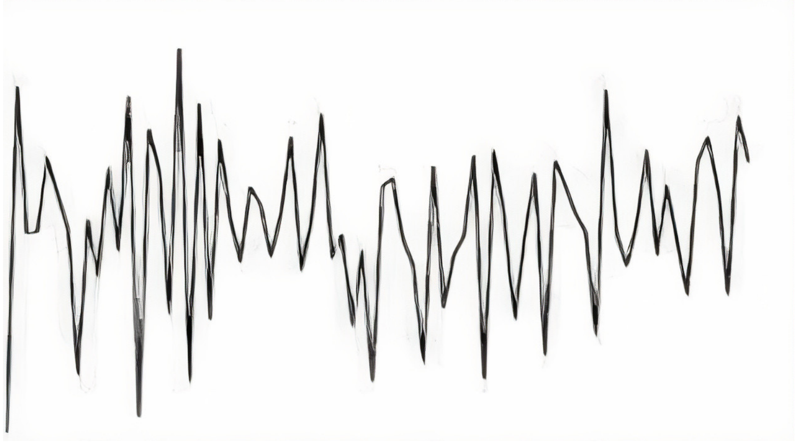
A classic example is the random walk model (RWM).

It is said that asset prices, such as stocks or exchange rates, follow a random walk. That is, they are NON-stationary.

Stochastic or random processes

Stochastic process

STATIONARY stochastic processes



NON-STATIONARY stochastic processes

White noise generation in STATA

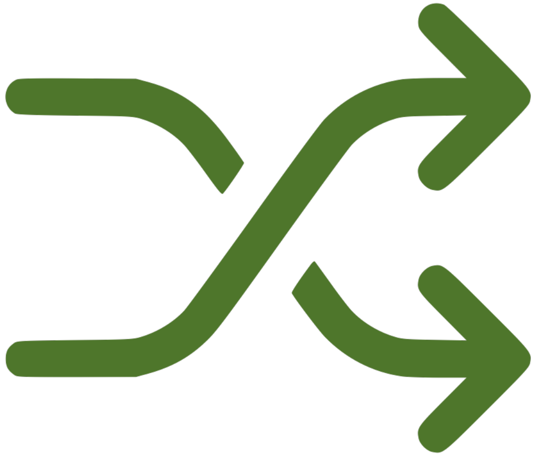
```
// set number of observations
set obs 276

// generate time
generate t=tm(1990m1)+_n-1

// set format
format t %tm

// set time series format
tsset t

// obtain graphs
drawnorm y1, n(276) means(0) sds(1)


twoway (line y1 t)

corrgram y1, lags(24)
```

Deterministic vs Stochastic

The distinction between stationary and non-stationary stochastic processes (or time series) is important for determining whether the TREND (the long-term evolution of the time series) is ________:

DETERMINISTIC

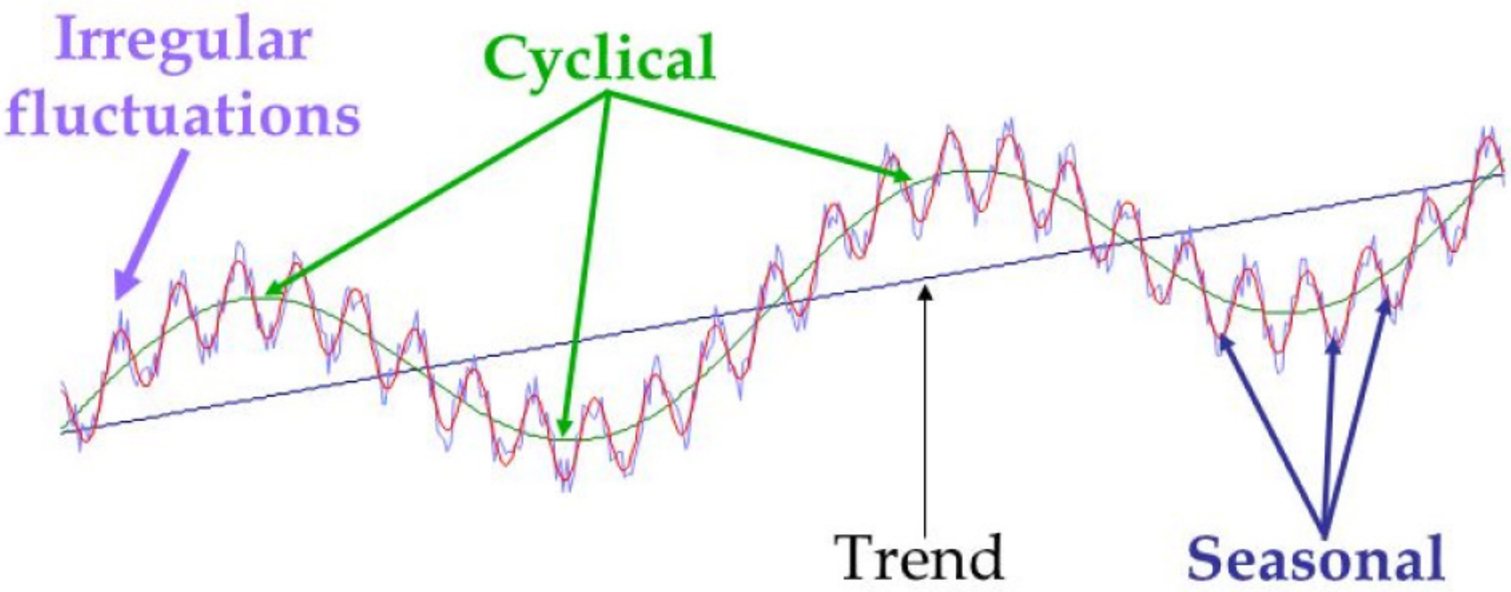
If the trend of a time series is entirely predictable and non-variable.



STOCHASTIC

The trend is not predictable.

Components of a Time Series



Trend. It can be defined as a long-term change in relation to the average level or the long-term change of the mean. The trend is identified by a smooth movement of the series over the long term.

Seasonality. The periodic fluctuation in time series within a given period. These fluctuations form a pattern that tends to repeat from one seasonal period to the next.

Cycles. Time series exhibit alternating sequences of points below and above the trend line. Long deviations from the trend due to factors other than seasonality.

Irregular movement. The movement that remains after accounting for trend, seasonal, and cyclical movements; random noise or error in a time series.

Components of a Time Series

Deterministic

Trend
Seasonality

Random

Irregular movement

Of these three components, the first two are deterministic, while the last one is random. Thus, the time series can be denoted as:

$$X_t = T_t + C_t + E_t + I_t$$

Where T is the trend, C is the cycle, E is the seasonal component, and I is the irregular or random component.

Components of a Time Series

Additive Model: Used when the magnitude of the seasonal fluctuations in the series does not vary as the trend does.

Multiplicative or Mixed Model: Used when the magnitude of the seasonal fluctuations in the series grows and/or decreases proportionally with the variations in the trend.

Additive

$$X_t = T_t + C_t + E_t + I_t$$

Multiplicative

$$X_t = T_t * C_t * E_t * I_t$$

Mixed

$$X_t = T_t * C_t * E_t + I_t$$

Trend models

Linear

$$T_t = \beta_0 + \beta_1 t$$

Quadratic

$$T_t = \beta_0 + \beta_1 t + \beta_2 t^2$$

Cubic

$$T_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3$$

Exponential

$$T_t = e^{\beta_0 + \beta_1 t}$$

Logistic

$$T_t = \frac{\beta_2}{1 + \beta_1 e^{-\beta_0 t}}$$

Autoregressive Processes AR(p)	Moving Average Process MA(q)	Autoregressive Moving Average Process ARMA(p,q)

