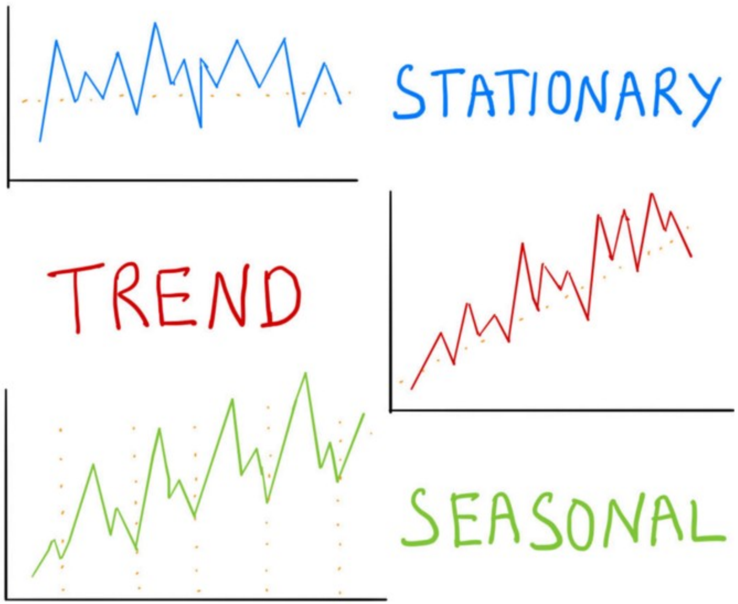
Time Series Analysis



Uriel Corona-Bermúdez

Material

References:

«Time Series Analysis and Its Applications With R Examples», Shumway, R. H. & Stoffer D. S

Evaluation:

Exam - 30%

Practices - 40%

Project - 30%

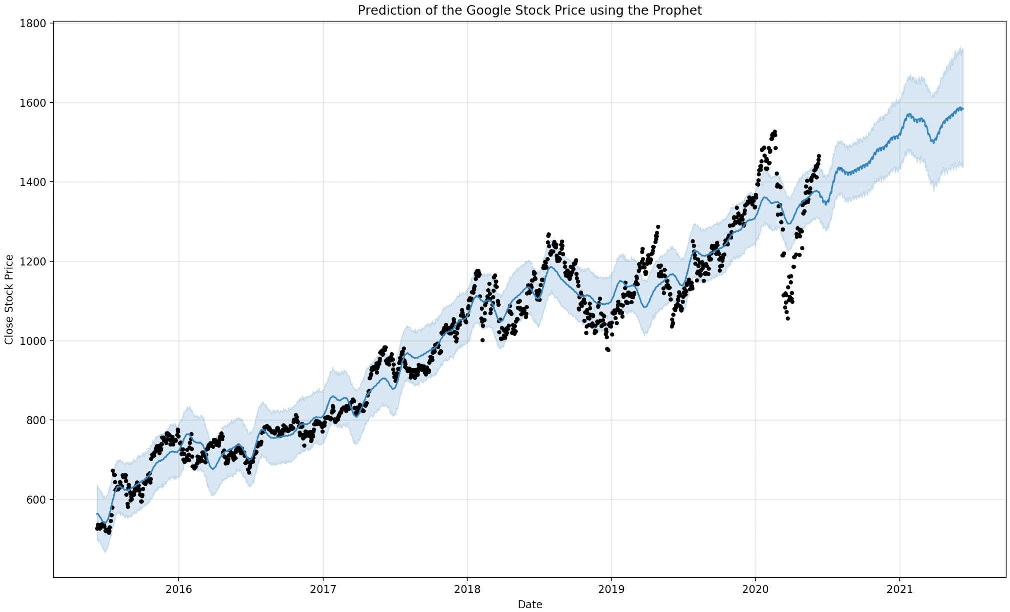
Motivation

Have you ever asked yourself what would happen if you have invested in Bitcoin in 2015?

What is a time series analyisis?

Time series analysis is a specific way of analyzing a sequence of data points collected over an interval of time.

The analysis can show how variables change over time.



How are time series applied in real life?

