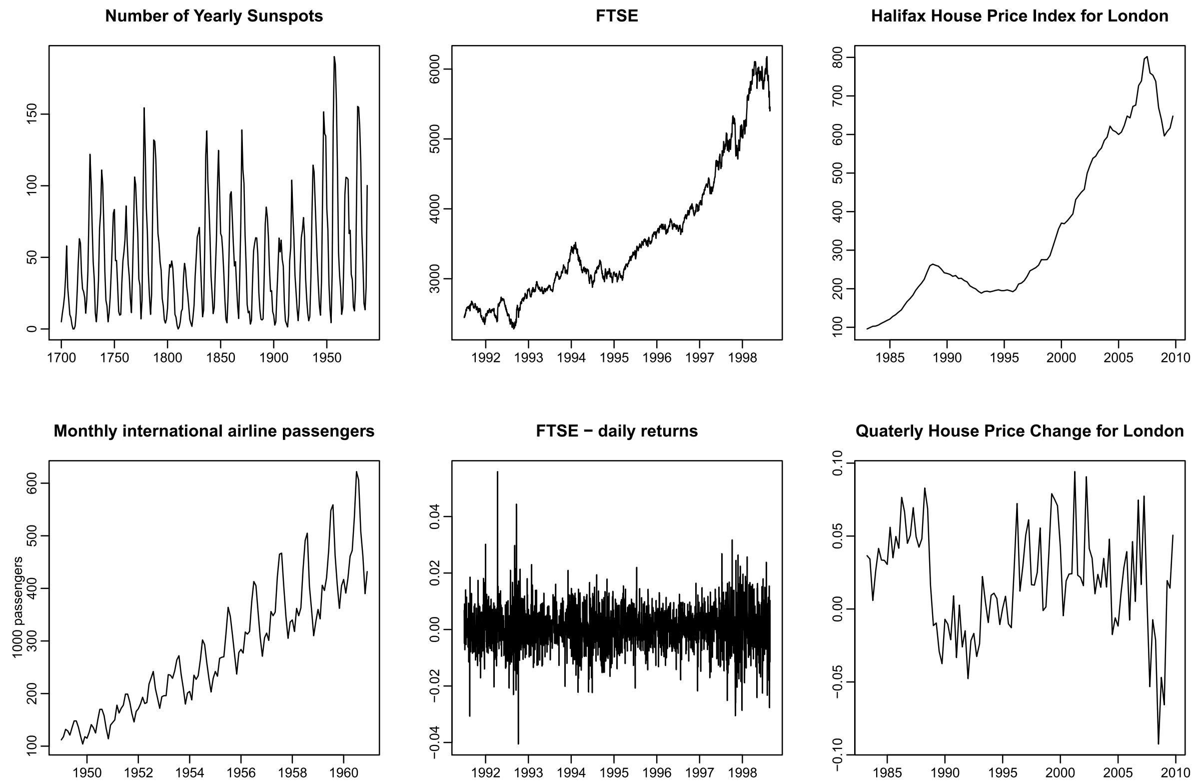
EE2 Mathematics – Probability & Statistics
Part 2: Time Series
2011/12
Lecture Notes