Autoregressive models are based on the idea that the current value of the series, X_t, can be explained as a function of past values $X_{t-1}, X_{t-2}, X_{t-p}$ where P determines the number of lags needed to predict a current value. The autoregressive model of order is given by:

$$X_t = \Phi_0 + \Phi_1 X_{t-1} + \Phi_2 X_{t-2} + \cdots + \Phi_p X_{t-p} + \varepsilon_t$$

Autoregressive Processes AR(p)

Moving Average Process MA(q)

Autoregressive Moving Average Process ARMA(p,q)

It is determined by an external source. These models assume linearity, where the current value of the series, , is influenced by the values of the external source. The moving average model of order q is given by:

$$X_t = \Theta_0 - \Theta_1 X_{\varepsilon-1} - \Theta_2 X_{\varepsilon-2} - \cdots - \Theta_q X_{\varepsilon-q} - \varepsilon_t$$

Stationary Linear Processes

Autoregressive Processes AR(p)

Moving Average Process MA(q)

Autoregressive Moving Average Process ARMA(p,q)

It is very likely that a time series, Xt, has characteristics of both AR and MA simultaneously, and therefore, is ARMA. Thus, X_t follows an ARMA (p, q) process, in which there will be p autoregressive terms and q moving average terms.

$$X_t = c + \underbrace{\Phi_1 X_{t-1} + \cdots + \Phi_p X_{t-p}}_{AR(p)} + \underbrace{\Theta_1 X_{\varepsilon-1} + \Theta_2 X_{\varepsilon-2} + \cdots + \Theta_q X_{\varepsilon-q} + \varepsilon_t}_{MA(q)}$$

Autoregressive Integrated Moving Average Process ARIMA (p,d,q)

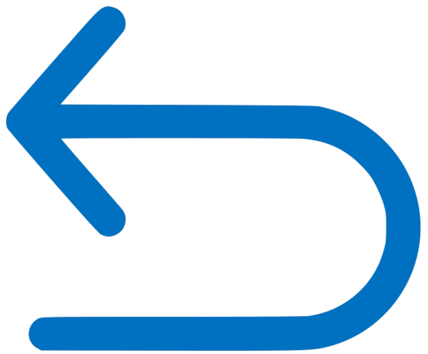
The time series models analyzed so far are based on the assumption of stationarity.

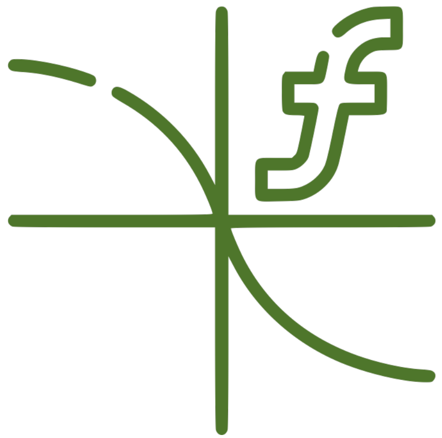
However, it is known that many time series, especially economic series, are not stationary because they may change levels over time or simply because the variance is not constant over time. These types of processes are considered integrated processes.

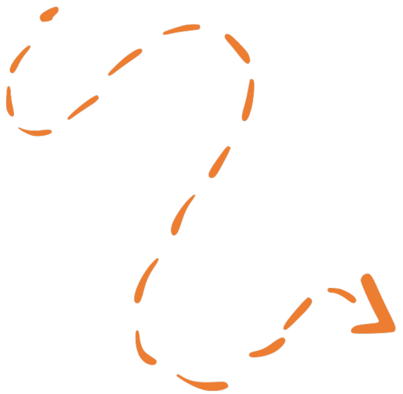
Therefore, a time series must be differentiated d times to make it stationary, and then an $ARMA(p,q)$ model can be applied to this differentiated series. It is said that the original series is $ARIMA(p,d,q)$ that is, an autoregressive integrated moving average time series.