Time series analysis helps organizations understand the underlying causes of trends or systemic patterns over time:

- Stock market
- Automated stock trading
- Weather data
- Quarterly sales
- Audio processing
- Video processing

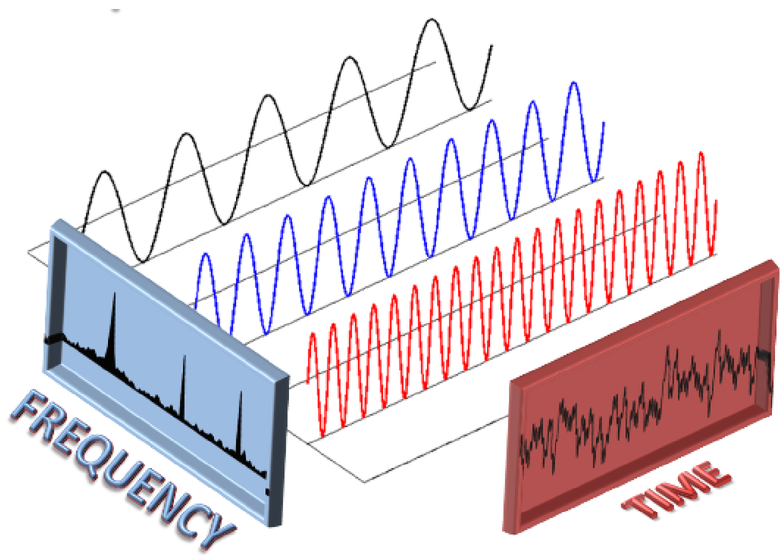
When are time series useful?

There are many tasks that can be performed using time series analysis, for example:

- Classification.
- Curve fitting.
- Descriptive analysis.
- Explanative analysis.
- Exploratory analysis.
- Forecasting.
- Intervention analysis.
- Segmentation.

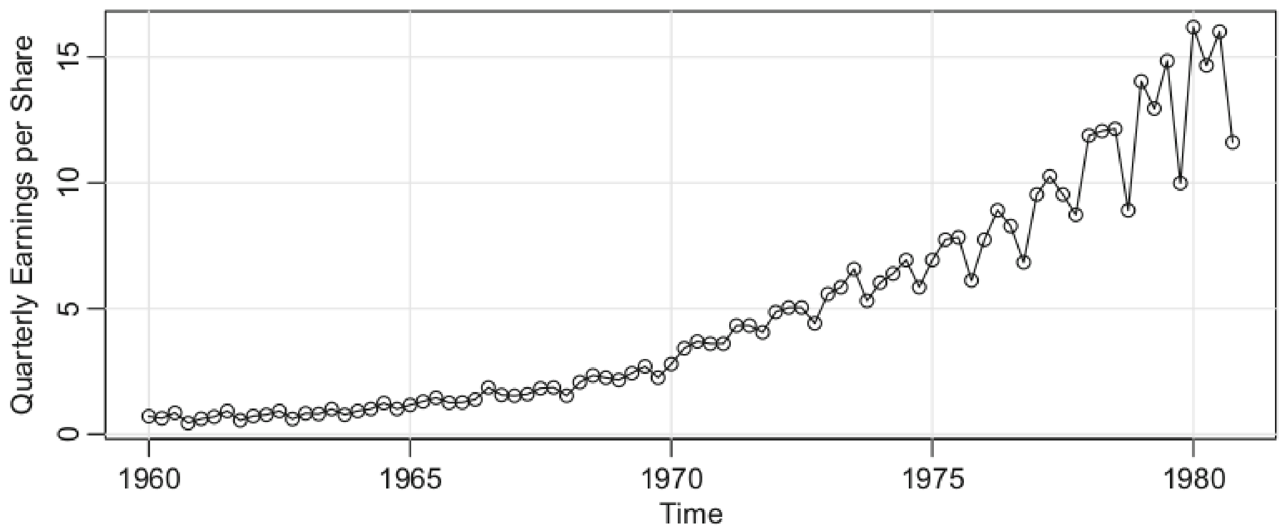
Characteristics of time series

- Correlation introduced by the sampling of adjacent points
- Time domain analysis approach
- Frecuency domain analysis approach



