# 3. Time series and stochastic processes

## Outline:

- Introduction
- The concept of the stochastic process
- Stationary processes
- White noise process
- Estimating the moments of stationary processes

## Recommended readings:

▷ Chapter 3 of D. Peña (2008).

▷ Chapter 2 of P.J. Brockwell and R.A. Davis (1996).

# Introduction

▷ The dynamic phenomena that we observe in a time series can be grouped into two classes:

- The first are those that take stable values in time around a constant level, without showing a long term increasing or decreasing trend. For example, yearly rainfall in a region, average yearly temperatures or the proportion of births corresponding to males. These processes are called **stationary**.

- A second class of processes are the **non-stationary** processes, which are those that can show trend, seasonality and other evolutionary effects over time. For example, the yearly income of a country, company sales or energy demand are series that evolve over time with more or less stable trends.

▷ In practice, the classification of a series as stationary or not depends on the period of observation, since the series can be stable in a short period and non-stationary in a longer one.

# The concept of the stochastic process

▷ A **stochastic process** is a set of random variables $\{z_t\}$ where the index $t$ takes values in a certain set $C$. In our case, this set is ordered and corresponds to moments of time (days, months, years, etc.).

▷ For each value of $t$ in set $C$ (for each point in time) a random variable, $z_t$, is defined and the observed values of the random variables at different times form a **time series**. That is, a series of $T$ data, $(z_1, \ldots, z_T)$, is a sample of size one of the vector of $T$ random variables ordered in time corresponding to the moments $t = 1, \ldots, T$, and the observed series is considered a result or **trajectory** of the stochastic process.

▷ The process is characterized by the joint probability distribution of the random variables $(z_1, \ldots, z_T)$, for any value of $T$. These distributions are called **finite-dimensional distributions** of the process. We say that we know the probabilistic structure of a stochastic process when we know these distributions, which determine the distribution of any subset of variables and, in particular, the **marginal distribution** of each variable.

# The concept of the stochastic process - Examples

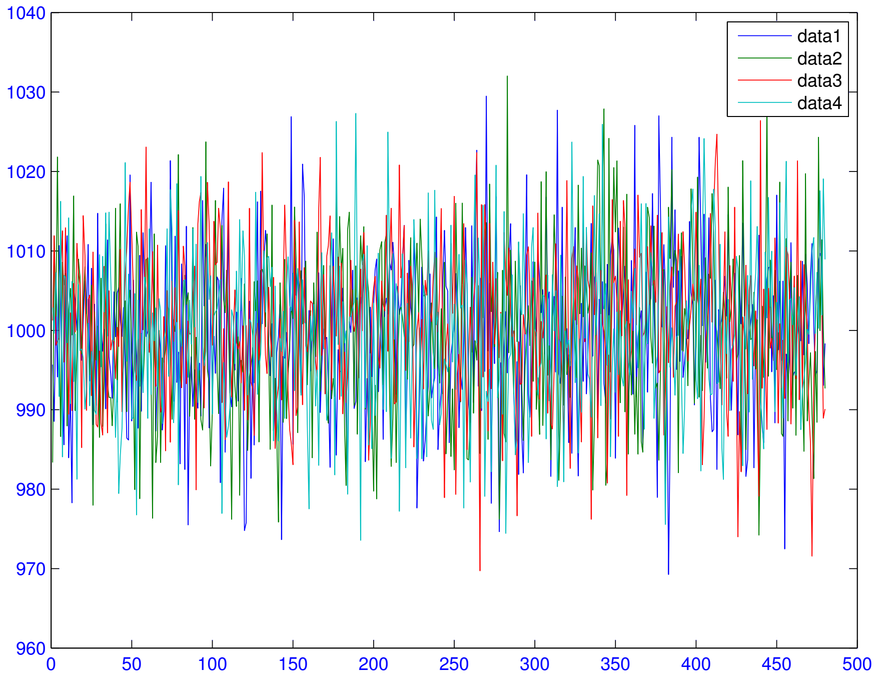
**Example 24.** *Let us assume that in an industrial company the temperatures of an oven are measured every minute from the time it starts up, at nine o'clock, until it is turned off, at five o'clock.*

*Every day 480 observations are obtained, which correspond to the oven temperature at times 9h 1m, 9h 2m,...,4h 59m, 5h, that we associate with the values* $t = 1, ..., 480.$

*The set of temperature measurements on any given day constitute the realization of a stochastic process* $\{z_t\}$, $t = 1, ..., 480$, *where* $z_t$ *is the random variable: oven temperature at time* $t$.

*If we have many trajectories of the process (data from many days) we can obtain the probability distribution of the variables that comprise it. To do this, we must assume that the oven temperature at a set time* $t_0$, *follows a definite probability distribution.*

# The concept of the stochastic process - Examples

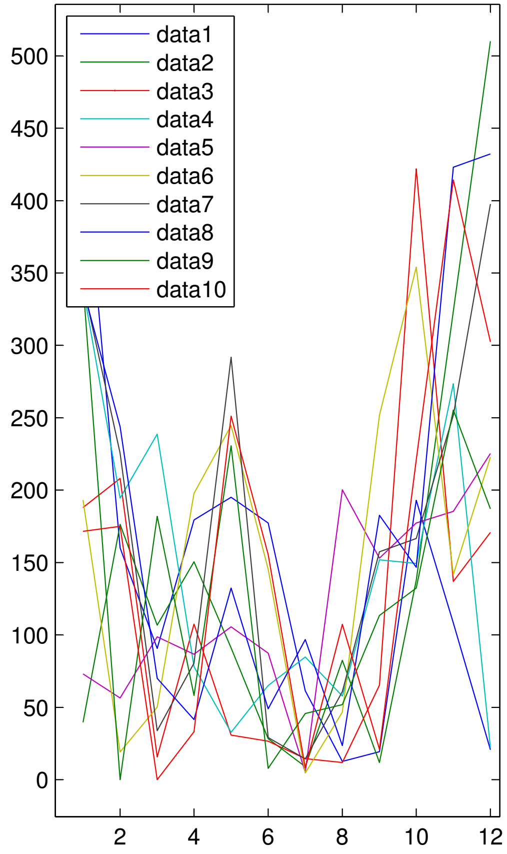


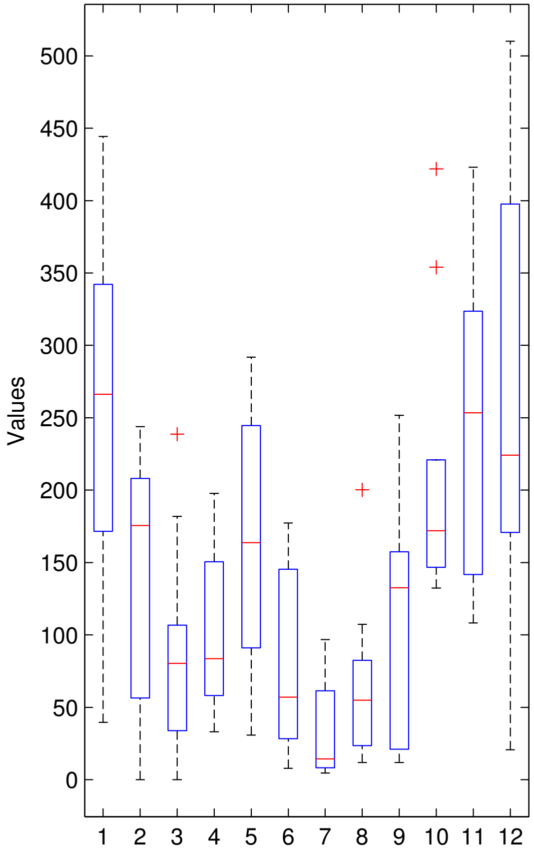

▷ *If this hypothesis is true, which implies that the situation of the oven is the same on the different days, we can obtain the distribution of each variable $z_t$, by studying the distribution of the $t-th$ observation on different days.*