Dr. Miltiadis Mavrakakis
Imperial College London

Chapter 1

Time series: Time Domain



1.1 Terms

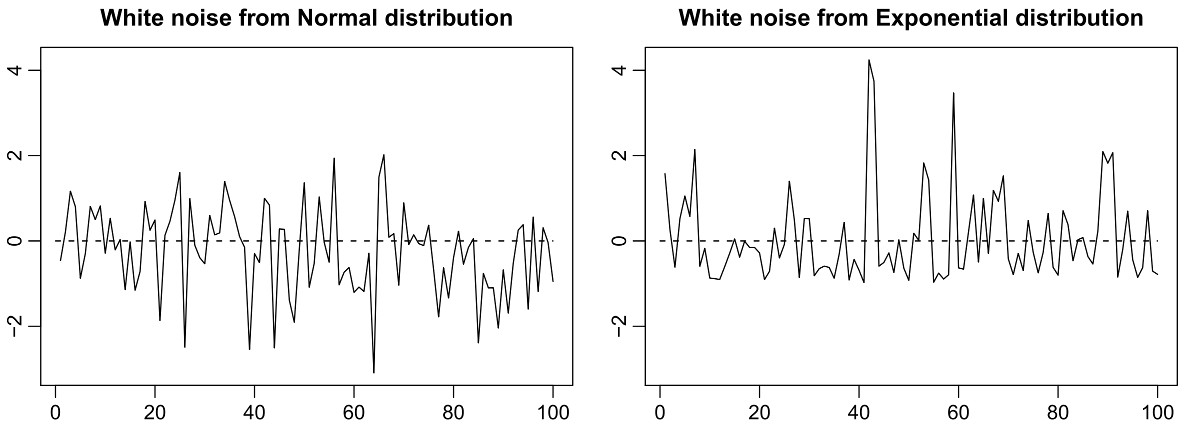
A time series is a sequence $\{y_t\}$ of random variables, e.g. $\{y_t\} = \{y_0, \dots, y_n\}$, $\{y_t\} = \{y_0, y_1, \dots\}$, $\{y_t\} = \{\dots, y_{-2}, y_{-1}, y_0, y_1, y_2, \dots\}$. y_t represents the observation made at 'time' t. Often, the realisation of these random variables is also called a time series.

The time intervals are taken to be equal here, though this is not necessary; also, 'time' may in fact be some other linear ordering, such as distance or similarity.

Example 1.1 (White noise). A time series $\{\varepsilon_t\}$ is called *white noise* if

- $\mathrm{E}(\varepsilon_t) = 0$ for all t,

- $\mathrm{Cov}(\varepsilon_t, \varepsilon_s) = 0$ for all $t \neq s$,
- $\mathrm{Var}(\varepsilon_t)$ does not depend on t, i.e. there exists σ_e^2 s.t. $\mathrm{Var}(\varepsilon_t) = \sigma_e^2$ for all t.



Example 1.2 (Simple random walk). The time series $\{y_t, t = 0, 1, \dots\}$ is called a *simple random walk* if

$$y_0 = 0 \quad \text{and} \quad y_t = y_{t-1} + \varepsilon_t, \quad t = 1, 2, ...$$

where $\{\varepsilon_t\}$ is white noise. Note that $y_t = \sum_{s=1}^{t} \varepsilon_s$, $\mathrm{E}(y_t) = 0$, $\mathrm{Var}(y_t) = t\,\mathrm{Var}(\varepsilon_1)$ and $\mathrm{Cov}(y_t, y_{t+s}) = t\,\mathrm{Var}(\varepsilon_1)$ (check!).

1.2 Stationarity

If the probability structure of a time series is constant in some way over time it is said to be *stationary*. The *mean function* or *trend* of the series $\{y_t\}$ is defined as

$$\mu_t = \mathrm{E}(y_t)$$

and the *covariance function* as

$$\gamma(t, t+s) = \mathrm{Cov}(y_t, y_{t+s})$$

If both μ_t and $\gamma(t, t+s)$ are independent of t the process is said to be *weakly* (or *second-order*) *stationary*. In this case $\gamma(t, t+s)$ depends only on the time difference s (known as the *lag*) and will be denoted by γ_s; this is commonly known as the *autocovariance function*. The variance of y_t is then γ_0, and the *autocorrelation function* is $\rho_s = \gamma_s/\gamma_0$. A plot of ρ_s versus s is known as a *correlogram*.

A time series $\{y_t\}$ is called *strictly stationary* if for all $m = 1, 2, \dots$ and all indices $t_1, \dots, t_m$, the joint distribution $(y_{t_1+s}, ..., y_{t_m+s})$ is the same as that of $(y_{t_1}, ..., y_{t_m})$ for any s. Strict stationarity implies (and is much stronger than) weak stationarity.

Most basic statistical methods rely only on weak stationarity, and we will use 'stationarity' to mean 'weak stationarity'.

