Series Temporales
Time Series

Alejandre Cabaña

Universitat Autònoma de Barcelona

GEA - GMAT 2020-2021

1er Quadrimestre

Time Series

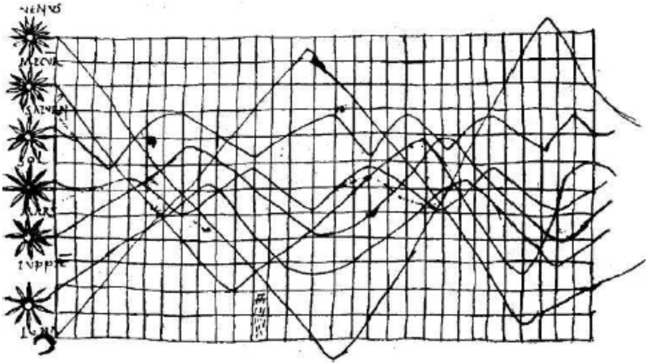


Figure 2: Planetary movements shown as cyclic inclinations over time, by an unknown astronomer, appearing in a 10^{th} century appendix to commentaries by A. T. Macrobius on Cicero's *In Somnium Scripionus*. Source: Funkhouser (1936, p. 261).

Introduction to Time Series Analysis

In Statistics and many other fields, a time series

$$x_{t_1}, x_{t_2}, \ldots$$

is a sequence of data points coming from a phenomenon that evolves over *time*[1] measured over (often uniform) intervals.

Applications

- Economics: monthly data for unemployment, hospital admissions, ...
- Finance: daily exchange rates, share prices and benefits...
- Environment: daily rainfall, air quality readings, ozone layer measurements...
- Medicine: brain wave activity, oxygen (or drugs) concentration in blood, daily number of CoVID-19 cases in a region...
- $\vdots$

[1] or space, or any other ordered magnitude

Technically speaking, by *time* we mean that the observations are recorded sequentially, but it might be spatial (1st plant in a row, 2nd plant in a row ...) or depth (1 cm down, 2 cm down,...) or any other **ordered** independent variable.

There can be univariate or multivariate observations, **numerical** (real numbers or integers) or categorical.

The observations can be taken in a **discrete** set $t_1, t_2, t_3, \ldots$ or continuously over some interval.

Recording times might be regular (hourly, monthly, ...) or irregular due to missing values or the nature of the data

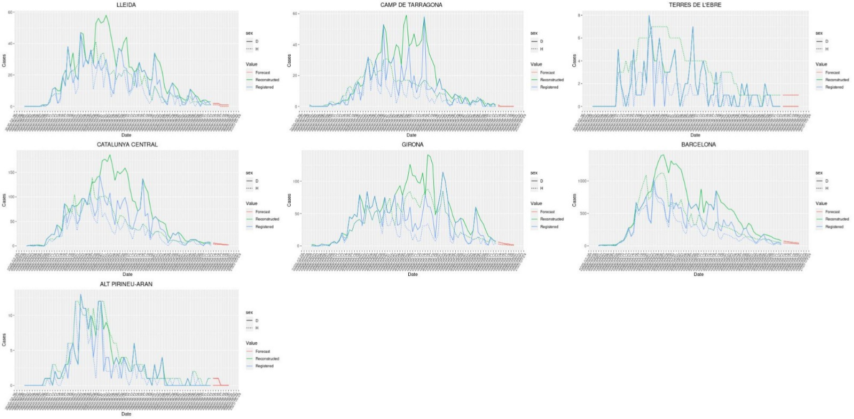
- Most bibliography and also most statistical packages deal with the classic time series, where the $x(t_i)$ are continuous random variables normally distributed; but the observations can be discrete (number of CoVID-19 cases)or categorical random variables, and different models shall be needed.
- In this course we will mostly concentrate in classical time series, studied through the basic Box and Jenkins Methodology.
- We will comment on simple forms of AR models for count time series data.
- Markov chain models are frequently used to describe categorical time series, that is, when the random variables take values in a countable set of possible states, and are dealt with in Stochastic processes courses.

Examples of time series data

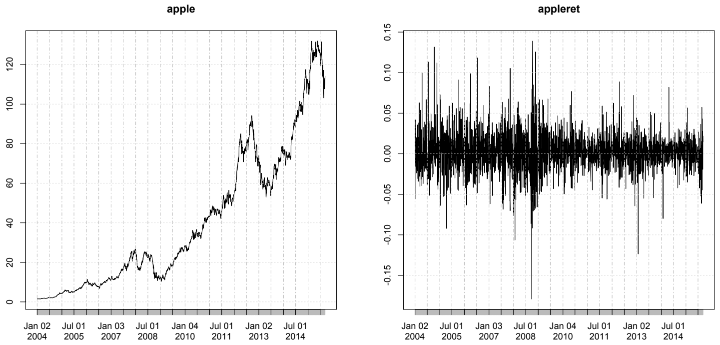
Monthly sales (in liters) of red wine by Australian winemakers from January 1980 to October 1991.

- Upward trend.
- Seasonal pattern with a peak in July, and a tough in January.

Number of CoVID-19 cases in Catalunya, by sanitary region



Apple Stock Prize and Rturns



Objectives of Time Series Analysis

What do we hope to achieve with time series analysis? As we have just seen, the examples can show very different behaviours. But in all cases, we need to:

1. Provide an interpretable model of the data
   - often involves a multivariate series
   - allows testing of scientific hypotheses
   - allows formulation of inverse problems for parameter determination
   - often not heavily emphasised in TSA
   - models depend on unknown parameters, so a main issue will be estimation
2. Predict future values
   - very common goal of analysis
   - predictive models often do not "explain"
3. Produce a compact description of the data: data compression in telecommunications

Modeling

We take the approach that the data are a realisation of random variables. However, it is clear that these observations are not I.I.D. and we rather interpret the time series as a realisation of a **stochastic process** $X(\omega, t)$ defined on $\Omega \times [a, b]$.

- For a fixed t we observe a random variable

- For a fixed ω_f we observe a sample path (a trajectory) or rather, the evaluations of $X(\omega_f, t_i)$ for a collection of time points t_i.

This kind of random processes $X_1, X_2, \ldots$ are determined by the specification of all the joint distributions of the random vectors $X_1, X_2, \ldots X_n$ for $n = 1, 2, \ldots$,or, equivalently, the probabilities

$$\mathbf{P}(X_1 \leq x_1, \ldots, X_n \leq x_n) \qquad -\infty < x_1, \ldots, x_n < \infty, \qquad n = 1, 2, \ldots$$

This plan is often too complicated to accomplish, hence second-order properties will be used. We shall see that for many intersting purposes, this simplification poses no limitations.

References

P.J. Brockwell and R.A. Davis: *Introduction to Time Series and Forecasting*. 2nd edit. Springer, 2002.

J.D. Cryer and K.S. Chan: *Time Series Analysis with Applications in R*. 2nd. edit. Springer, 2008

R.H. Shumway, and D.S. Stoffer: *Time Series Analysis and its Applications*. 3rd. edit. Springer,2011.

R. Tsay *Analysis of Financial Time Series*, 3rd Edition, Wiley 2010

C. Weiss,*An Introduction to Discrete-Valued Time Series*, Wiley (2018)