▷ *Analogously, we can study the joint distribution of two consecutive variables $(z_{t_0}, z_{t_0+1})$, taking the pairs of observed values at times $t_0$ and $t_0 + 1$ on different days. And so on ...*

# The concept of the stochastic process - Examples

**Example 25.** *The figure shows 10 realizations of the rainfall series in Santiago de Compostela over the 12 months of the year. We have 10 values of 12 random variables, one for each month, and the trajectory of the 12 values in a given year represents a realization of the stochastic process.*



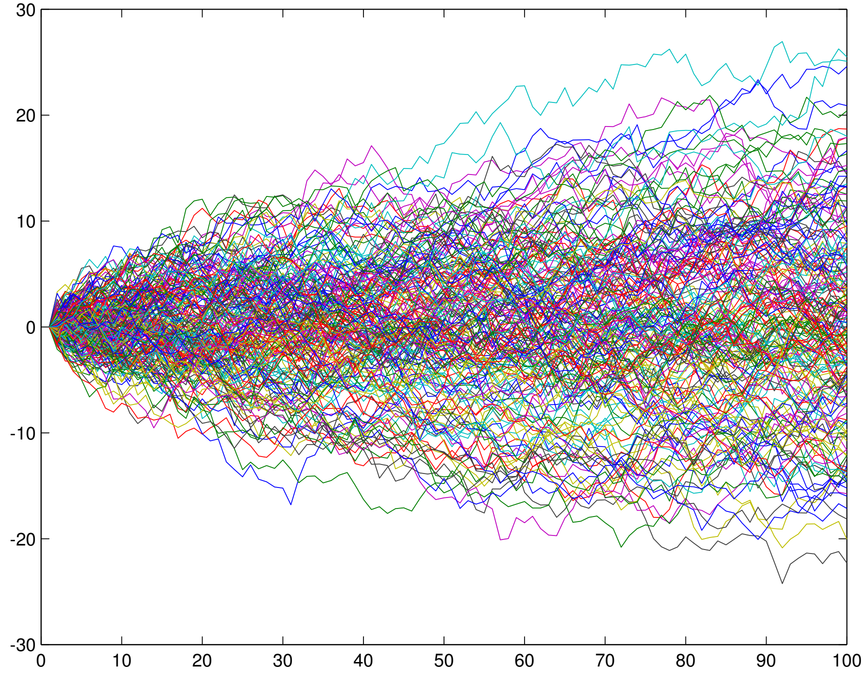



# The concept of the stochastic process - Examples

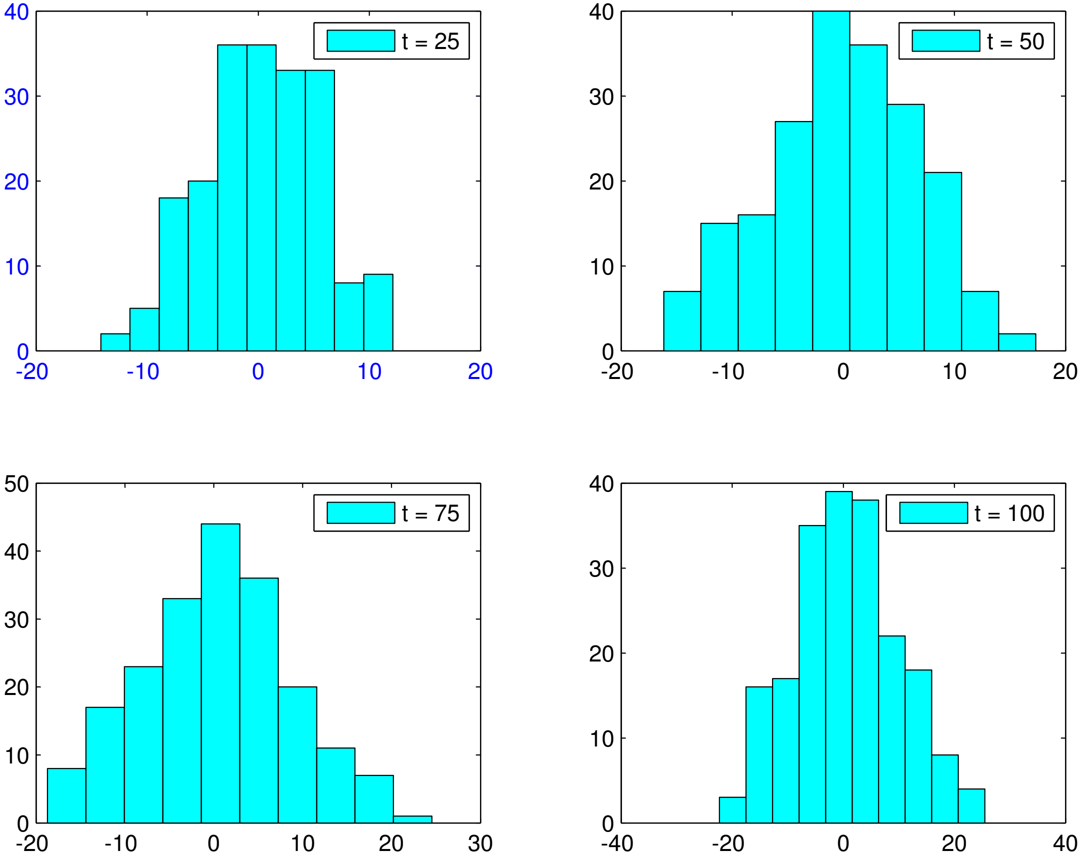
**Example 26.** *Let us consider the stochastic process defined by:*

$$z_t = z_{t-1} + a_t \tag{20}$$

*which we assume begins at* $z_0 = 0$, *and where the* $a_t$ *are i.i.d.* $\mathcal{N}(0, \sigma^2)$. *This process is known as a* **random walk** *and the figure shows 200 realizations of the process carried out by computer.*



# The concept of the stochastic process - Examples




Observe that the mean of these distributions is approximately zero in the four cases, but the variance of the distribution increases with the time period being considered. This result coincides with the previous figure, where we observe that the realizations of the process tend to move away from the initial value over time.

# Properties of marginal distributions

▷ The **mean function** of the process refers to a function of time representing the expected value of the marginal distributions $z_t$ for each instant:

$$E\left[z_t\right] = \mu_t. \tag{21}$$

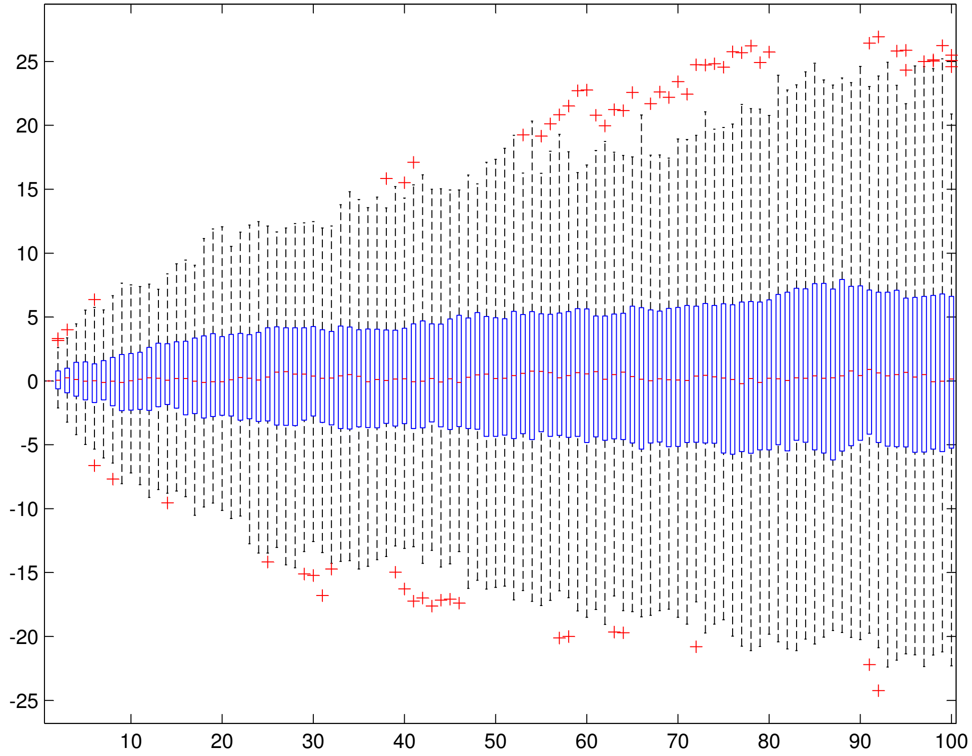
▷ An important particular case, due to its simplicity, arises when all the variables have the same mean and thus the mean function is a constant. The realizations of the process show no trend and we say that the process is **stable in the mean**.