1.3 Autoregressive and moving average processes

In this subsection we will be looking at time series $\{y_t, t = 0, \pm 1, \pm 2, \dots\}$

1.3.1 MA(q)

A *moving average process of order* q, denoted by MA(q), is defined by

$$y_t = \alpha_0\varepsilon_t + \alpha_1\varepsilon_{t-1} + \alpha_2\varepsilon_{t-2} + \ldots + \alpha_q\varepsilon_{t-q} = \sum_{j=0}^{q}\alpha_j\varepsilon_{t-j}$$

where $\{\varepsilon_t\}$ is white noise with $\mathrm{Var}(\varepsilon_t) = \sigma_\varepsilon^2$.

The model can be written as

$$y_t = \sigma_\varepsilon \sum_{j=0}^{p}\alpha_j f_{t-j}$$

where $f_j = \varepsilon_j/\sigma_\varepsilon$. There is a lack of identifiability here, e.g. σ_ε could be doubled and the α_js halved without changing the model equation. A standard way of avoiding this problem is to take $\alpha_0 = 1$.

For the mean we have $\mu_t = \mathrm{E}(y_t) = 0$ for all t, and the covariance function is given by

$$\mathrm{Cov}(y_t, y_{t+s}) = \left\{ \begin{array}{ll} 0 & \text{for } s > q \\ (\alpha_0\alpha_s + \alpha_1\alpha_{s+1} + \ldots + \alpha_{q-s}\alpha_q)\sigma_\varepsilon^2 & \text{for } s = 0, 1, 2, \ldots, q \end{array} \right\} = \gamma_s$$

This is independent of t and so the process is (weakly) stationary.

Example 1.3. MA(1) process:

$$y_t = \alpha_0\varepsilon_t + \alpha_1\varepsilon_{t-1} \quad \text{with} \quad \sigma_\varepsilon^2 = 1$$

For this process

$$\gamma_0 = \alpha_0^2 + \alpha_1^2, \quad \gamma_1 = \alpha_0\alpha_1 \quad \text{and} \quad \gamma_s = 0 \text{ for } s > 1$$

The first-order autocorrelation is $\rho_1 = \alpha_0\alpha_1/(\alpha_0^2 + \alpha_1^2)$.

Note that γ_0 and γ_1 are symmetric in α_0 and α_1, which means that the MA(1) process with α_0 and α_1 interchanged has the same variance-covariance structure as the original.

1.3.2 AR(1)

An time series $\{y_t\}$ is called *autoregressive process of order 1* (AR(1)) if

$$y_t = \beta y_{t-1} + \varepsilon_t,$$

where β is some constant and $\{\varepsilon_t\}$ is white noise with mean 0 and variance σ_ε^2.

Under what conditions is $\{y_t\}$ stationary?

Taking expectations,

$$\mu_t = \beta\mu_{t-1}$$

Thus the process has constant mean iff $\beta = 1$ or $\mu_t = 0$ for all t.Furthermore,

$$\underbrace{\mathrm{Var}(y_t)}_{=:\sigma_t^2} = \beta^2\underbrace{\mathrm{Var}(y_{t-1})}_{=:\sigma_{t-1}^2} + \sigma_\varepsilon^2,$$

since y_{t-1} and ε_t are uncorrelated. Thus, for $\sigma_t = \text{Var}(y_t)$ to be constant we need

$$\sigma_t^2 = \frac{\sigma_\varepsilon^2}{1-\beta^2},$$

which implies $\mid \beta \mid < 1$. Lastly, consider the covariance function. Repeated substitution yields

$$y_t = \beta\{\beta y_{t-2} + \varepsilon_{t-1}\} + \varepsilon_t = \beta^2 y_{t-2} + \beta\varepsilon_{t-1} + \varepsilon_t = ... = \sum_{j=0}^{s-1} \beta^j \varepsilon_{t-j} + \beta^s y_{t-s}.$$

Then we have

$$\text{Cov}(y_t, y_{t-s}) = \sum_{j=0}^{s-1} \beta^j \,\text{Cov}(\varepsilon_{t-j}, y_{t-s}) + \beta^s \,\text{Var}(y_{t-s}) = 0 + \beta^s \frac{\sigma_\varepsilon^2}{1-\beta^2}$$

using the fact that the ε_{t-j} are all uncorrelated with y_{t-s}. Thus, $\{y_t\}$ is weakly stationary if $\mid \beta \mid < 1$. Its autocovariance function is $\gamma_s = \beta^s \sigma_\varepsilon^2/(1-\beta^2)$ and its autocorrelation function is $\rho_s = \beta^s$.