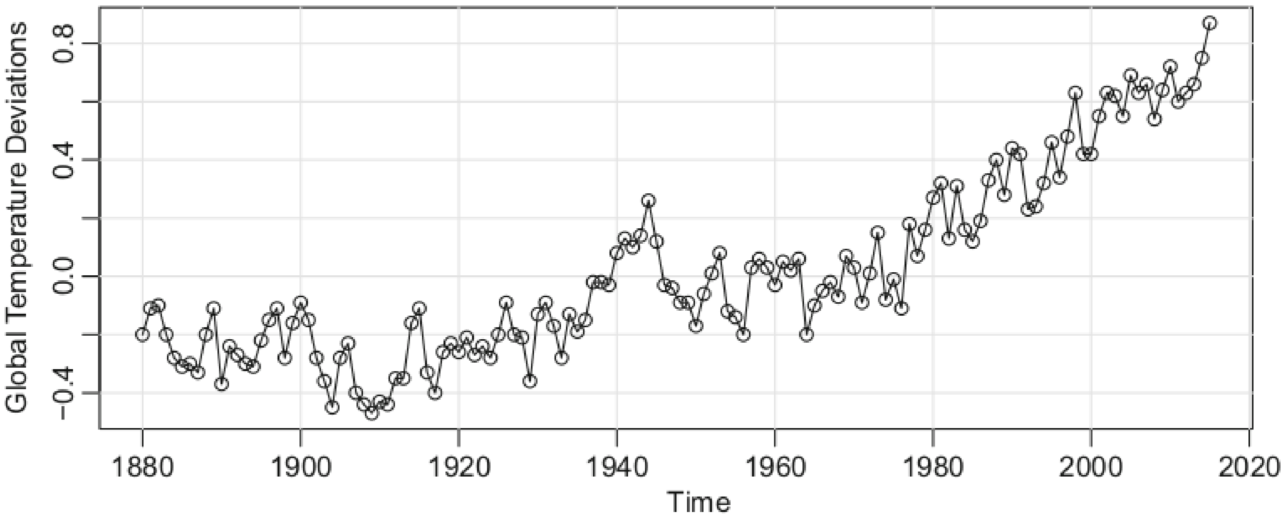
Example 1: Johnson & Johnson quarterly earnings

Figure shows quarterly earnings per share for the U.S. company Johnson & Johnson. There are 84 quarters (21 years) measured from the first quarter of 1960 to the last quarter of 1980. Modeling such series begins by observing the primary patterns in the time history. In this case, note the gradually increasing underlying trend and the rather regular variation superimposed on the trend that seems to repeat over quarters.



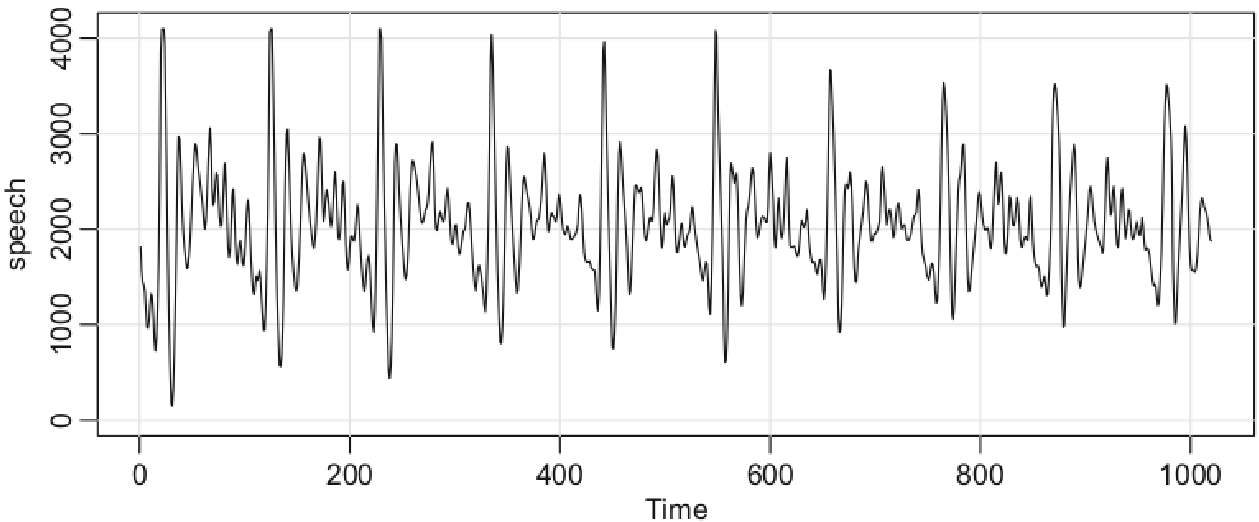
Example 2: Global warming

The data are the global mean land-ocean temperature. In particular, the data are deviations, measured in degrees centigrade, from the 1951–1980 average. We note an apparent upward trend in the series during the latter part of the twentieth century that has been used as an argument for the global warming hypothesis. Note also the leveling off at about 1935 and then another rather sharp upward trend at about 1970.