Non-Stationary Linear Processes

$$ARIMA(p, d, q)$$







p denotes the number of autoregressive terms

q denotes the number of times the series must be differentiated to make it stationary

d denotes the number of invertible moving average terms

Sometimes in a time series, the values that a variable takes over time are not independent of each other, but a specific value depends on previous values. There are two graphical ways to measure this dependence of variables.

Autocorrelation function (ACF) and correlogram Autocorrelation

measures the correlation between two variables separated by k periods.

Partial Autocorrelation Function (PACF)

measures the correlation between two variables separated by k periods when the dependence created by the intermediate lags between them is not considered.

Simulation of AR(1) process

```
// set number of observations
set obs 276

// generate time
generate t=tm(1990m1)+_n-1

// set format
format t %tm

// set time series format
tsset t

// draw sample
drawnorm u1, n(276) means(0) sds(1)

// gen y
gen y=.
replace y = 0 in 1
replace y = 0.9*l.y+u1 in 2/276

// graph
twoway (line y t)
```

Time Series Decomposition

```
// Hodrick-Prescott filter for cycle and trend decomposition
tsfilter hp wpi_ciclo = wpi, trend(wpi_tendencia)smooth(16000)

// Smoothing parameter
// Lambda = 1600(trimestral)129600(mensual)6.25(anual)

// Smoothing through Moving Averages
tssmooth ma wpi_suavizada=wpi, window(4)
```

Stages of time series analysis

1. **Identification**
2. **Estimation**
3. **Verification**
4. **Forecasting**

The time series technique (ARIMA models), although known for a long time, was not systematized until the late 1960s by G.E.P. Box and G.M. Jenkins.

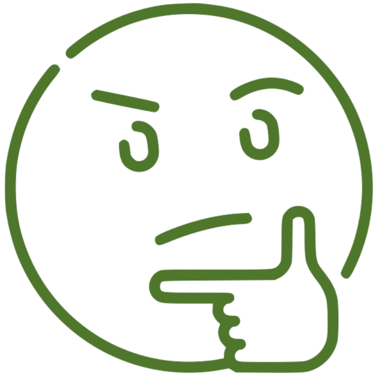
If the series is weakly stationary, stage (i) is immediately proceeded; otherwise, the series must be "pre-processed" in order to be transformed into stationary realizations.

Identification

Estimation & verification

Forecasting

Graphically represent the series, in addition to its autocorrelation function (ACF) and partial autocorrelation function (PACF).



The series plot indicates whether the series is stationary or not. Depending on the reasons why the series is not stationary, we will have to apply the following procedures until it becomes stationary.

🚨 *Note: If it has a trend we will take regular differences until it disappears. Normally, the order of difference is 1, and rarely will it be greater than 3.*

Identification

Estimation & verification

Forecasting

Assuming that we have stationary series, the identification aims to:

Determine the type of model to apply (AR, MA, or ARMA) and the order of the parameters "p" and "q".

The parameters p, d, and q denote:

1. The order of the autoregressive part of the process,

2. The degree of differentiation required to transform a non-stationary series into a stationary one,

3. The order of the moving average part of the process, respectively.

Protocol for identifying ARIMA models (according to Box-Jenkins steps)

Identification

Estimation & verification

Forecasting

By observing the two graphs of the ACF and PACF of the transformed series, we can get an idea of the model that describes our series, or at least which are the first candidates we should test.

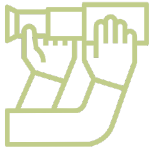
To analytically (not visually) verify a model, several candidate ARIMA(p,d,q) models are often fitted, and we will choose as a good model:

- The one with residuals similar to white noise,
- Lower values of AIC (Akaike Information Criterion) and
- BIC (Bayesian Information Criterion) compared to the rest of the candidate models.

Protocol for identifying ARIMA models (according to Box-Jenkins steps)

Identification

Estimation & verification



Forecasting

One of the reasons for the popularity of the ARIMA model-building process is its success in prediction. ARIMA models are good for making short-term predictions.