▷ If, on the contrary, the means change over time, the observations at different moments will reveal that change.

▷ On many occasions we only have one realization of the stochastic process and we have to deduce from that whether the mean function of the process is, or is not, constant over time.

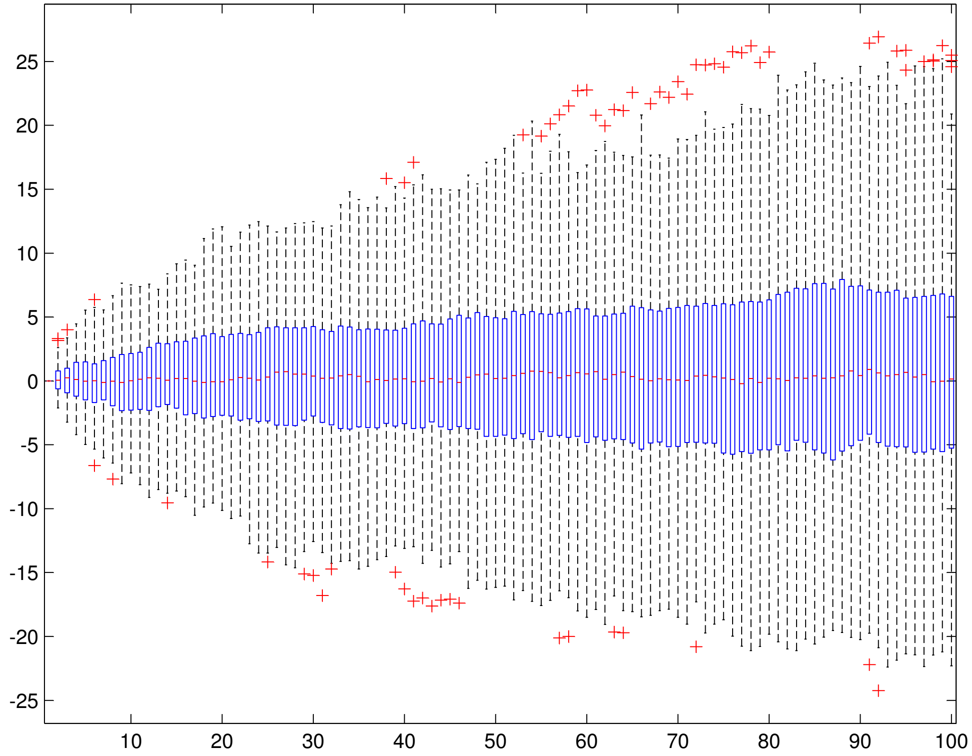
# Properties of marginal distributions - Examples

**Example 27.** *The mean of the random walk variables seem to be constant and near zero. This can be deduced from the equation of the process, (20), since for the first variable we can write* $E[z_1] = 0 + E(a_1) = 0$ *and for the rest we can use that if the expectation of* $z_{t-1}$ *is zero so is that of* $z_t$*, since:* $E[z_t] = E[z_{t-1}] + E(a_t) = 0.$



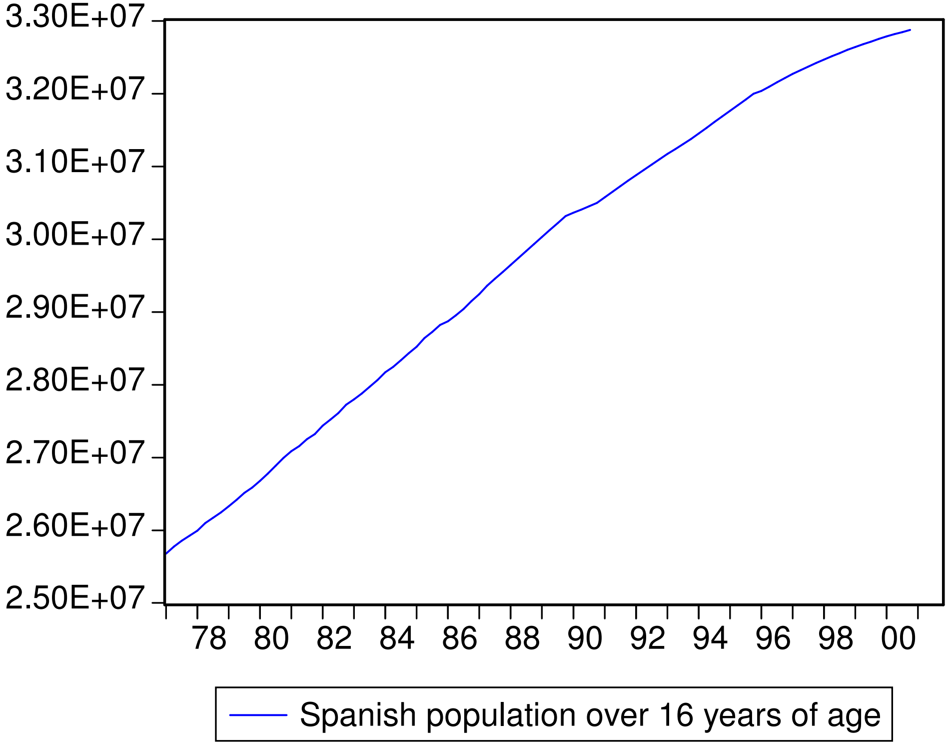
# Properties of marginal distributions - Examples

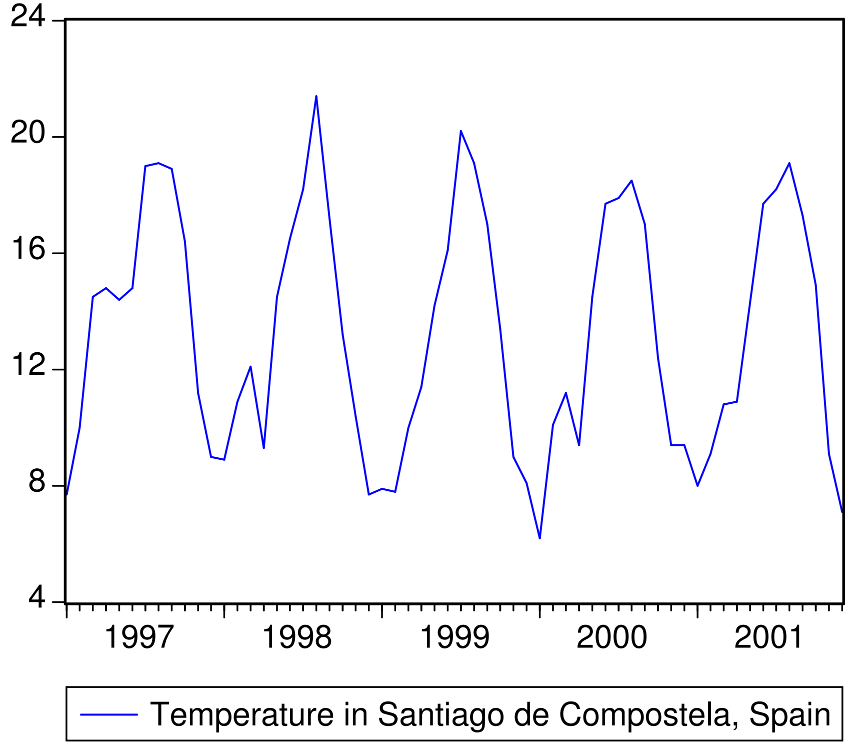
**Example 28.** *The mean of the random walk variables seem to be constant and near zero. This can be deduced from the equation of the process, (20), since for the first variable we can write* $E[z_1] = 0 + E(a_1) = 0$ *and for the rest we can use that if the expectation of* $z_{t-1}$ *is zero so is that of* $z_t$*, since:* $E[z_t] = E[z_{t-1}] + E(a_t) = 0$.