1.3.3 AR(p)

An *autoregressive process of order p*, denoted by AR(p), is defined by

$$y_t = \beta_1 y_{t-1} + \beta_2 y_{t-2} + ... + \beta_p y_{t-p} + \varepsilon_t$$

where $\{\varepsilon_t\}$ is white noise.

The equation may be written in terms of the *backshift operator* B, defined by $By_t = y_{t-1}$ (so $B^2 y_t = By_{t-1} = y_{t-2}$, etc.). Thus, the AR($p$) can be written as

$$y_t = \sum_{j=1}^{p} \beta_j B^j y_t + \varepsilon_t \quad \text{ or } \beta(B) y_t = \varepsilon_t$$

where

$$\beta(B) = 1 - \sum_{j=1}^{p} \beta_j B^j$$

is a polynomial in B of order p.

Example 1.4. Consider the AR(2) process

$$6y_t - y_{t-1} - y_{t-2} = 5\varepsilon_t$$

This can be written as

$$(2-B)(3+B)y_t = 5\varepsilon_t$$

so (calculating symbolically)

$$\begin{aligned} y_t &= 5\{(2-B)(3+B)\}^{-1}\varepsilon_t = ((2-B)^{-1} + (3+B)^{-1})\varepsilon_t \\ &= \frac{1}{2}\sum_{j=0}^{\infty}(B/2)^j \varepsilon_t + \frac{1}{3}\sum_{j=0}^{\infty}(-B/3)^j \varepsilon_t \\ &= \sum_{j=0}^{\infty}\{(\frac{1}{2})^{j+1} - (-\frac{1}{3})^{j+1}\}\varepsilon_{t-j} \end{aligned}$$

Hence,

$$\gamma_0 = \mathrm{Var}(y_t) = \sigma_\varepsilon^2 \sum_{j=0}^{\infty} \{(\frac{1}{2})^{j+1} - (-\frac{1}{3})^{j+1}\}^2 = ... \quad \text{(geometric series)}$$
$$= (\frac{1}{3} + \frac{2}{7} + \frac{1}{8})\sigma_\varepsilon^2 = \frac{125}{168}\sigma_\varepsilon^2$$

and

$$\gamma_s = \mathrm{Cov}(y_t, y_{t+s})$$
$$= \sigma_\varepsilon^2 \sum_{j=0}^{\infty} \{(\frac{1}{2})^{j+1} - (-\frac{1}{3})^{j+1}\}\{(\frac{1}{2})^{j+s+1} - (-\frac{1}{3})^{j+s+1}\} = ...$$
$$= \sigma_\varepsilon^2 \{\frac{10}{21}(\frac{1}{2})^s + \frac{15}{56}(-\frac{1}{3})^s\}$$

1.3.4 ARMA(p,q)

An *autoregressive-moving average process of order (p,q)*, denoted by ARMA(p, q), is defined by

$$y_t = \sum_{j=1}^{p} \beta_j y_{t-j} + \sum_{j=0}^{q} \alpha_j \varepsilon_{t-j}$$

where $\{\varepsilon_t\}$ is white noise with variance σ_ε^2.

Example 1.5. ARMA(1,1) process. For this model take

$$\beta(B) = 1 - \beta B \text{ and } \alpha(B) = 1 + \alpha B \text{ with } \beta_1 = \beta, \alpha_0 = 1, \alpha_1 = \alpha$$

Then we have

$$(1 - \beta B)y_t = (1 + \alpha B)\varepsilon_t \text{ or } y_t = \beta y_{t-1} + \varepsilon_t + \alpha\varepsilon_{t-1}$$

Formal manipulations give

$$y_t = (1 - \beta B)^{-1}(1 + \alpha B)\varepsilon_t = \sum_{j=0}^{\infty} (\beta B)^j (1 + \alpha B)\varepsilon_t$$
$$= \varepsilon_t + \sum_{j=1}^{\infty} (\beta^j + \beta^{j-1}\alpha)\varepsilon_{t-j} = \varepsilon_t + (\alpha + \beta)\sum_{j=0}^{\infty} \beta^j \varepsilon_{t-j-1}$$