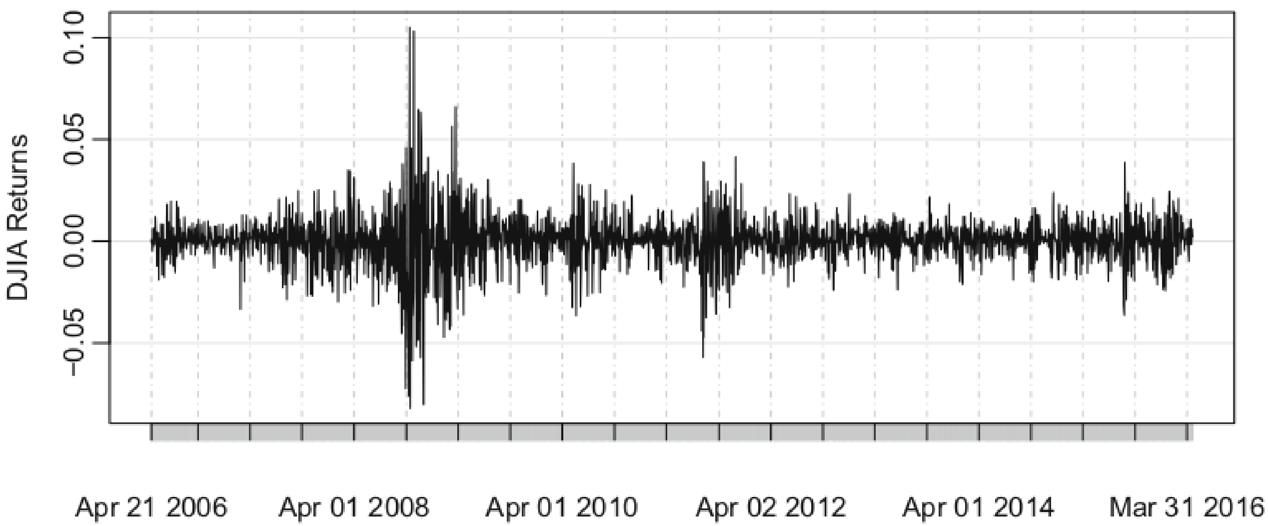
Example 3: Speech data

Figure shows a small .1 second (1000 point) sample of recorded speech for the phrase aaa...hhh, and we note the repetitive nature of the signal and the rather regular periodicities. Spectral analysis can be used in this context to produce a signature of this phrase that can be compared with signatures of various library syllables to look for a match.



Example 4: Dow Jones Industrial Average

Figure shows the daily returns (or percent change) of the Dow Jones Industrial Average (DJIA) from April 20, 2006 to April 20, 2016. It is easy to spot the financial crisis of 2008 in the figure. The mean of the series appears to be stable with an average return of approximately zero, however, highly volatile (variable) periods tend to be clustered together. A problem in the analysis of these type of financial data is to forecast the volatility of future returns.



The primary objective of time series analysis is to develop mathematical models that provide plausible descriptions for sample data.

Definición 1. *A time series can be defined as a collection of random variables indexed according to the order they are obtained in time.*

Conventionals

- To display a sample time series graphically by plotting the values of the random variables on the vertical axis, or ordinate, with the time scale as the abscissa.
- To connect the values at adjacent time periods to reconstruct visually some original hypothetical continuous time series that might have produced these values as a discrete sample. The appearance of data can be changed completely by adopting an insufficient sampling rate.

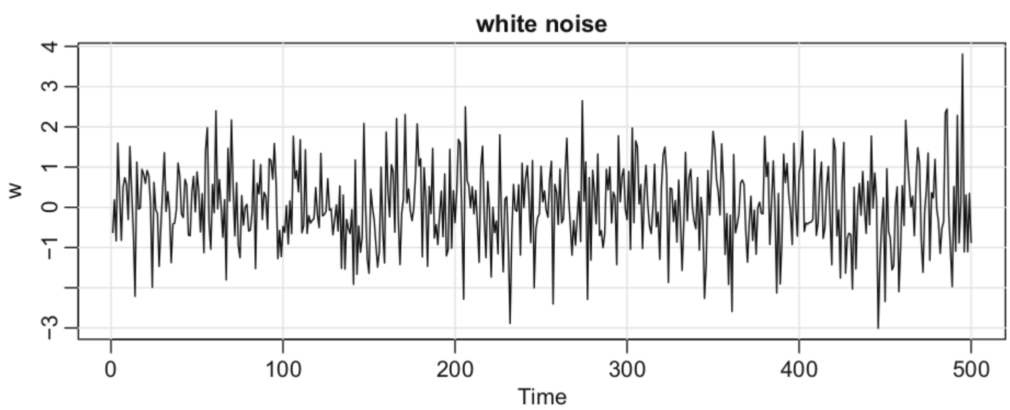
The fundamental visual characteristic distinguishing the different series is their differing degrees of smoothness. One possible explanation for this smoothness is that it is being induced by the supposition that adjacent points in time are correlated, so the value of the series at time t, say, x_t, depends in some way on the past values x_{t-1}, x_{t-2}, ... This model expresses a fundamental way in which we might think about generating realistic-looking time series.

Statisticals models: White noise

A simple kind of generated series might be a collection of uncorrelated random variables, w_t, with mean 0 and finite variance σ_w^2.The time series generated from uncorrelated variables is used as a model for noise, and it is called, *white noise*.

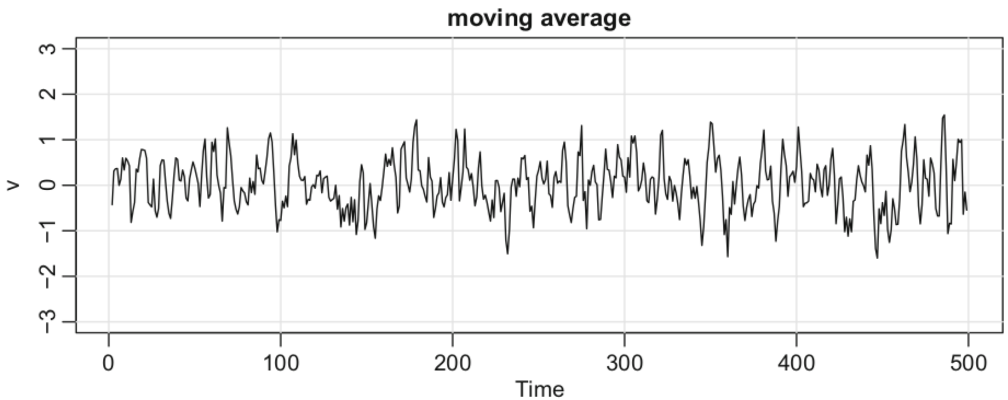
$$w_t \backsim \mathrm{wn}(0, \sigma_w^2)$$

$$w_t \sim i.i.d\, N(0, \sigma_w^2)$$



We might replace the white noise series w_t by a moving average that smooths the series.

$$v_t = \frac{1}{3}(w_{t-1} + w_t + w_{t+1})$$

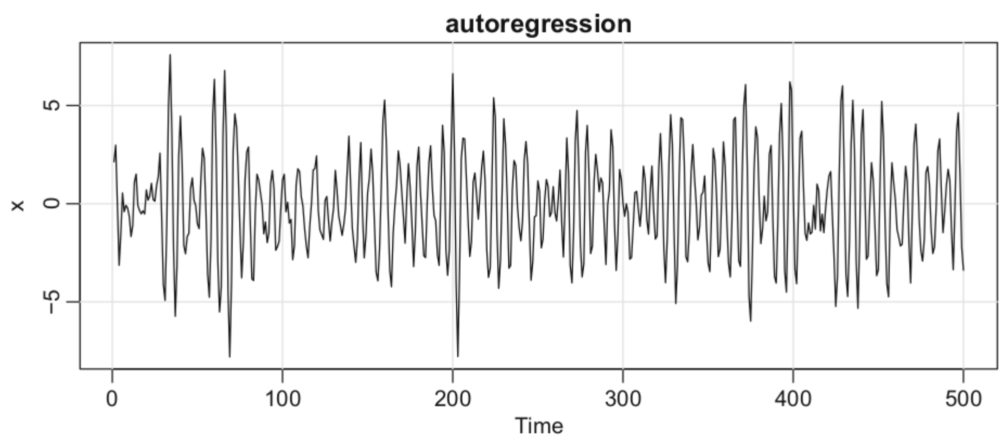


Autoregressions

Suppose we consider the white noise series w_t as input and calculate the output using the second-order equation:

$$x_t = x_{t-1} - 0.9x_{t-2} + w_t$$

The equation represents a regression of the current value x_t of a time series as a function of the past two values of the series.



Random walk with drift

A model for analyzing trend is the random walk with drift model given by:

$$x_t = \delta + x_{t-1} + w_t$$

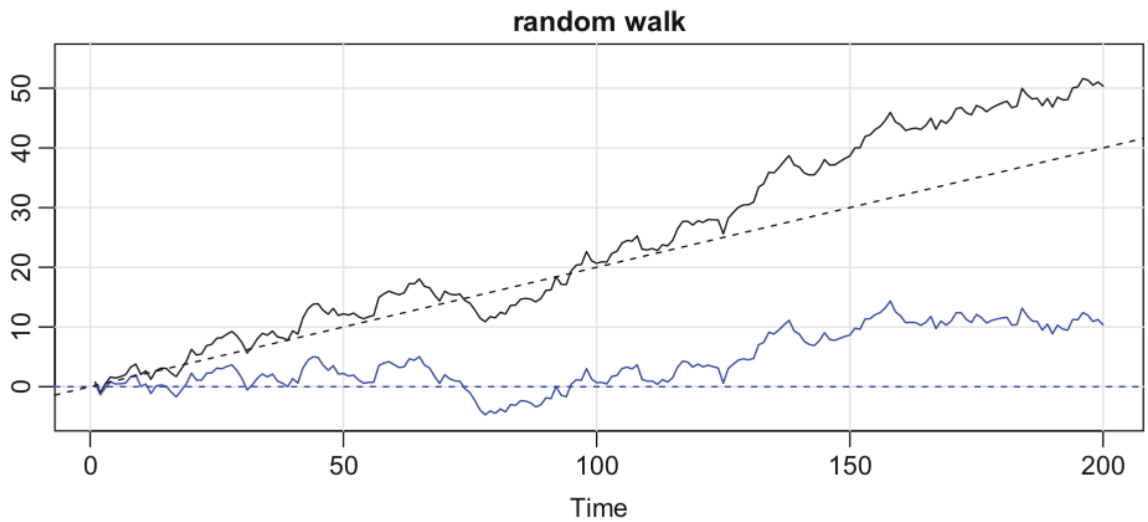
for $t = 1, 2, \ldots$, with initial condition $x_0 = 0$, and where w_t is white noise. The constant δ is called the *drift*, and when $\delta = 0$ is called simply a *random walk*. This can be rewritten as:

$$x_t = \delta t + \sum_{j=1}^{t} w_j$$

Example: Random walk with drift

Figure shows 200 observations generated from the model with $\delta = 0$ and $\delta = 0.2$, and with $\sigma_w = 1$. For comparison, we also superimposed the straight line $0.2t$ on the graph.

$$x_t = \delta t + \sum_{j=1}^{t} w_j$$



Signal in noise

Many realistic models for generating time series assume an underlying signal with some consistent periodic variation, contaminated by adding a random noise. Consider the next model:

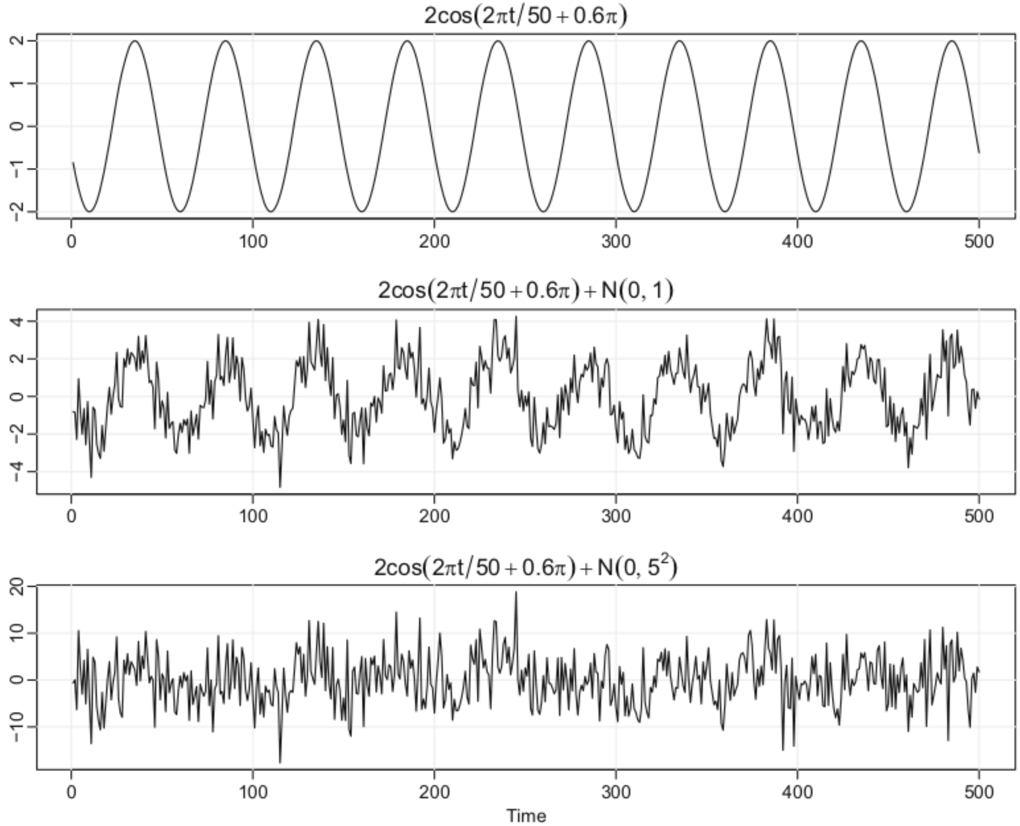
$$x_t = 2\cos\left(2\pi\frac{t+15}{50}\right) + w_t$$

We note that a sinusoidal waveform can be written as:

$$\mathrm{A}\cos(2\pi\omega t + \phi)$$

The degree to which the signal is obscured depends on the amplitude of the signal and the size of σ_w. The ratio of the amplitude of the signal to σ_w (or some function of the ratio) is sometimes called the *signal-to-noise ratio (SNR)*; the larger the *SNR*, the easier it is to detect the signal.

Example: Singal in noise



Why to study the previous models?

In the previous examples, we have tried to motivate the use of various combinations of random variables emulating real time series data. Some observed characteristics were:

- Smoothness characteristics of observed time series were introduced by combining the random variables in various ways.
- Averaging independent random variables over adjacent time points.
- Looking at the output of difference equations that respond to white noise inputs.

Measures of dependence

A complete description of a time series, observed as a collection of n random variables at arbitrary time points $t_1, t_2, \ldots, t_n$, for any positive integer n, is provided by the joint distribution function, evaluated as the probability that the values of the series are jointly less than the n constants, $c_1, c_2, \ldots, c_n$; i.e.,

$$F_{t_1,t_2,\ldots,t_n}(c_1, c_2, \ldots, c_n) = \mathbf{Pr}(x_{t_1} \leq c_1, x_{t_2} \leq c_2, \ldots, x_{t_n} \leq c_n)$$

The marginal distribution functions

The marginal distribution functions for the joint distribution function are:

$$F_t(x) = \boldsymbol{P}\{x_t \leq x\}$$

or the corresponding marginal density functions

$$f_t(x) = \frac{\partial F_t(x)}{\partial x}$$

when they exist, are often informative for examining the marginal behavior of a series.

The statistical problem...

Unfortunately, these multidimensional distribution functions cannot usually be written easily unless the random variables are jointly normal, in which case the joint density has the well-known form.

Although the joint distribution function describes the data completely, it is an unwieldy tool for displaying and analyzing time series data, because, the joint distribution must be evaluated as a function of n arguments, so any plotting of the corresponding multivariate density functions is virtually impossible.

The next measures could be helpful to describe a time series when the joint distribution function is not available.

The mean function (expected value)

Another informative marginal descriptive measure is the mean function.

Definición 2. *The mean function is defined as:*

$$\mu_{x_t} = \mathbb{E}[x_t] = \int_{-\infty}^{\infty} x\, f_t(x)\, dx$$

Definición 3. *A **linear map** or **linear function** $f(x)$ is a function that satisfies the two properties*

- *Additivity:* $f(x+y) = f(x) + f(y)$
- *Homogeneity:* $f(\alpha x) = \alpha f(x)$

Prove that expected value is a linear operator.

White noise

What is the expected value of white noise?

$$w_t \backsim \mathcal{N}(0, \sigma_t^2)$$

$$f_t(x) = \frac{1}{\sigma_t \sqrt{2\pi}} \exp\left\{ -\frac{1}{2\sigma_t^2}(x - \mu_t)^2 \right\}$$

What is the expected value of a moving averages series?

$$v_t = \frac{1}{3}(w_{t-1} + w_t + w_{t+1})$$

Random walk with drift

What is the expected value of a randomwalk with drift?

$$x_t = \delta t + \sum_{j=1}^{t} w_j$$

Signal plus noise

What is the expected value of a signal plus noise?

$$x_t = 2\cos\left(2\pi \frac{t+15}{50}\right) + w_t$$

Definición 4. *The autocovariance function is defined as:*

$$\gamma_x(s,t) = \mathrm{cov}(x_s, x_t) = \mathbb{E}[(x_s - \mu_s)(x_t - \mu_t)]$$

The autocovariance measures the linear dependence between two points on the same series observed at different times.

What is the covariance of x_t with itself?

What is the autocovariance of white noise?

$$\gamma_w(s,t) = \mathrm{cov}(w_s, w_t)$$

What is the autocovariance of moving average?

$$\gamma_v(s,t) = \mathrm{cov}(v_s, v_t) = \mathrm{cov}\left\{\frac{1}{3}(w_{s-1} + w_s + w_{s+1}), \frac{1}{3}(w_{t-1}, w_t, w_{t+1})\right\}$$

$$\gamma_v(s,t) = \begin{cases} \frac{3}{9}\sigma_w^2 & s = t \\ \frac{2}{9}\sigma_w^2 & |s-t| = 1 \\ \frac{1}{9}\sigma_w^2 & |s-t| = 2 \\ 0 & |s-t| > 0 \end{cases}$$

Random walk

What is the autocovariance of a random walk?

$$\gamma_x(s,t) = \text{cov}(x_s, x_t) = \text{cov}\left(\sum_{j=1}^{s} w_j, \sum_{k=1}^{t} w_k\right)$$

Definición 5. *The autocorrelation function (ACF) is defined as:*

$$\rho(s,t) = \frac{\gamma(s,t)}{\sqrt{\gamma(s,s)\gamma(t,t)}}$$

The ACF measures the linear predictability of the series at time t, say x_t, using only the value x_s.

Teorema 6. *Let X and Y be random variables on a common probability space. Then:*

$$|E(\mathrm{XY})| \leq \sqrt{E(X^2)E(Y^2)}$$

Using the Cauchy-Schwarz inequality prove that:

$$-1 \leq \rho(s,t) \leq 1$$

If we can predict x_t perfectly from x_s through a linear relationship, $x_t = \beta_0 + \beta_1 x_s$, then the correlation will be $+1$ when $\beta_1 > 0$, and -1 when $\beta_1 < 0$.

What it means geometrically?

Hence, we have a rough measure of the ability to forecast the series at time t from the value at time s.

Often, we would like to measure the predictability of another series y_t from the series x_s. Assuming both series have finite variances, we have the following definition.

Definición 7. *The cross-covariance function between two series, x_t and y_t is:*

$$\gamma_{xy}(s,t) = \mathrm{cov}(x_s, y_t) = E[(x_s - \mu_{x_s})(y_t - \mu_{y_t})]$$

Why could be important to mix the analysis of two time series?

Definición 8. *The cross-correlation function is given by:*

$$\rho_{xy}(s,t) = \frac{\gamma_{xy}(s,t)}{\sqrt{\gamma_x(s,s)\gamma_y(t,t)}}$$

What is the main difference bvetween auto-functions and cross-functions?