# Properties of marginal distributions - Examples

**Example 29.** *The population series in the left figure clearly shows an unstable mean. The temperature series in the right figure does not have a constant mean either, since the average temperature is different in different months of the year.*



# Properties of marginal distributions

▷ The **variance function** of the process gives the variance at each point in time:

$$Var(z_t) = \sigma_t^2 \tag{22}$$

and we say that the process is **stable in the variance** if the variability is constant over time.

▷ A process can be stable in the mean but not in the variance and vice versa. For example, the random walk has a constant mean, as we saw earlier, but the variance is not constant over time. See the figure in Example 27. In fact, let us assume that the variance of $a_t$ is $\sigma^2$. Hence, the variable $z_2$ will have that:

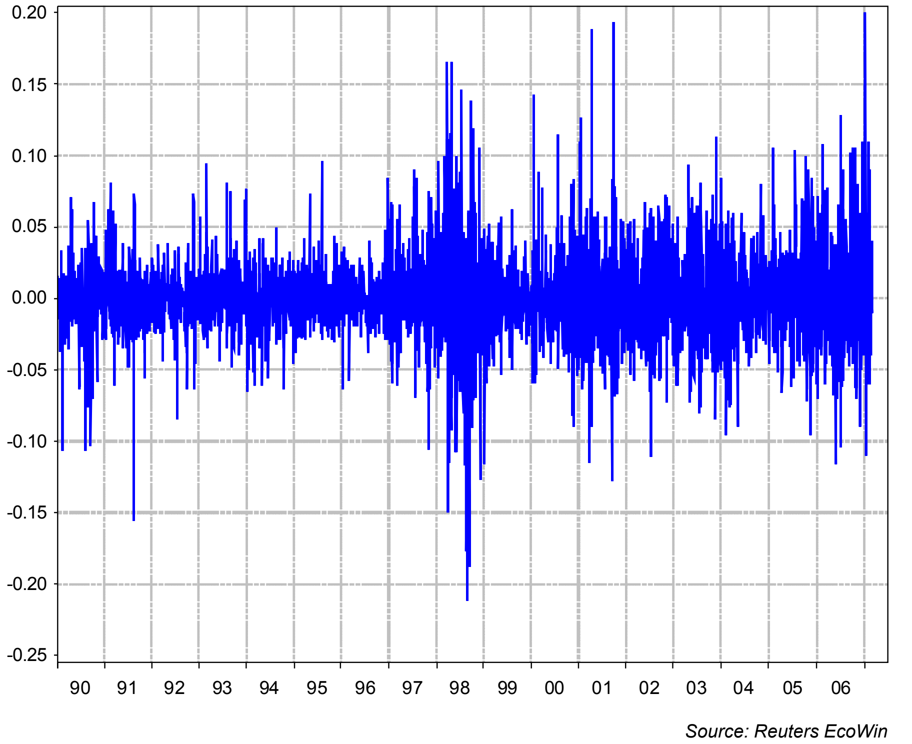
$$Var(z_2) = E(z_2^2) = E(z_1^2 + a_2^2 + 2z_1a_2) = 2\sigma^2$$

since the variables $z_1$ and $a_2$ are independent because $z_1$ depends only on $a_1$, which is independent of $a_2$. In general, the variance of $z_t$ is $t\sigma^2$, and it increases linearly over time.

# Properties of marginal distributions - Examples

▷ When we have a single realization, the apparent variability of the series may be approximately constant or change over time.

**Example 30.** *The series in figure seems to have a constant mean but variability appears to be greater in certain periods than in others.*
*Thus, the variance of the series might not be constant (* **heteroscedatic series** *).*



# Properties of marginal distributions

▷ The structure of linear dependence between random variables is represented by the covariance and correlation functions. We use the term **autocovariance function** of the process to refer to the covariance between two variables of the process at any two given times:

$$\gamma(t, t+j) = Cov(z_t, z_{t+j}) = E[(z_t - \mu_t)(z_{t+j} - \mu_{t+j})] \tag{23}$$

In particular, we have

$$\gamma(t, t) = Var(z_t) = \sigma_t^2.$$

▷ The mean function and the autocovariance functions play the same role in a stochastic process as the mean and variance for a scalar variable.

▷ The autocovariances have dimensions, the squares of the series, thus it is not advisable to use them for comparing series measured in different units.

# Properties of marginal distributions

▷ We can obtain a non-dimensional measurement of the linear dependence generalizing on the idea of the linear correlation coefficient between two variables.

▷ The **autocorrelation coefficient** of order $(t, t+j)$ is the correlation coefficient between the variables $z_t$ and $z_{t+j}$ and the **autocorrelation function** is the function of the two plots that describe these coefficients for any two values of the variables. This function is

$$\rho(t, t+j) = \frac{Cov\,(t, t+j)}{\sigma_t \sigma_{t+j}} = \frac{\gamma(t, t+j)}{\gamma^{1/2}(t, t)\gamma^{1/2}(t+j, t+j)}. \qquad (24)$$

In particular, we have

$$\rho(t, t) = \frac{Cov\,(t, t)}{\sigma_t \sigma_t} = 1.$$

# Properties of conditional distributions

▷ In addition to the study of marginal distributions in stochastic processes it is usually of great interest to study conditional distributions.

▷ An important group of processes are the **Markov processes** (or Markovian) which have the following property

$$f(z_{t+1}|z_t, ..., z_1) = f(z_{t+1}|z_t), \qquad t = 1, 2, ...,$$

that is, the distribution of the random variable at any time given the previous values of the process depends only on the last value observed.

▷ Intuitively a process is Markovian if, knowing the current value of the process, the distribution of the future values depends only on this value and not on the path followed up to that point.

**Example 31.** *The Gaussian random walk, shown in (20), is a Markovian process where*

$$f(z_{t+1}|z_t, ..., z_1) = f(z_{t+1}|z_t) = \mathcal{N}(z_t, \sigma^2).$$

# Properties of conditional distributions

▷ A property that is weaker than the Markovian is that the conditional expectation depends solely on the last value observed. In particular, when the process verifies that:

$$E(z_t|z_{t-1},...,z_1) = E(z_t|z_{t-1}) = z_{t-1} \tag{25}$$

the process is called a **Martingale**.

▷ A random walk, for example, is a martingale.

▷ We can see that a martingale is not necessarily a Markov process, because for that not only the expectation but rather the whole distribution must depend only on the last value observed.

▷ Moreover, a Markov process does not imply that condition (25) is satisfied, since $E(z_t|z_{t-1},...,z_1)$ can be any function $g(z_{t-1})$ of the last value observed.

# Properties of conditional distributions

▷ It is interesting to notice the differences between conditional distributions and the marginal distributions studied in this section.

- The marginal distribution of $z_t$ represents what we know about a variable without knowing anything about its trajectory until time $t$.
- The conditional distribution of $z_t$ given $z_{t-1}, ..., z_{t-k}$ represents what we know about a variable when we know the $k$ previous values of the process.
- For example, in the random walk (20) the marginal mean is constant and equal to zero whereas the conditional mean is equal to the last observed value and the marginal variance grows with time, while the conditional variance is constant.
- In time series conditional distributions are of greater interest than marginal ones because they define the predictions that we can make about the future knowing the past as well as the uncertainty of these predictions.

# Stationary processes

▷ Obtaining the probability distributions of the process is possible in some situations, for example with climatic variables, where we can assume that each year a realization of the same process is observed, or techniques that can be generated in a laboratory.

▷ Nevertheless, in many situations of interest, such as with economic or social variables, we can only observe one realization of the process.

▷ For example, if we observe the series of yearly growth in the wealth of a country it is not possible to go back in time to generate another realization.

▷ The stochastic process exists conceptually, but it is not possible to obtain successive samples or independent realizations of it.

▷ In order to be able to estimate the "transversal" characteristics of the process (means, variance, etc.) from its "longitudinal" evolution we must assume that the "transversal" properties (distribution of the variables at each instant in time) are stable over time. This leads to the concept of **stationarity**.

# Stationary processes

▷ We say that a stochastic process (time series) is **stationary** in the **strict sense** if:

(1) the marginal distributions of all the variables are identical;

(2) the finite-dimensional distributions of any set of variables depend only on the lags between them.

▷ The first condition establishes that, in particular, the mean and the variance of all the variables are the same.

▷ The second condition imposes that the dependence between variables depends only on their lags, that is, v.g. the same dependency exists between the variables $(z_t, z_{t+j})$ as between the variables $(z_{t+k}, z_{t+j+k})$.

▷ These two conditions can be summarized by establishing that the joint distribution of any set of variables is not changed if we translate the variables in time, that is:

$$F(z_i, z_j, \ldots, z_k) = F(z_{i+h}, z_{j+h}, \ldots, z_{k+h}).$$

# Stationary processes

▷ Strict stationarity is a very strong condition, since to prove it we must have the joint distributions for any set of variables in the process. A weaker property, but one which is easier to prove, is **weak sense stationarity**. A process is **stationary** in the **weak sense** if, for all $t$:

1. $\mu_t = \mu = cte,$

2. $\sigma_t^2 = \sigma^2 = cte,$

3. $\gamma(t, t-k) = E\left[(z_t - \mu)(z_{t-k} - \mu)\right] = \gamma_k \qquad k = 0, \pm 1, \pm 2, \ldots$

▷ The first two conditions indicate that the mean and variance are constant.

▷ The third indicates that the covariance between two variables depends only on their separation.

# Stationary processes

▷ In a stationary process the autocovariances and autocorrelations depend only on the lag between the observations and, in particular, the relationship between $z_t$ and $z_{t-k}$, is always equal to the relationship between $z_t$ and $z_{t+k}$.

▷ As a result, in stationary processes:

$$Cov(z_t, z_{t+k}) = Cov(z_{t+j}, z_{t+k+j}) = \gamma_k, \qquad j = 0 \pm 1, \pm\ 2, ...$$

and for autocorrelations as well since:

$$\rho_k = \frac{Cov(z_t, z_{t-k})}{\sqrt{var(z_t) var(z_{t-k})}} = \frac{\gamma_k}{\gamma_0}.$$

▷ To summarize, in stationary processes, we have that $\gamma_0 = \sigma^2$, and $\gamma_k = \gamma_{-k}$. For the autocorrelations $\rho_k = \rho_{-k}$.

# Stationary processes

▷ We use the term **covariance matrix** of the stationary process of order $k$, $\Gamma_k$, for the square and symmetric matrix of order $k$ that has the variances in its principal diagonal and in the next diagonals the autocovariances:

$$\Gamma_k = E\left[\begin{bmatrix} z_t - \mu \\ z_{t-1} - \mu \\ \cdots \\ z_{t-k} - \mu \end{bmatrix}\begin{bmatrix} z_t - \mu & z_{t-1} - \mu & \cdots & z_{t-k} - \mu \end{bmatrix}\right] =$$

$$= \begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{k-1} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{k-2} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{k-1} & \gamma_{k-2} & \cdots & \gamma_0 \end{bmatrix} \qquad (26)$$

▷ The square matrices which, like this matrix $\Gamma_k$, have the same elements in each diagonal are called Toeplitz matrices.

# Stationary processes

▷ We use the term **autocorrelation function** (***ACF***) to refer to the representation of the autocorrelation coefficients of the process as a function of the lag and the term **autocorrelation matrix** for the square and symmetric Toeplitz matrix with ones in the diagonal and the autocorrelation coefficients outside the diagonal:

$$R_k = \begin{bmatrix} 1 & \rho_1 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \cdots & \rho_{k-2} \\ \vdots & \vdots & \cdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \cdots & 1 \end{bmatrix}. \qquad (27)$$

▷ Weak stationarity does not guarantee complete stability of the process. For example, the distribution of the variables $z_t$ may be changing over time. Nevertheless, if we assume that these variables *jointly* have a normal $n$-dimensional distribution, then weak and strict stationarity coincide.

▷ From this point on, for simplicity's sake, we will use the expression **stationary** process to refer to a stationary process in the weak sense.

# Combinations of stationary processes

▷ An important property of stationary processes is that they are stable under linear combinations, that is, the processes obtained using linear combinations of stationary processes are also stationary.

▷ Let $\mathbf{z}_t = (z_{1t}, ..., z_{kt})$ be a vector of $k$ stationary processes where we assume that the autocovariances depend only on the lag, and the covariances between two components at two times depend only on the two components considered and the lag between the times. In these conditions, the vector series is stationary.

▷ Let us consider the scalar process defined by the vector of constants $\mathbf{c}' = (c_1, ..., c_k)$ :

$$y_t = \mathbf{c}'\mathbf{z}_t = c_1 z_{1t} + ... + c_k z_{kt}$$

which will be a linear combination of the components of the vector $\mathbf{z}_t$.

▷ We will prove that $y_t$ is also stationary.

# Combinations of stationary processes

- The expectation of this process is

$$E(y_t) = c_1 E(z_{1t}) + ... + c_k E(z_{kt}) = \mathbf{c}'\mu$$

where $\mu = (\mu_{1,...,} \mu_k)$ is the vector of means of the components. Since the expectations $E(z_{it}) = \mu_i$ are constants, so is $E(y_t)$.

- The variance of the process $y_t$ is:

$$var(y_t) = E(\mathbf{c}'(\mathbf{z}_t - \mu)(\mathbf{z}_t - \mu)\mathbf{c}) = \mathbf{c}'\Gamma_z \mathbf{c} \tag{28}$$

where $\Gamma_z$ is the covariance matrix between the components of the vector at the same time. Since the components are stationary, the covariance matrix between them is also constant.

# Combinations of stationary processes

- Analogously, it is proved that

$$cov(y_t y_{t+k}) = \mathbf{c}'\Gamma_z(k)\mathbf{c}$$

where $\Gamma_z(k)$ contains the covariances between the components at different times which, by hypothesis, depend only on the lag.

▷ Therefore, process $y_t$ is stationary.

▷ If we define the linear combination of the lagged values of the scalar process, $z_t$, using:

$$y_t = c_1 z_t + c_2 z_{t-1} + ... + c_k z_{t-k} = \mathbf{c}'\mathbf{z}_{t,k}$$

where $\mathbf{z}_{t,k} = (z_t, z_{t-1}, ..., z_{t-k})'$, the variance of the variable $y_t$ is given by $\mathbf{c}'\Gamma_k\mathbf{c}$, where $\Gamma_k$ has the expression (26), and it is a non-negative number. This implies that the autocovariance matrix of order $k$ of the process $z_t$ is non-negative definite.

# White noise process

▷ A very important stationary process is that defined by the conditions:

1a. $E[z_t] = 0$, for $t = 1, 2, ...$

2a. $Var(z_t) = \sigma^2$ for $t = 1, 2, ...$

3a. $Cov(z_t, z_{t-k}) = 0$ for $k = \pm 1, \pm 2, ...$

which is called the **white noise process**.

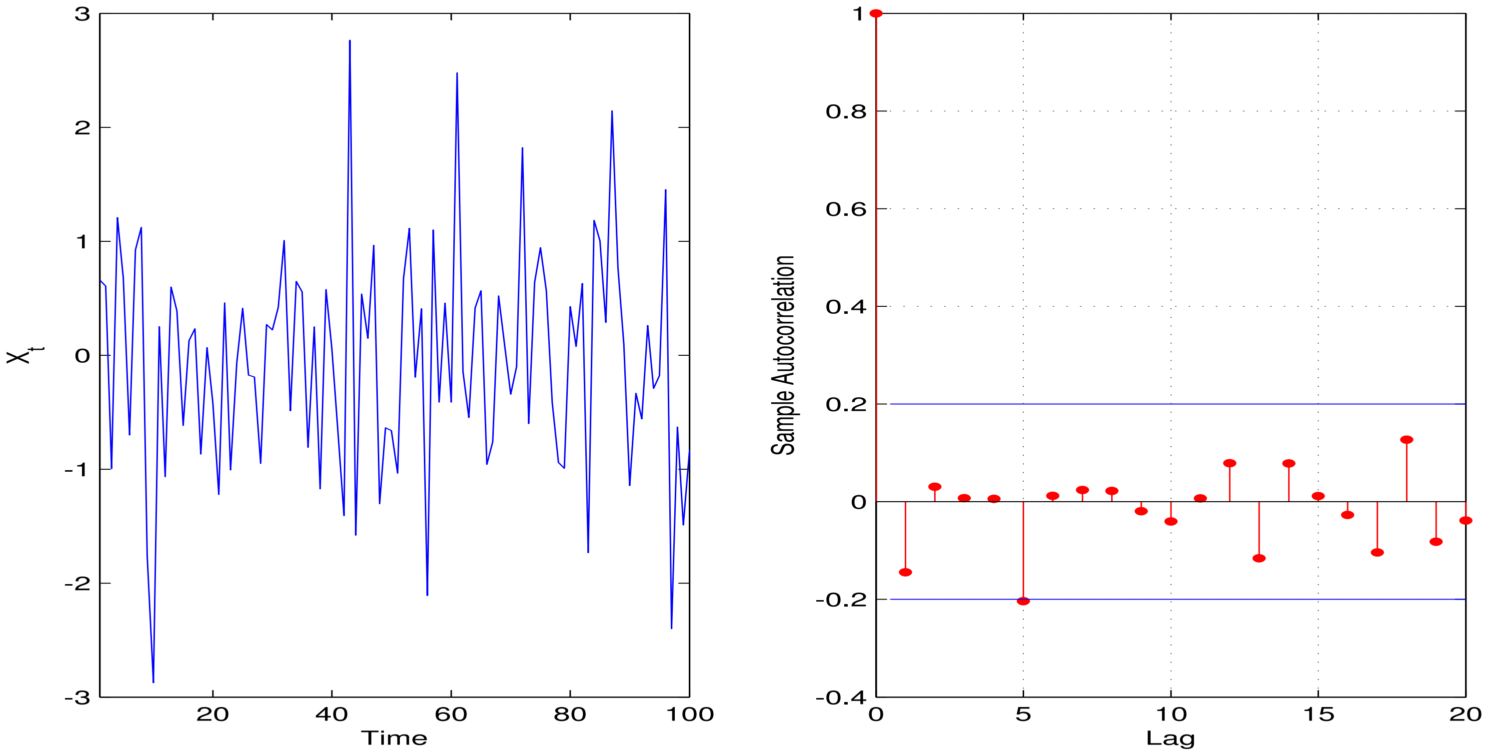
▷ First condition establishes that the expectation is always constant and equal to zero.

▷ Second condition establishes that variance is constant.

▷ Third condition establishes that the variables of the process are uncorrelated for all lags.

# White noise process - Example

**Example 32.** *If we generate random normal numbers of mean zero and constant variance with a computer and we place them in a sequence, we will have a white noise process.*




▷ In these processes knowing past values provides no information about the future since the process has "no memory".

# White noise process

▷ An equivalent condition to define a white noise is to assume a stationary process with marginal finite variance $\sigma^2$ and the condition:

1b. $E(z_t|z_{t-1}, ..., z_1) = 0$ for every $t$,

since it can be proved that the conditions (1a) and (3a) are then verified automatically.

▷ When the process is defined using condition (1b) it is usually called, for historical reasons, the **martingale difference process**.

▷ The reason for this is that if process $y_t$ is a martingale, defined by the condition (25), process $z_t = y_t - y_{t-1}$ verifies:

$$E(z_t|z_{t-1}, ..., z_1) = E(y_t|z_{t-1}, ..., z_1) - E(y_{t-1}|z_{t-1}, ..., z_1) = 0,$$

then $z_t$ verifies (1b) and its a martingale difference process.

# White noise process

▷ A white noise process is not necessarily stationary in the strict sense nor does it have to be formed by independent variables, since only non-correlation is required.

▷ If we impose the additional condition that the variables of the process are independent, and not just uncorrelated, we call this a **strict white noise** process.

▷ If we assume that the variables have a normal distribution, the non-correlation guarantees independence, and the process is strict white noise. We call the resulting process the **normal white noise** process.

▷ Normality is a strong condition and we can have strict white noise processes with variables that have non-normal distributions. For example, a process of uniform independent variables is a strict white noise process, but not a normal white noise one.

# Estimating the moments of stationary processes

▷ Let us assume a stationary process with mean $\mu$, variance $\sigma^2$ and covariances $\gamma_k =$ from which we observe one realization $(z_1, ..., z_T)$. We are going to study how to estimate the mean, variance, covariances and autocorrelations of the process from this single realization.

## Estimating the mean

▷ An unbiased estimator of the population mean is the **sample mean**. Letting $\overline{z}$ be the sample mean:

$$\overline{z} = \frac{\sum_{t=1}^{T} z_t}{T},$$

it is proved that:

$$E(\overline{z}) = \frac{\sum_{t=1}^{T} E(z_t)}{T} = \mu.$$

# Estimating the moments of stationary processes

## Estimating the mean

▷ For independent data the variance of the sample mean as an estimator of the population mean is $\sigma^2/T$. As a result, when the sample size is increased the mean square error estimation, given by

$$E(\overline{z} - \mu)^2$$

which coincides with the variance of $\overline{z}$, tends to zero.

▷ In a stationary stochastic process this property is not necessarily true, and it is possible that when the sample size is increased the variance of the estimate of the mean does not tend to zero.

Example

# Estimating the moments of stationary processes

## Estimating the mean

▷ From here on we will assume that the process is **ergodic** for the estimation of the mean and this implies that:

$$\lim_{T\to\infty} E(\overline{z}-\mu)^2 \to 0.$$

▷ To identify the conditions for a process to be ergodic for the estimation of the mean, we are going to calculate the mean square error of the sample mean for stationary processes:

$$MSE(\overline{z}) = var(\overline{z}) = E(\overline{z}-\mu)^2 = \frac{1}{T^2}E(\sum_{t=1}^{T}(z_t-\mu))^2$$

which we can write as:

$$var(\overline{z}) = \frac{1}{T^2}\left[\sum\nolimits_{t=1}^{T} E(z_t-\mu)^2 + 2\sum\nolimits_{i=1}^{T}\sum_{j=i+1}^{T} E((z_i-\mu)(z_j-\mu))\right].$$

# Estimating the moments of stationary processes

## Estimating the mean

▷ The first summation within the square brackets is $T\sigma^2$. The second double summation contains $T-1$ times the covariances of order one, $T-2$ times the covariances of order 2 and, in general, $T-i$ times those of order $i$. As a result, the variance of the sample mean can be written as:

$$var(\overline{z}) = \frac{1}{T}\left[\sigma^2 + 2\sum_{i=1}^{T-1}(1-\frac{i}{T})\gamma_i\right] \quad (29)$$

▷ The condition for $var(\overline{z})$ to tend to zero when $T$ increases is that the summation converges to a constant when $T$ increases. A necessary (although not sufficient) condition for the sum to converge is:

$$\lim_{i\to\infty}\gamma_i \to 0$$

which assumes that the dependence between observations tends to zero when the lag is increased.

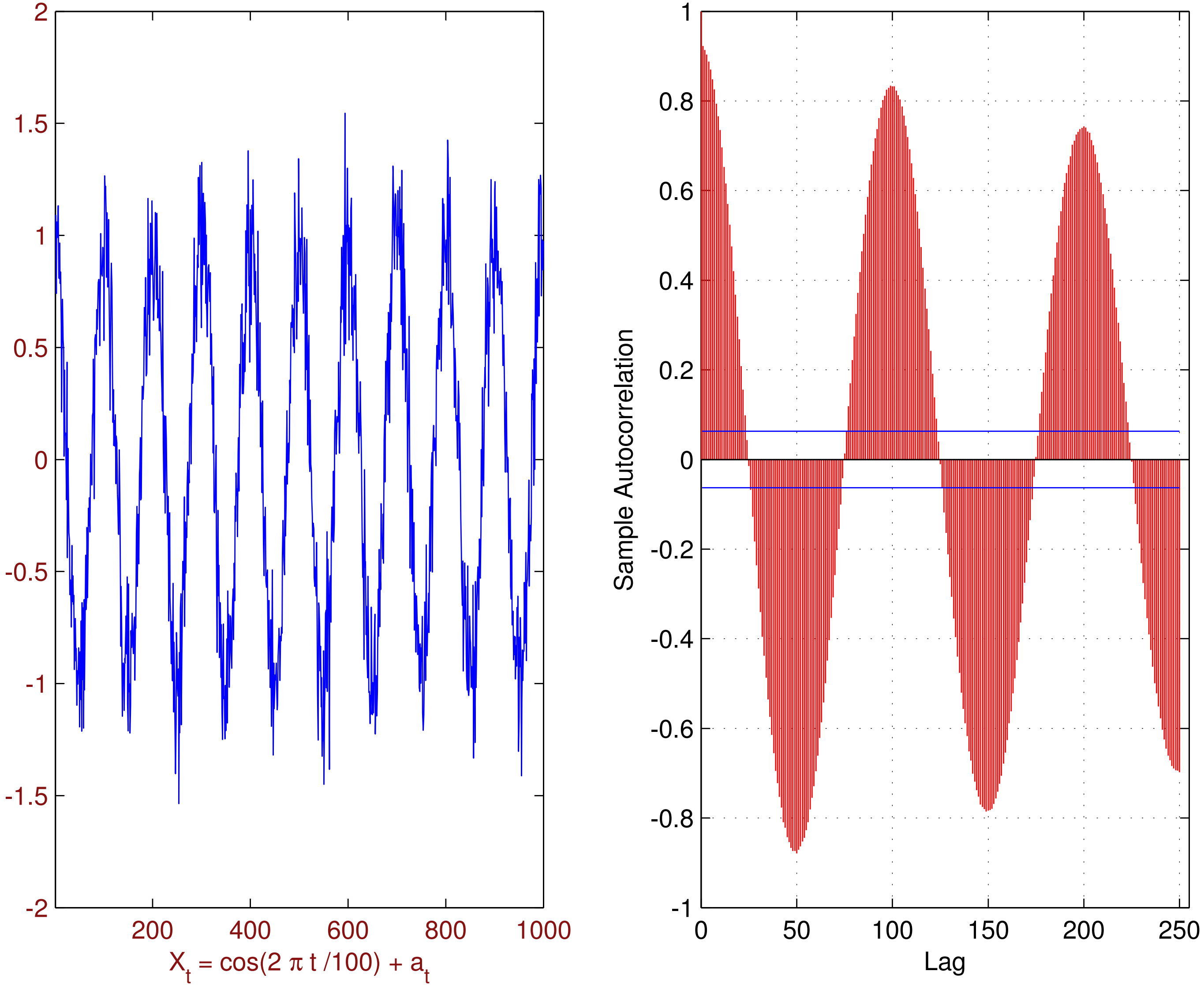
# Estimating the moments of stationary processes

▷ To summarize, the property of ergodicity, which can be applied to the estimation of any parameter, is satisfied if the new observations of the process provide additional information on the parameter, such that when the sample size increases the estimation error tends to zero.

▷ In the estimation of the mean this does not occur if there is a dependence so strong that new observations are predictable from the past and do not provide new information for estimating the mean.

▷ From here on we will assume that the processes we consider are ergodic, meaning that in practice we have eliminated possible deterministic trend terms of type $A\cos(\omega t + \delta)$ that might exist.

# Example of non-ergodic processes



# Estimating the moments of stationary processes

## Estimation of autocovariances and autocorrelations

▷ If the mean of the process is known, the estimator of the autocovariances of order $k$ is:

$$\widetilde{\gamma}_k = \frac{1}{T-k}\sum\nolimits_{t=k+1}^{T}(z_t-\mu)(z_{t-k}-\mu) \tag{30}$$

and it is easy to prove when expectations are taken that this estimator is unbiased for estimating $\gamma_k = E((z_t-\mu)(z_{t-k}-\mu))$.

▷ Nevertheless, when $\mu$ is unknown and we replace it with its estimator, $\overline{z}$, it is proved that the resulting estimator is biased.

▷ An alternative estimator of $\gamma_k$, which has better properties when $\mu$ is unknown, is:

$$\widehat{\gamma}_k = \frac{1}{T}\sum\nolimits_{t=k+1}^{T}(z_t-\overline{z})(z_{t-k}-\overline{z}). \tag{31}$$

Although $\widehat{\gamma}_k$ is also a biased estimator of the population autocovariance it has less squared error of estimation than the previous one.

# Estimating the moments of stationary processes

## Estimation of autocovariances and autocorrelations

▷ An additional advantage to using the estimator (31) is that the sample autocovariance matrix:

$$\widehat{\Gamma}_k = \begin{bmatrix} \widehat{\gamma}_0 & \widehat{\gamma}_1 & \cdots & \widehat{\gamma}_{k-1} \\ \widehat{\gamma}_1 & \widehat{\gamma}_0 & \cdots & \widehat{\gamma}_{k-2} \\ . & . & \cdots & . \\ \widehat{\gamma}_{k-1} & \widehat{\gamma}_{k-2} & \cdots & \widehat{\gamma}_0 \end{bmatrix}$$

is always non-negative definite.

▷ This may not occur if, instead of dividing by the sample size, we do so using the number of terms in the sum, as in estimator (30).

▷ Lets remember this property is necessary for the estimated covariances to be able to correspond to a stationary process.

# Estimating the moments of stationary processes

## Estimation of autocovariances and autocorrelations

▷ The autocorrelations are estimated by

$$r_k = \widehat{\gamma}_k / \widehat{\gamma}_0$$

and we can estimate the autocorrelation function representing the estimated correlation coefficients as a function of the lag. This representation is called a **correlogram** or **sample autocorrelation function**.

▷ The vector of sample autocorrelations, $\mathbf{r} = (r_1, ..., r_k)'$ has an approximately normal distribution for large $T$ with mean $\rho$, the theoretical vector of autocorrelations, and covariance matrix $\mathbf{V}_\rho / T$.

▷ The terms of the matrix $\mathbf{V}_\rho$ are given by Bartlett's formula.

# Estimating the moments of stationary processes

## Estimation of autocovariances and autocorrelations

▷ Bartlett's formula have simple expression if the process only has the first $q$ autocorrelation coefficients different from zero, the variances of the estimated autocorrelations are approximated by:

$$var(r_k) = \frac{T-k}{T(T+2)}(1 + 2\sum_{j=1}^{q}\rho_j^2), \qquad k > q \tag{32}$$

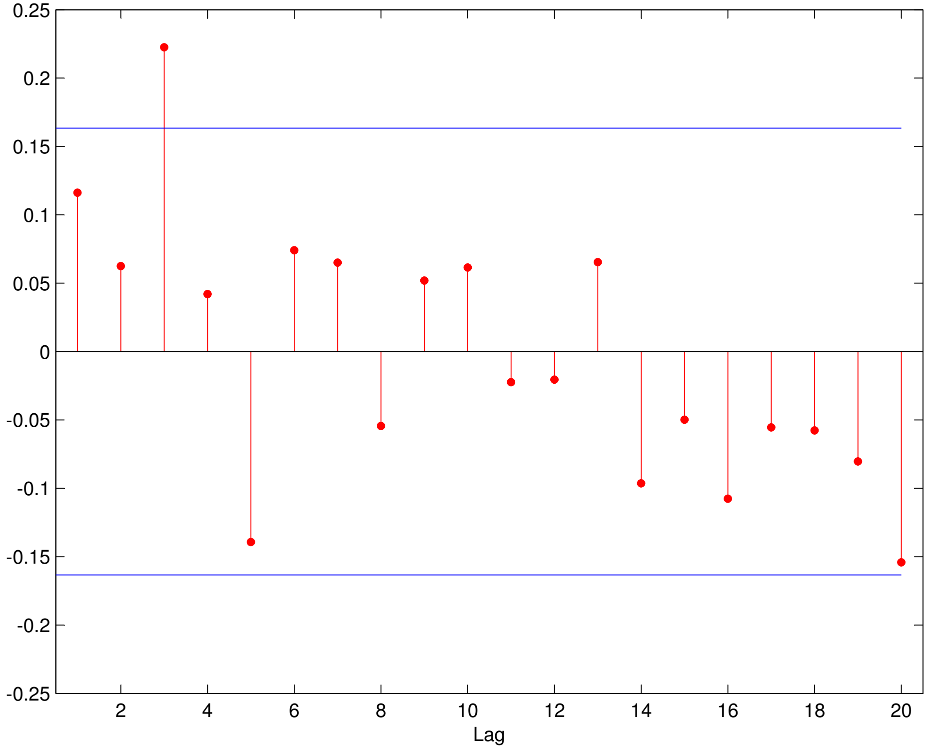
Therefore, if all the autocorrelations of the process are null, the asymptotic variance of the estimated or sample autocorrelation coefficients is:

$$var(r_k) = \frac{T-k}{T(T+2)},$$

which can be approximated, for large $T$, by $1/T$.

# Estimation of autocorrelations - Examples

**Example 33.** *The figure shows the correlogram of a white noise series including two lines parallel to the* $x$ *axis at* $\pm 2/\sqrt{T}$. *These lines provide, approximately, a 95% confidence interval for the sample autocorrelation coefficients if the series has been generated by a white noise process.*

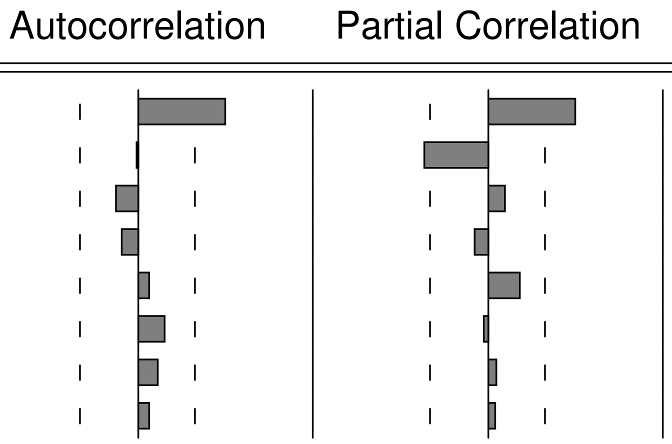


# Estimation of autocorrelations - Examples

**Example 34.** *We are going to calculate the sample autocorrelation function or correlogram for the data on the leagues sailed by Columbus.*

Correlogram of DAILYLEAGUES

Date: 01/28/08 Time: 16:28
Sample: 1 36
Included observations: 34

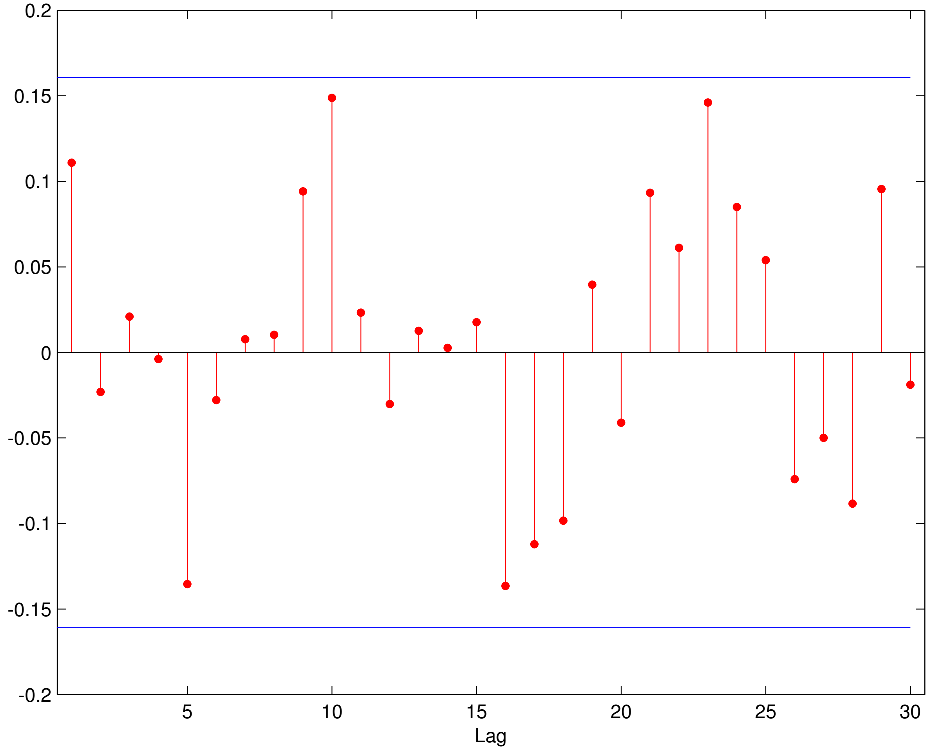



| | AC | PAC | Q-Stat | Prob |
|---|---|---|---|---|
| 1 | 0.520 | 0.520 | 10.038 | 0.002 |
| 2 | -0.003 | -0.375 | 10.038 | 0.007 |
| 3 | -0.130 | 0.106 | 10.710 | 0.013 |
| 4 | -0.091 | -0.079 | 11.047 | 0.026 |
| 5 | 0.071 | 0.193 | 11.260 | 0.046 |
| 6 | 0.158 | -0.019 | 12.352 | 0.055 |
| 7 | 0.116 | 0.056 | 12.964 | 0.073 |
| 8 | 0.073 | 0.043 | 13.214 | 0.105 |

Only the first autocorrelation coefficient seems to be significant, indicating that the leagues travelled on a given day depend on those sailed the day before, but not on the days prior to that.

# Estimation of autocorrelations - Examples

**Example 35.** *The figure shows the autocorrelation function for the series of the Madrid Stock Exchange. Notice that all the autocorrelation coefficients are small, and within the bands of* $2/\sqrt{T}$. *This suggests a white noise process, where past performance gives us no information for predicting future performance.*



# Summary

- We have sawn the basic concepts of stochastic processes: realizations, mean function, variance function, autocorrelation function, stationarity, ergodicity, martingale, etcetera.

- We have consideres different definitions of stationary processes (in strict and weak sense).

- We have studied the simplest stationary process: White Noise.

- We have sawn how to estimate the mean, the autocovariances and the autocorrelations functions (theoretically and using EViews).