Assuming that $\mid \beta \mid < 1$, we have

$$\mathrm{Var}(y_t) = \sigma_\varepsilon^2 + (\alpha + \beta)^2 \sum_{j=0}^{\infty} \beta^{2j}\sigma_\varepsilon^2 = \sigma_\varepsilon^2\{1 + (\alpha + \beta)^2/(1 - \beta^2)\} = \sigma_\varepsilon^2(1 + \alpha^2 + 2\alpha\beta)/(1 - \beta^2)$$

For the covariance function first note that

$$y_{t+s} = \varepsilon_{t+s} + (\alpha + \beta)\sum_{j=0}^{\infty} \beta^j \varepsilon_{t+s-j-1}$$
$$= \varepsilon_{t+s} + (\alpha + \beta)\sum_{j=0}^{s-2} \beta^j \varepsilon_{t+s-j-1} + (\alpha + \beta)\beta^{s-1}\varepsilon_t + (\alpha + \beta)\sum_{j=0}^{\infty} \beta^{j+s}\varepsilon_{t-j-1}$$

then

$$
\begin{aligned}
cov(y_t, y_{t+s}) = \mathrm{E}(y_t y_{t+s}) &= \sigma_\varepsilon^2(\alpha+\beta)\beta^{s-1} + \sigma_\varepsilon^2(\alpha+\beta)^2 \sum_{j=0}^{\infty} \beta^{2j+s} \\
&= \sigma_\varepsilon^2(\alpha+\beta)\beta^{s-1} + \sigma_\varepsilon^2(\alpha+\beta)^2\beta^s/(1-\beta^2) = \sigma_\varepsilon^2(\alpha+\beta)\beta^{s-1}(1+\alpha\beta)/(1-\beta^2)
\end{aligned}
$$

This is independent of t so the process is (weakly) stationary.

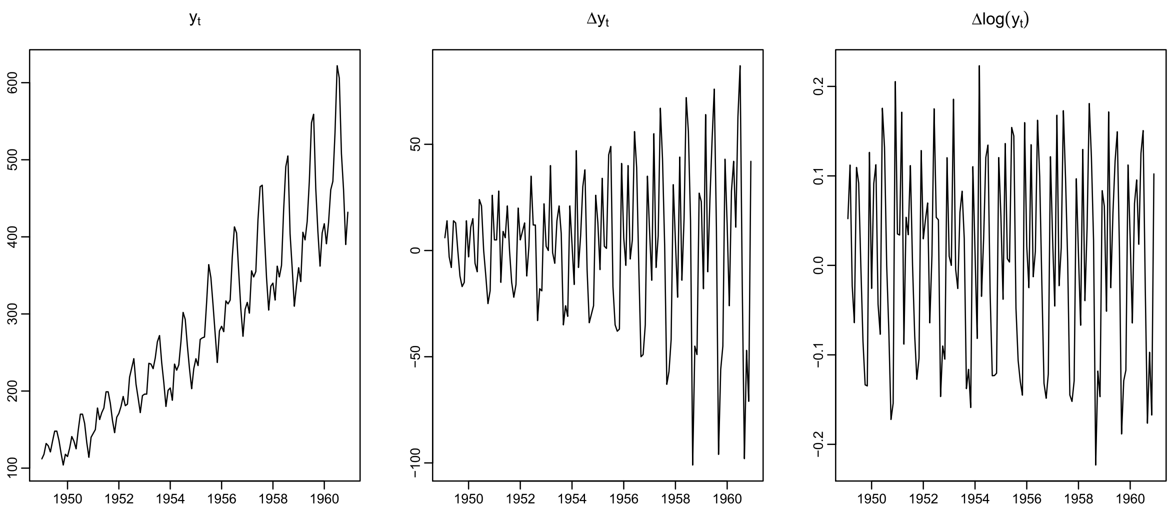
1.4 Simple analysis of data

When a *realisation* $\{y_1, \ldots, y_n\}$ of a time series is available for inspection some basic descriptive methods can be applied.

First, a plot of y_t vs t should throw some light on the form of the mean function, or trend, μ_t. For this it might be better to *smooth* the series first, e.g. by applying a *moving average* such as $\frac{1}{3}(y_{t-1} + y_t + y_{t+1})$. The trend, once identified, can be subtracted out, to give a zero-mean series, $\{y_t - \mu_t\}$.

Trends can be removed via differencing. If $y_t = at + \epsilon_t$ where ϵ_t stationary then $\Delta y_t = y_t - By_t = y_t - y_{t-1} = a + \epsilon_t - \epsilon_{t-1}$ where $\Delta = 1 - B$ is the differencing operator.

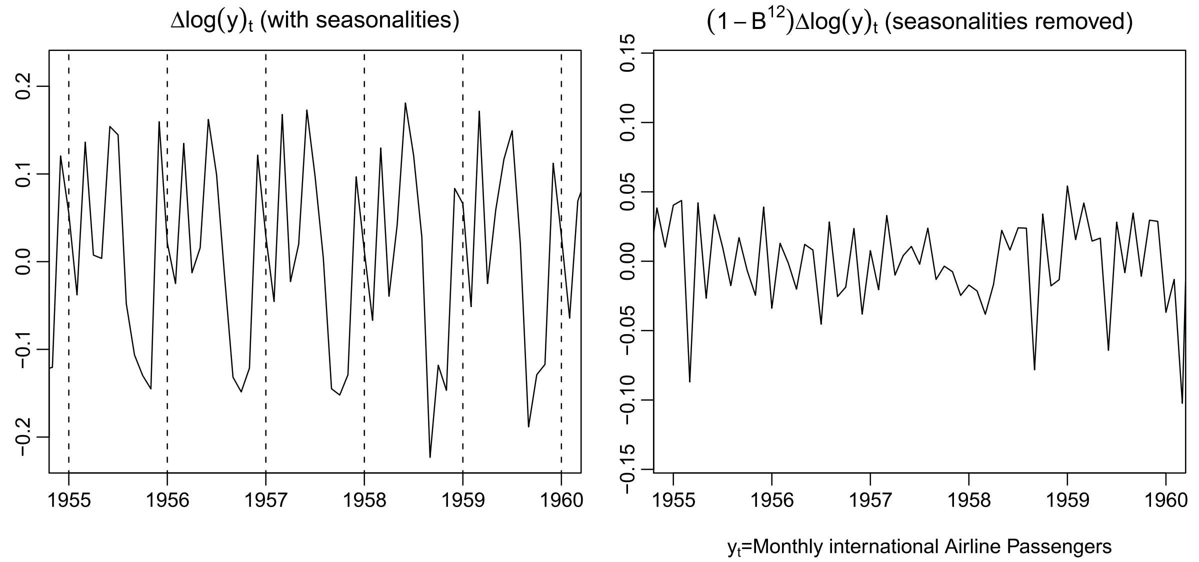
Example:



y_t=Monthly international Airline Passengers

Seasonalities can also be removed by differencing. Suppose $y_t = x_t + \epsilon_t$ where x_t has a period of length s and ϵ_t is stationary. Then $(I - B^s)y_t = y_t - y_{t-s} = x_t - x_{t-s} + \epsilon_t - \epsilon_{t-s} = \epsilon_t - \epsilon_{t-s}$ which is stationary.

Example:

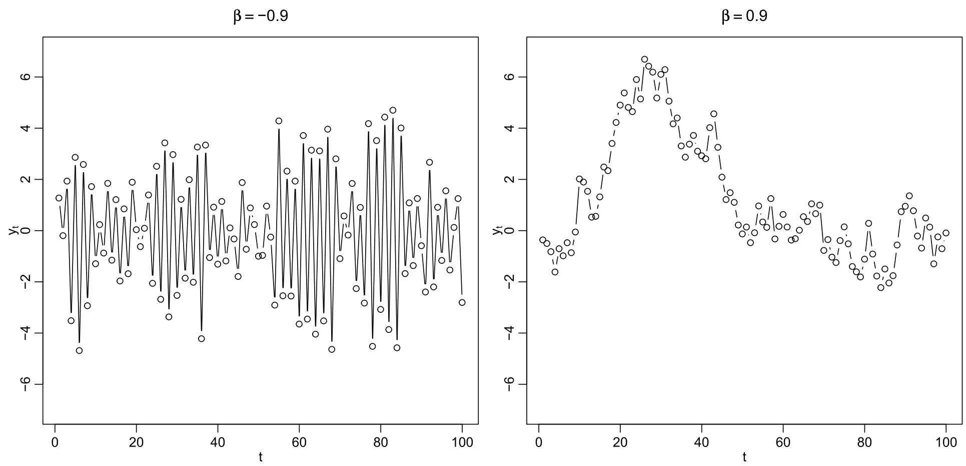


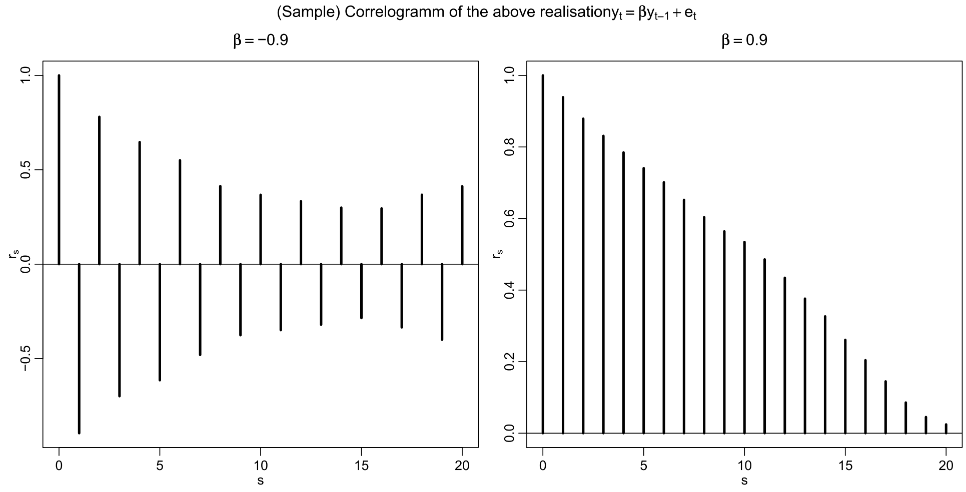
Assuming now that we have a stationary time series, it makes sense to compute the *sample autocovariance function* as

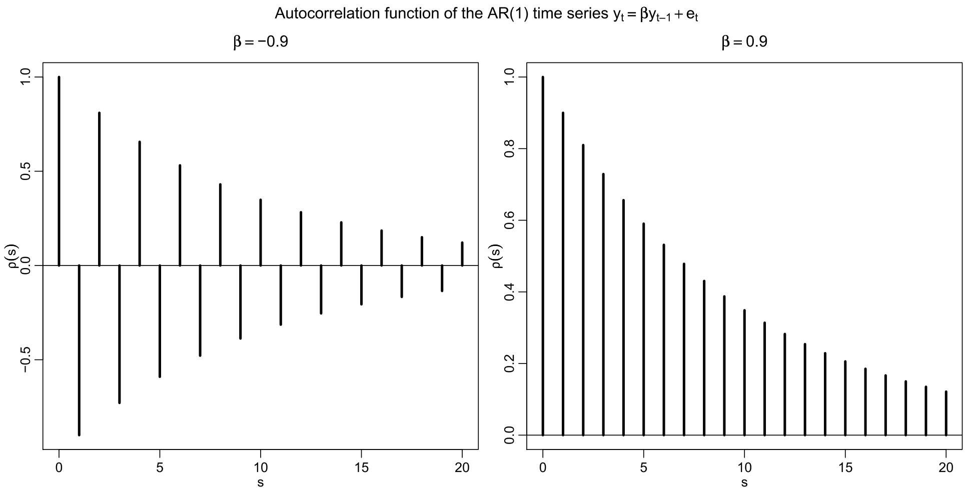
$$c_s = n^{-1} \sum_{t=1}^{n-s} (y_t - \bar{y})(y_{t+s} - \bar{y}).$$

Then, the *sample autocorrelation function* is given by $r_s = c_s/c_0$ and this can be plotted against s, this plot being known as the *correlogram.*

Example:







Correlogram for the examples: