# Time Series Analysis

## Integrated and long-memory processes

Andrés M. Alonso Carolina García-Martos

Universidad Carlos III de Madrid

Universidad Politécnica de Madrid

June – July, 2012

**Outline:**

- Introduction
- Integrated processes
- The random walk
- The simple exponential smoothing process
- Integrated process of order two
- ARIMA processes
- Integrated processes and trends
- Long-memory processes

**Recommended readings:**

▷ Chapter 6 of Brockwell and Davis (1996).

▷ Chapter 15 of Hamilton (1994).

# Introduction

▷ In this section we begin our study of **non-stationary processes**:

- A process can be non-stationary in the mean, the variance, the autocorrelations, or in other characteristics of the distribution of variables.
  - When the level of the series is not stable in time, in particular showing increasing or decreasing trends, we say that the series is not stable in the mean.
  - When the variability or autocorrelations change with time, we say that the series is not stationary in the variance or autocovariance.
  - Finally, if the distribution of the variable at each point in time varied over time, we say that the series is not stationary in distribution.
- The most important non-stationary processes are the integrated processes, which have the basic property that by differentiating them we obtain stationary processes.

# Introduction

▷ An important property that distinguishes integrated processes from stationary ones is the form in which dependency disappears over time.

- In ARMA stationary processes the autocorrelations diminish geometrically, and practically reach zero in few lags.
- In the integrated processes, the autocorrelations diminish linearly over time and it is possible to find autocorrelation coefficients different from zero even for very high lags.

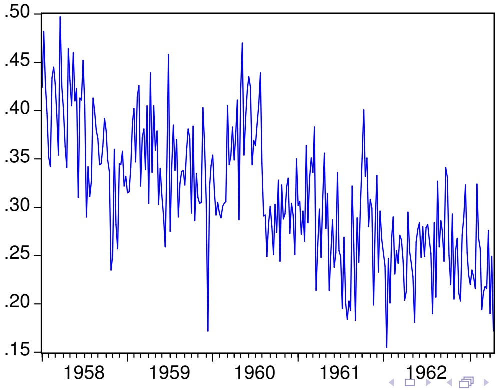
▷ There is a class of stationary processes where the autocorrelations decay much more slowly over time than in the case of the ARMA processes or in the integrated processes. These are known as long-memory processes.

▷ In addition to their theoretical interest, these processes can closely approximate behavior observed in long climatological or financial series.

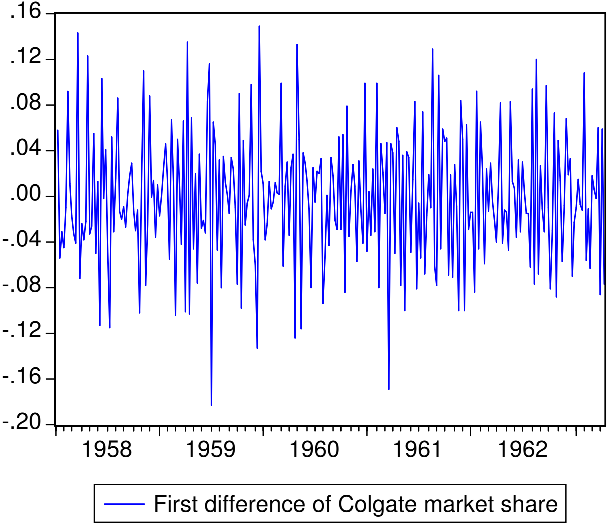
▷ Most real series are not stationary, and their average level varies over time.

## Example 46

*The figure shows the weekly market share for Colgate toothpaste. The series, which we will denote by $z_t$, shows a clearly decreasing trend and thus is not stationary.*



▷ The figure shows the first difference in this series, that is, the series of variations in market share from one week to the next. If we let $w_t = \nabla z_t$ denote this new series, we see that its values oscillate around a constant mean and seem to correspond to a stationary series.



# Integrated processes

▷ We conclude that the series $z_t$ of Colgate market share seems to be an integrated series, which is transformed into a stationary one by means of differentiation.

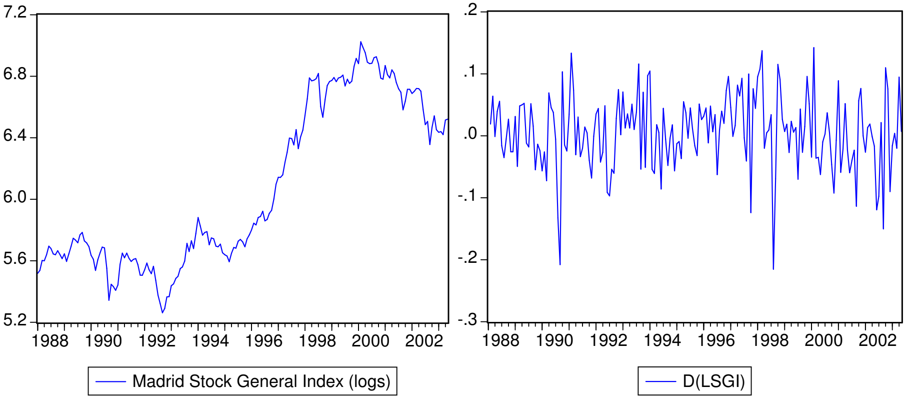
▷ We say then that it is **integrated of order one**; the number of differences needed to obtain a stationary process being the order of integration.

▷ Economic series are not usually stationary but their relative differences, or the differences when we measure the variable in logarithms, are stationary.

## Example 47

*The next figures show another non-stationary series and its difference, the logarithm of the Madrid Stock Exchange general index:*

- *The series is not stationary in the mean since its level varies over time.*
- *If we take a difference in this series we obtain the series of the Madrid Stock Exchange returns using the General Index. It seems that the monthly return is stable over time.*

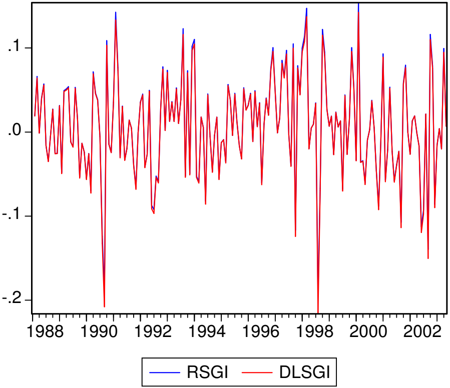



▷ Examples 46 and 47 show first order integrated processes, that is, they are non-stationary but their growth, or first differences, are stationary.

▷ If $z_t$ is the series of the general stock index, then $\nabla \log z_t$ is:

$$\nabla \log z_t = \log \frac{z_t}{z_{t-1}} = \log(1 + \frac{z_t - z_{t-1}}{z_{t-1}}) \approx \frac{\nabla z_t}{z_{t-1}}$$

and we prove that the first difference of the logarithm of a variable is approximately equal to its relative growth.




▷ When the original variable, $z_t$, is a series of prices, then the series $\nabla \log z_t$ is defined as the stock return.

# Integrated processes

▷ It is sometimes necessary to differentiate more than once to obtain a stationary process.

▷ For example, if we let $z_t$ denote the original series, and its variations (or growth) measured by $w_t = \nabla z_t$ are not stationary, but the variations of growth, measured by $\nabla^2 z_t$, are. That is, the series:

$$y_t = \nabla \omega_t = \omega_t - \omega_{t-1} = z_t - 2z_{t-1} + z_{t-2} = \nabla^2 z_t \tag{98}$$

which represents the growth of series $\omega_t$ is stationary.

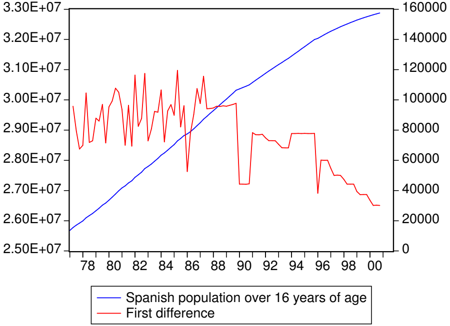
▷ Series $y_t$ is called the second difference of the original series, $z_t$, and we say that $z_t$ is **integrated of order 2**, or with stationary second order increments.

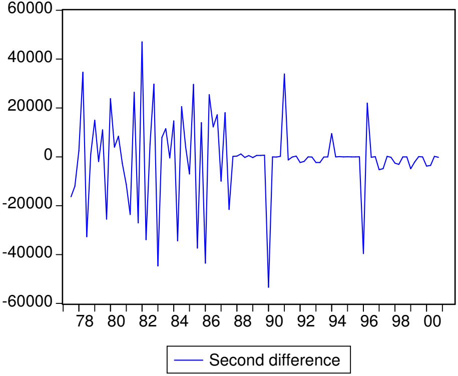
▷ Generalizing, we say that a process is **integrated of order h**$\geq 0$, and we denote it by $I(h)$, when upon differentiating it $h$ times a stationary process is obtained.

▷ A stationary process is, therefore, always $I(0)$.

## Example 48

*The figures show a series and its difference which are non-stationary. Nevertheless, the difference of this last series, which corresponds to the second difference of the original does have a stable level.*



# The random walk

▷ We have seen that finite MA processes are always stationary and that the AR are only so if the roots of $\phi(B) = 0$ lie outside the unit circle.

▷ Let us consider the AR(I):

$$z_t = c + \phi z_{t-1} + a_t. \tag{99}$$

- If $|\phi| < 1$ the process is stationary.
- If $|\phi| > 1$ it is easy to see that the process is explosive process and the values of the variable grow without limit to the infinite.

▷ Since explosive processes are not frequent in practice, values of the AR parameter greater than the unit are not generally useful for representing real series.

▷ An interesting case is when $|\phi| = 1$. Then, the process is not stationary, but neither is it explosive, and it belongs to a class of first order integrated processes.

# The random walk

▷ Indeed, it is straightforward to show that the first difference of an AR(1) with $|\phi| = 1$:

$$w_t = \nabla z_t = c + a_t,$$

is in fact a stationary process.

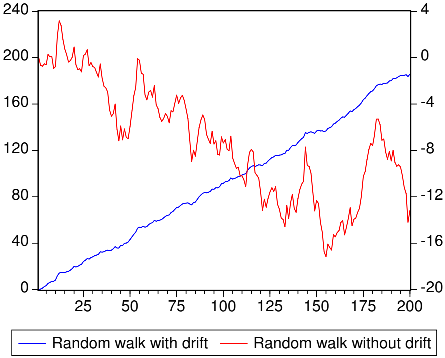
▷ This process is called a **random walk**, and was used as an example of a Markov process and a Martingale process.

▷ An important characteristic that distinguishes stationary from non-stationary processes is the role of the constants:

- In a stationary process the constant is unimportant, and we have subtracted its mean from the observations and worked with zero mean processes. Both the form of the process and its basic properties are the same whether the mean is zero or different from zero.
- Nevertheless, in a non-stationary process the constants, if they exist, are very important and represent some permanent property of the process.

## Example 49

*The figure shows two simulations of model (99). In the first $c = 0$, it is said that the process does not have drift, and the level of the series oscillates in time. In the second $c = 1$, it is said that the process has a drift equal to one, and the process shows a linear trend of slope $c$.*




▷ The level of the series is that of the previous period plus $c$, which produces deterministic linear growth.

Datafile `2randomwalk.xls`

# The random walk

▷ We see that the graph of the series is totally different in the two cases.

▷ To calculate the descriptive measures of this process we assume that it starts at $t = 0$. Then, successively replacing $z_t$ with $z_{t-1}$ we have:

$$z_t = ct + a_t + a_{t-1} + a_{t-2} + ... + a_1$$

and taking expectations:

$$E(z_t) = ct.$$

▷ If $c \neq 0$, the process has a mean that increases linearly over time, but if the process does not have drift, $c = 0$, the mean is constant and equal to zero.

▷ Its variance is:

$$Var\,(z_t) = E(a_t + a_{t-1} + a_{t-2} + ... + a_1)^2 = \sigma^2 t \qquad (100)$$

and we see that the variance increases over time and tends to the infinite with $t$. This property indicates that increasing time increases uncertainty about the situation of the process.

# The random walk

▷ To calculate the autocovariances we use the expression of the process for $t + k$:

$$z_{t+k} = c(t + k) + a_{t+k} + ... + a_t + ... + a_1$$

and using the notation $Cov(t, t + k) = Cov(Z_t, Z_{t+k})$ we have:

$$Cov(t, t + k) = E\left[(z_t - ct)(z_{t+k} - c(t + k))\right] = \sigma^2 t. \tag{101}$$

▷ Note that the autocovariances also increase over time and are not only a function of the lag, as in stationary processes, but that they depend on the times at which they were calculated.

▷ Particularly:

$$Cov(t, t - k) = \sigma^2(t - k) \neq Cov(t, t + k).$$

# The random walk

▷ The autocorrelation is obtained dividing (101) by the standard deviations of the variables that are obtained from the expression (100), which results in:

$$\rho(t, t+k) = \frac{t}{\sqrt{t(t+k)}} = (1 + \frac{k}{t})^{-1/2}. \qquad (102)$$

▷ This expression indicates that if $t$ is large, the coefficients of the autocorrelation function will be close to one and will decay in an approximately linear form with $k$.

▷ Indeed, if we assume that the process starts in the distant past such that $k/t$ is small, then the function $(1 + k/t)^{-1/2}$ can be approximated using a first order Taylor expansion such as:

$$\rho(t, t+k) \simeq 1 - \frac{k}{2t}. \qquad (103)$$

▷ This equation indicates that assuming fixed $t$, if we look at the autocorrelation as a function of $k$ we obtain a straight line with slope $(-1/2t)$.

# The random walk

▷ Taking logarithms in (102) we have

$$\log \rho(t, t+k) = -\frac{1}{2} \log(1 + \frac{k}{t}) \approx -\frac{k}{2t}$$

and we see that if we express the correlations in logarithms they also decay linearly with lag $k$.

▷ This behavior of the autocorrelation function of an integrated process can be thought of as an extreme case of an AR(1) when $\phi$ is close to the unit.

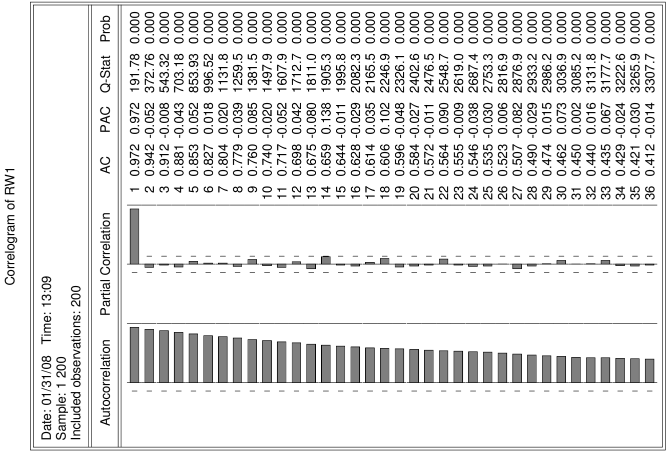
▷ Indeed, if $\phi = 1 - \epsilon$, with small $\epsilon$, we have

$$\rho_k = (1-\epsilon)^k \approx 1 - \epsilon k$$

for small $\epsilon$.

## Example 50

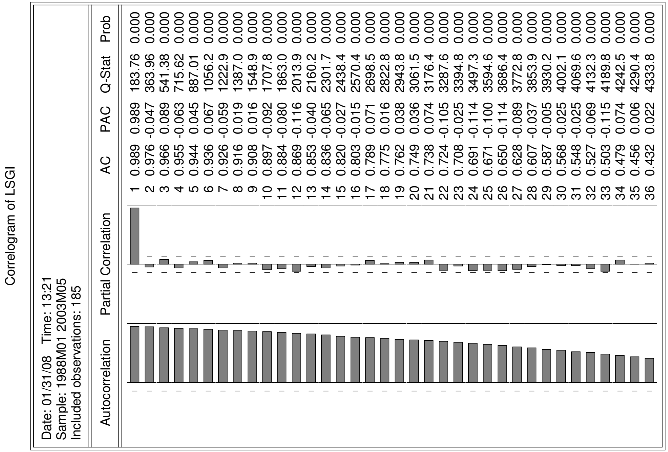
*The figure shows a graph of the autocorrelation function of a simulation of 200 observations that follow a random walk. Here, we can clearly see the linear decay of the ACF.*




| | AC | PAC | Q-Stat | Prob |
|---|---|---|---|---|
| 1 | 0.972 | 0.972 | 191.78 | 0.000 |
| 2 | 0.942 | -0.052 | 372.76 | 0.000 |
| 3 | 0.912 | -0.008 | 543.32 | 0.000 |
| 4 | 0.881 | -0.043 | 703.18 | 0.000 |
| 5 | 0.853 | 0.052 | 853.93 | 0.000 |
| 6 | 0.827 | 0.018 | 996.52 | 0.000 |
| 7 | 0.804 | 0.020 | 1131.8 | 0.000 |
| 8 | 0.779 | -0.039 | 1259.5 | 0.000 |
| 9 | 0.760 | 0.085 | 1381.5 | 0.000 |
| 10 | 0.740 | -0.020 | 1497.9 | 0.000 |
| 11 | 0.717 | -0.052 | 1607.9 | 0.000 |
| 12 | 0.698 | 0.042 | 1712.7 | 0.000 |
| 13 | 0.675 | -0.080 | 1811.0 | 0.000 |
| 14 | 0.659 | 0.138 | 1905.3 | 0.000 |
| 15 | 0.644 | -0.011 | 1995.8 | 0.000 |
| 16 | 0.628 | -0.029 | 2082.3 | 0.000 |
| 17 | 0.614 | 0.035 | 2165.5 | 0.000 |
| 18 | 0.606 | 0.102 | 2246.9 | 0.000 |
| 19 | 0.596 | -0.048 | 2326.1 | 0.000 |
| 20 | 0.584 | -0.027 | 2402.6 | 0.000 |
| 21 | 0.572 | -0.011 | 2476.5 | 0.000 |
| 22 | 0.564 | 0.090 | 2548.7 | 0.000 |
| 23 | 0.555 | -0.009 | 2619.0 | 0.000 |
| 24 | 0.546 | -0.038 | 2687.4 | 0.000 |
| 25 | 0.535 | -0.030 | 2753.3 | 0.000 |
| 26 | 0.523 | 0.006 | 2816.9 | 0.000 |
| 27 | 0.507 | -0.082 | 2876.9 | 0.000 |
| 28 | 0.490 | -0.029 | 2933.2 | 0.000 |
| 29 | 0.474 | 0.015 | 2986.2 | 0.000 |
| 30 | 0.462 | 0.073 | 3036.9 | 0.000 |
| 31 | 0.450 | 0.002 | 3085.2 | 0.000 |
| 32 | 0.440 | 0.016 | 3131.8 | 0.000 |
| 33 | 0.435 | 0.067 | 3177.7 | 0.000 |
| 34 | 0.429 | -0.024 | 3222.6 | 0.000 |
| 35 | 0.421 | -0.030 | 3265.9 | 0.000 |
| 36 | 0.412 | -0.014 | 3307.7 | 0.000 |

# Example 51

*The figure shows the autocorrelation function of series in Example 47. Notice that the correlations decline linearly, indicating a non-stationary process.*




| | AC | PAC | Q-Stat | Prob |
|---|---|---|---|---|
| 1 | 0.989 | 0.989 | 183.76 | 0.000 |
| 2 | 0.976 | -0.047 | 363.96 | 0.000 |
| 3 | 0.966 | 0.089 | 541.38 | 0.000 |
| 4 | 0.955 | -0.063 | 715.62 | 0.000 |
| 5 | 0.944 | 0.045 | 887.01 | 0.000 |
| 6 | 0.936 | 0.067 | 1056.2 | 0.000 |
| 7 | 0.926 | -0.059 | 1222.9 | 0.000 |
| 8 | 0.916 | 0.019 | 1387.0 | 0.000 |
| 9 | 0.908 | 0.016 | 1548.9 | 0.000 |
| 10 | 0.897 | -0.092 | 1707.8 | 0.000 |
| 11 | 0.884 | -0.080 | 1863.0 | 0.000 |
| 12 | 0.869 | -0.116 | 2013.9 | 0.000 |
| 13 | 0.853 | -0.040 | 2160.2 | 0.000 |
| 14 | 0.836 | -0.065 | 2301.7 | 0.000 |
| 15 | 0.820 | -0.027 | 2438.4 | 0.000 |
| 16 | 0.803 | -0.015 | 2570.4 | 0.000 |
| 17 | 0.789 | 0.071 | 2698.5 | 0.000 |
| 18 | 0.775 | 0.016 | 2822.8 | 0.000 |
| 19 | 0.762 | 0.038 | 2943.8 | 0.000 |
| 20 | 0.749 | 0.036 | 3061.5 | 0.000 |
| 21 | 0.738 | 0.074 | 3176.4 | 0.000 |
| 22 | 0.724 | -0.105 | 3287.6 | 0.000 |
| 23 | 0.708 | -0.025 | 3394.8 | 0.000 |
| 24 | 0.691 | -0.114 | 3497.3 | 0.000 |
| 25 | 0.671 | -0.100 | 3594.6 | 0.000 |
| 26 | 0.650 | -0.114 | 3686.4 | 0.000 |
| 27 | 0.628 | -0.089 | 3772.8 | 0.000 |
| 28 | 0.607 | -0.037 | 3853.9 | 0.000 |
| 29 | 0.587 | -0.005 | 3930.2 | 0.000 |
| 30 | 0.568 | -0.025 | 4002.1 | 0.000 |
| 31 | 0.548 | -0.025 | 4069.6 | 0.000 |
| 32 | 0.527 | -0.069 | 4132.3 | 0.000 |
| 33 | 0.503 | -0.115 | 4189.8 | 0.000 |
| 34 | 0.479 | 0.074 | 4242.5 | 0.000 |
| 35 | 0.456 | 0.006 | 4290.4 | 0.000 |
| 36 | 0.432 | 0.022 | 4333.8 | 0.000 |

▷ This behavior is observed both in the autocorrelations of the series as well as in the logarithms of the autocorrelations.

# The simple exponential smoothing process

▷ A non-stationary process that has been widely used for forecasting and which can be considered a generalization of the random walk is:

$$\nabla z_t = c + (1 - \theta B)\, a_t. \tag{104}$$

In this process if we make $\theta = 0$ we obtain a random walk. It can also be seen as an extreme case of the ARMA(1,1) when the autoregressive coefficient is the unit.

▷ For simplicity we assume $c = 0$ and $|\theta| < 1$ and inverting the MA part, this process is written:

$$z_t = (1 - \theta) z_{t-1} + \theta(1 - \theta) z_{t-2} + \theta^2 (1 - \theta) z_{t-3} + \ldots + a_t. \tag{105}$$

▷ In the usual case where $0 < \theta < 1$, all the coefficients of the above AR($\infty$) are positive and add up to the unit. These coefficients form an infinite geometric progression with the first term $(1 - \theta)$ and ratio $\theta$, hence their sum is one.

# The simple exponential smoothing process

▷ The equation (105) indicates that the observed value at each time is the weighted mean of the values of the series at earlier times, with coefficients that decay geometrically with the lag:

- If $\theta$ is close to the unit the mean is calculated using many coefficients, but each of them with little weight.
  - For example, if $\theta = .9$ the weights of the lags are .1, .09, .081, . . . and the weights decay very slowly.
- If $\theta$ is close to zero, the mean is calculated by weighting only the last values observed.
  - For example, if $\theta = .1$ the weights of the lags are .9, .09, .009,. . . and practically only the the last lags are taken into account, with very different weights.
- This is known as the **simple exponential smoothing process** and was introduced intuitively in the first session.

# The simple exponential smoothing process

▷ To calculate the descriptive measures of the process we assume that it started at $t = 0$, with $z_0 = a_0 = 0$ and $c = 0$. Thus:

$$\begin{array}{rcl} z_1 & = & a_1 \\ z_2 & = & z_1 + a_2 - \theta a_1 = a_2 + (1-\theta)\, a_1 \\ \vdots & \vdots & \qquad\qquad \vdots \\ z_t & = & a_t + (1-\theta)\, a_{t-1} + ... + (1-\theta)\, a_1. \end{array}$$

▷ The mean of the process is zero, as in the random walk when $c = 0$. Its variance is:

$$Var(z_t) = \sigma^2(1 + (t-1)(1-\theta)^2)$$

and the process is non-stationary in the variance, since the function depends on the time.

# The simple exponential smoothing process

▷ The covariances are calculated as:

$$Cov(t, t+k) = E(z_t z_{t+k}) =$$
$$= E\left[(a_t + (1-\theta)\sum_{i=1}^{t-1} a_{t-i})(a_{t+k} + (1-\theta)\sum_{i=1}^{t+k-1} a_{t+k-i})\right]$$

and, taking into account that $E(a_i a_j) = 0$, if $i \neq j$, we obtain:

$$Cov(t, t+k) = \sigma^2(1-\theta)(1+(t-1)(1-\theta)).$$

▷ Utilizing this expression and that of the variance we obtain the autocorrelation function:

$$\rho(t, t+k) = \frac{(1-\theta)(1+(t-1)(1-\theta))}{\sqrt{(1+(t-1)(1-\theta)^2)\,(1+(t+k-1)(1-\theta)^2)}}.$$

# The simple exponential smoothing process

▷ Assuming that $t$ is large, such that $t-1 \approx t$, and that $\theta$ is not very close to one, such that $1+t(1-\theta)^2 \approx t(1-\theta)^2$, then:

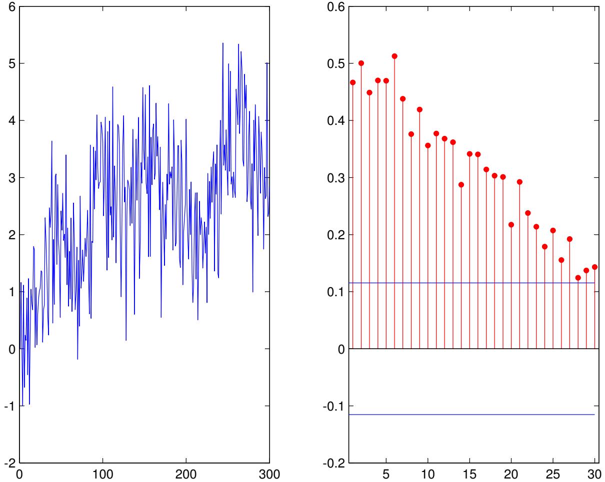
$$\rho(t, t+k) = \frac{t}{\sqrt{t(t+k)}}$$

▷ That is the same autocorrelation function (102) obtained for the random walk, i.e., the coefficients *show approximately linear decay* when the lag increases.

▷ Nevertheless, if $\theta$ is close to one, the coefficients do not necessarily have to be large, although we should always observe a linear decay.

## Example 52

*The figure shows a series generated using $\theta = .9$. Notice that although the autocorrelations are no longer close to one, they show the expected characteristic of linear decay.*



# Integrated process of order two

▷ Many real series with trend are represented with a model of two differences, or integrated of order two. A simple model that appears in many applications is:

$$\nabla^2 z_t = (1 - \theta B) a_t. \tag{106}$$

▷ To justify this model, we assume a random walk but where the drift changes with time, that is, the process:

$$\nabla z_t = c_t + u_t.$$

▷ This equation indicates that $\nabla z_t$ has a mean, which is the growth of $z_t$, that evolves over time. Successively substituting in the above equation and assuming that the process starts at $t = 0$ and that $z_0 = u_0 = 0$, we obtain:

$$z_t = (c_t + \ldots + c_1) + u_t + \ldots + u_1.$$

# Integrated process of order two

▷ Let us assume that the evolution of the growth coefficient at each time, $c_t$, is smooth, and such that

$$c_t = c_{t-1} + \epsilon_t,$$

where $\epsilon_t$ is a white noise process, independent of $u_t$.

▷ Thus:

$$\nabla^2 z_t = \nabla z_t - \nabla z_{t-1} = c_t + u_t - (c_{t-1} + u_{t-1}) = \\ = \epsilon_t + u_t - u_{t-1} = (1 - \theta B) a_t,$$

since the sum of white noise and a non-invertible MA(1) will be an invertible MA(1).

▷ We conclude, therefore, that process (106) is a generalization of the random walk that allows the drift to vary smoothly over time.

▷ In general, integrated processes of order two can be seen as a generalization of integrated processes of order one but where the slope of the growth line, instead of being fixed, varies over time.

# ARIMA processes

▷ The two processes discussed, RW and IMA, were obtained by accepting that the root of the AR part of the AR(1) and ARMA(1,1) processes is the unit. This idea can be generalized for any ARMA process, allowing one or various roots of the AR operator to be the unit:

$$(1-\phi_1 B - \ldots - \phi_p B^p)(1-B)^d z_t = c + (1-\theta_1 B - \ldots - \theta_q B^q) a_t$$

which we call **ARIMA** $(p, d, q)$ **processes**.

- $p$ is the order of the autoregressive stationary part.
- $d$ is the number of unit roots (order of integration of the process).
- $q$ is the order of the moving average part.

▷ Using the difference operator, $\nabla = 1 - B$, the above process is usually written as:

$$\phi_p(B)\nabla^d z_t = c + \theta_q(B) a_t. \tag{107}$$

# ARIMA processes

▷ ARIMA stands for **A**utoregressive **I**ntegrated **M**oving **A**verage, where "integrated" indicates that letting $\omega_t = \nabla^d z_t$ denote the stationary process, $z_t$ is obtained as a sum (integration) of $\omega_t$. Indeed, if

$$\omega_t = (1 - B)z_t$$

since, defining

$$(1 - B)^{-1} = 1 + B + B^2 + B^3 + \ldots$$

results in:

$$z_t = (1 - B)^{-1}\omega_t = \sum_{j=-\infty}^{t} \omega_t.$$

▷ We have seen two examples of ARIMA processes in two earlier examples: the random walk is an ARIMA(0,1,0) model and the simple exponential smoothing is the ARIMA(0, 1, 1) or IMA(1, l).

▷ Both are characterized by the fact that the autocorrelation function has slowly decaying coefficients.

▷ All non-stationary ARIMA processes have this general property.

▷ To prove it, recall that the correlogram of an ARMA(p,q) satisfies the equation for $k > q$:

$$\phi_p(B)\rho_k = 0 \qquad k > q$$

whose solution is of type:

$$\rho_k = \sum_{i=1}^{p} A_i G_i^k$$

where $G_i^{-1}$ are the roots of $\phi_p(B) = 0$ and $|G_i| < 1$.

▷ If one of these roots $G_i$ is very close to the unit, writing $G_i = 1 - \varepsilon$, with very small $\varepsilon$, for large $k$ the terms $A_j G_j^k$ will be zero due to the other roots (since $G_j^k \longrightarrow 0$) and we have that, approximately:

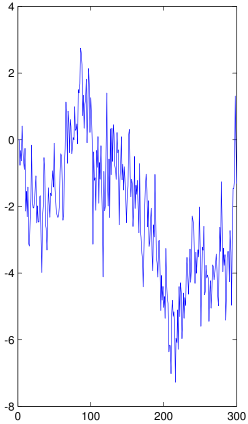
$$\rho_k = A_i(1-\varepsilon)^k \simeq A_i(1-k\varepsilon) \quad \text{for large } k$$

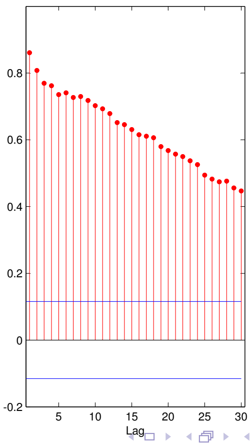
▷ As a result, the *ACF* will have positive coefficients that will fade out approximately linearly and may be different from zero for high values of $k$.

▷ This property of persistence of positive values in the correlogram (even though they are small) and linear decay characterizes the non-stationary processes.

## Example 53

*The figure shows a realization of the series* $(1 - .4B)\nabla z_t = (1 - .8B)a_t$ *and its estimated autocorrelation function. The linear decay of the autocorrelation coefficients is again observed.*



# Integrated processes and trends

▷ Not all non-stationary processes are integrated, but integrated processes cover many interesting cases that we find in practice.

▷ It is easy to prove that any process which is the sum of a polynomial trend and a stationary process will be integrated. For example, let us take the case of a deterministic linear growth process:

$$z_t = b + ct + u_t \tag{108}$$

where $u_t$ is white noise.

▷ This process is integrated of order one or $I(1)$, since taking the first difference of the series $z_t$:

$$\begin{aligned} \omega_t &= z_t - z_{t-1} = b + ct + u_t - (b + c(t-1) + u_{t-1}) \\ &= c + u_t - u_{t-1} \end{aligned}$$

since $c$ is constant and $\nabla u_t$ is a stationary process (being the difference of two stationary processes, see section 3.4), process $\omega_t$ is stationary as well.

# Integrated processes and trends

▷ Therefore, $z_t$ defined by (108), is integrated of order one, and its first difference follows model (104) with $a_t = u_t$ and parameter $\theta$ equal to the unit.

▷ The process is not invertible, and this property tells us that we have a deterministic linear component in the model.

▷ We observe that by differentiating the process the remaining constant is the slope of the deterministic linear growth.

▷ This result is valid for processes generated as a sum of a polynomial trend of order $h$ and any stationary process, $u_t$:

$$z_t = \mu_t + u_t \tag{109}$$

where

$$\mu_t = a + bt + ct^2 + ... + dt^h.$$

▷ If we differentiate this process $h$ times a stationary process is obtained with a non-invertible moving average part.

# Integrated processes and trends

▷ For example, let us assume a quadratic trend. Then, taking two differences in the process, we have:

$$\nabla^2 z_t = c + \nabla^2 u_t$$

and the process $\eta_t = \nabla^2 u_t$ has a non-invertible moving average.

▷ For example, if $u_t$ is an AR(1), $(1 - \phi B) u_t = a_t$, we have

$$\eta_t = (1 - \phi B)^{-1} \nabla^2 a_t$$

and $\eta_t$ is a non-invertible ARMA(1,2) process.

▷ This example shows us that if a series has a deterministic polynomial trend and it is modelled by an ARIMA process we have: (1) non-invertible moving average components; (2) a stationary series with a mean different from zero and a constant that is the highest order coefficient in the deterministic trend.

▷ Again, we see the importance of constants in integrated processes, since they represent deterministic effects.

# Integrated processes and trends

▷ Processes with a deterministic linear trend are infrequent in practice, but we can generalize the above model somewhat allowing the slope at each time to vary a little with respect to the previous value. We can write the process with a linear trend but with variable components, as:

$$z_t = \mu_t + v_t. \tag{110}$$

▷ With this model the value of the series at each point in time is the sum of its level, $\mu_t$, and a white noise process, $v_t$.

▷ We assume that the level evolves as:

$$\mu_t = \mu_{t-1} + c + \varepsilon_t. \tag{111}$$

▷ The process has a trend $c$, because the level at time t, $\mu_t$ is obtained from the previous time adding a constant $c$, which is the slope, plus a small random component, $\varepsilon_t$.

# Integrated processes and trends

▷ In this process the first difference is:

$$\nabla z_t = c + \varepsilon_t + \nabla v_t$$

and since the sum of an MA(1) process and white noise is an MA(1), we again obtain process (104), but now with a $\theta$ parameter smaller than the unit, as in (104).

▷ The above model can be further generalized by making the deterministic slope $c$, change with time, but with certain inertia. To do that, we replace expression (111) with:

$$\mu_t = \mu_{t-1} + \beta_t + \varepsilon_t$$

and

$$\beta_t = \beta_{t-1} + \upsilon_t$$

where the processes $\varepsilon_t$ and $\upsilon_t$ are independent white noises. Now the level grows linearly, but with a slope, $\beta_t$, that changes over time.

# Integrated processes and trends

▷ In this representation if the variance of $\upsilon_t$ is very small the series has an almost constant trend, and we return to the above model. If we take the difference in the original series, we have

$$\nabla z_t = \beta_t + \varepsilon_t + \nabla u_t$$

and taking a second difference

$$\nabla^2 z_t = \upsilon_t + \nabla \varepsilon_t + \nabla^2 u_t = (1 - \theta_1 B - \theta_2 B^2) a_t \tag{112}$$

since if we sum MA processes a new MA process is obtained.

▷ Depending on the variances of the three white noise processes we can have different situations: (i) if $V(u_t)$ is much smaller than that of the other noise processes the term $\nabla^2 u_t$ may not be taken into account in the right hand side of (112) and the sum will be an MA(1); (ii) if the term $\upsilon_t$ is the dominant one and we do not take the other two into account, the series is approximately the model $\nabla^2 z_t = \upsilon_t$, and (iii) if $V(\upsilon_t)$ is zero, then $\beta_t = \beta_{t-1} = \beta$ and we return to the model which has a deterministic trend.

# Integrated processes and trends

▷ It is important to understand the difference between the two types of models presented here:

- Model (108) is pure deterministic, since at each time, $t$, the expected value of the series is determined, and is $b + ct$. Moreover, the expected value of the series at $z_t$ knowing the value at $z_{t-1}$ is still the same, $b + ct$. In this model, knowing the previous values does not modify the predictions.

- Model (110) is more general, because although at each point in time it also has linear growth equal to $c$, the level of the series at each point in time is not determined from the beginning and can vary randomly. This means that the expected value of the series at $z_t$ knowing the value at $z_{t-1}$ is $z_{t-1} + c$, and depends on the observed value.

▷ To summarize, the ARIMA models, by including differences, incorporate stochastic trends and, in extreme cases, deterministic trends. In this case, the differentiated series will have a mean different from zero and non-invertible moving average terms may appear.

# Long-memory processes

▷ There is a class of stationary processes that are easily confused with non-stationary integrated processes. These are **long-memory processes**, which are characterized by having many autocorrelation coefficients with small coefficients and which decay very slowly.

▷ This property has been observed in some stationary meteorological or environmental series (but also in financial series). The slow decay of the autocorrelations could lead to modelling these series as if they were integrated of order one, by taking a difference.

▷ Nevertheless the decay is different from that of integrated processes: it is much faster for the first lags than in an integrated process, but slower for high lags.

▷ This structure means that in a long-memory process we observe many small autocorrelation coefficients and they decay very slowly for high lags, as opposed to what usually happens with integrated processes, where the decay is linear.

# Long-memory processes

▷ The simplest long-memory process is obtained by generalizing the random walk using:

$$(1-B)^d z_t = a_t \tag{113}$$

where now the parameter $d$ instead of being an integer is a real number.

▷ To define the operator $(1-B)^d$ when $d$ is not an integer, we start from the expression of Newton's binomial theorem. When $d$ is a positive integer, it is verified that:

$$(1-B)^d = \sum_{i=0}^{d} \left( \begin{array}{c} d \\ i \end{array} \right) (1)^{d-i}(-B)^i = \sum_{i=0}^{d} \alpha_i B^i, \tag{114}$$

where the coefficients of this sum are defined by

$$\alpha_i = \left( \begin{array}{c} d \\ i \end{array} \right) (-1)^i = \frac{d!}{(d-i)!i!}(-1)^i = \frac{\Gamma(d+1)}{\Gamma(i+1)\Gamma(d-i+1)}(-1)^i,$$

where $\Gamma(a+1) = a\Gamma(a)$ is the Gamma function which for integer values coincides with the factorial function.

▷ To generalize this definition when $d$ is not an integer, we can take advantage of the fact that the Gamma function is defined for any real number, and for real $d$ write:

$$(1-B)^d = \sum_{i=0}^{\infty} \alpha_i B^i, \tag{115}$$

where the coefficients of this infinite expansion are given by

$$\alpha_i = \frac{\Gamma(d+1)}{\Gamma(i+1)\Gamma(d-i+1)}(-1)^i.$$

▷ Moreover, it can be proved that when $d$ is an integer this definition is consistent with (114), since all terms greater than $d$ in this infinite summation are zero, such that for an integer $d$ (115) is reduced to (114).

▷ We observe that for $i = 0$ the result is $\alpha_0 = 1$, and using the properties of the Gamma function it is shown that:

$$\alpha_i = \prod\nolimits_{0<j\leq i} \frac{j-1-d}{j}, \qquad \text{for } i = 1, 2, ...$$

▷ Using this expression we can write the fractional process (113) by means of the AR($\infty$) representation developing the operator $(1-B)^d$, to obtain:

$$z_t = -\sum\nolimits_{i=0}^{\infty} \alpha_i z_{t-i} + a_t. \tag{116}$$

It can be proved that this process also admits an MA($\infty$) representation of the form

$$z_t = \sum\nolimits_{i=0}^{\infty} \psi_i a_{t-i},$$

where the coefficients verify Wold's representation of a stationary MA($\infty$) series with finite variance, that is, $\sum \psi_i^2 < \infty$.

## Long-memory processes

▷ These coefficients are obtained using

$$\psi_i = \frac{\Gamma(d+i)}{\Gamma(i+1)\Gamma(d)}.$$

It is proved that the autocorrelation function of the long-memory fractional process (113) is

$$\rho(k) = \frac{\Gamma(d+k)\Gamma(1-d)}{\Gamma(k+1-d)\Gamma(d)}.$$

▷ Particularly for $k = 1$, we obtain

$$\rho(1) = \frac{d\Gamma(d)\Gamma(1-d)}{(1-d)\Gamma(1-d)\Gamma(d)} = \frac{d}{1-d},$$

for $k = 2$

$$\rho(2) = \frac{d(d+1)}{(2-d)(1-d)} = \rho(1)\frac{(d+1)}{(2-d)}$$

and, in general

$$\rho(k) = \rho(k-1)\frac{(d+k+1)}{(k-d)}.$$

# Long-memory processes

▷ For large $k$, we can approximate the correlations by:

$$\rho(k) \approx \frac{\Gamma(1-d)}{\Gamma(d)} |k|^{2d-1}$$

and taking logarithms

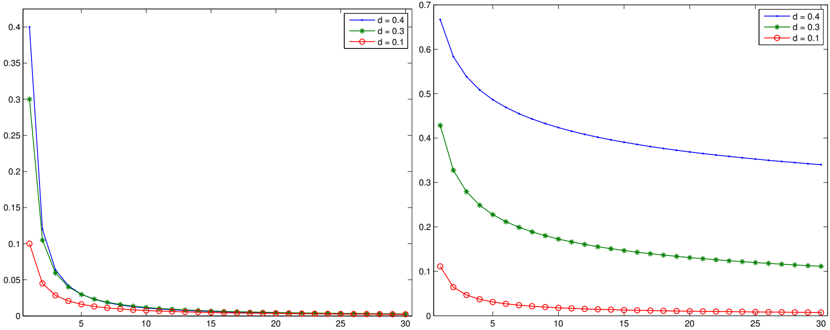
$$\log \rho(k) \approx a + (2d-1) \log |k| .$$

▷ Therefore, if we represent $\log \rho(k)$ with respect to $\log|k|$ for high values of $k$ we obtain a line with slope $2d - 1$.

▷ This property differentiates a long-memory process from a non-stationary integrated process, where we have seen that $\log \rho(k)$ decays linearly with $k$ and not with $\log k$.

# Example 54

*The left figure shows the coefficients -$\alpha_i$ of the AR representation given by (116) for several values of $d$ and right figure the autocorrelation function for those values of $d$.*




▷ Notice that the AR coefficients quickly become very small, and that the AC function has many non-null coefficients that decay slowly.

▷ These properties make it easy to confuse this process with a high order AR($p$), and if $d$ is close to .5 the high values and slow decay of the AC function can suggest an I(1) process.

# Long-memory processes

▷ Generalizing on the above, we can consider stationary ARMA processes, but ones that include long-memory as well, defining

$$(1-\phi_1 B - \dots - \phi_p B^p)(1-B)^d z_t = c + (1-\theta_1 B - \dots - \theta_q B^q) a_t$$

where $d$ is not necessarily an integer.

▷ If $d < 0.5$ the process is stationary, whereas if $d \geq 0.5$ it is non-stationary.

▷ These models are called ARFIMA$(p,d,q)$, **autoregressive fractionally integrated moving average**.

▷ If the process is stationary it has the property whereby for high lags, when the ARMA structure disappears, the long-memory characteristic will appear. If it is not stationary, taking a difference will convert it to stationary, but with long-memory.

# Long-memory processes

▷ Long-memory processes may appear when short-memory series are added under certain circumstances. Granger (1980) proved that if we add $N$ AR(1) independent processes as a limit we obtain a long-memory process.

▷ We known that the aggregation of independent AR processes leads to ARMA processes, and this result is expected with few summands.

▷ Granger's results indicate that in the limit we will have a long-memory process and suggest that these processes can be approximated by an ARMA with orders $p$ and $q$ that are high and similar.

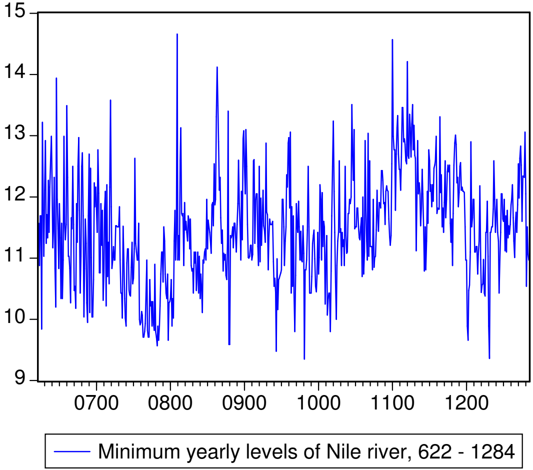
▷ With series that are not very long it is difficult to differentiate a long-memory process from an ARMA. Nevertheless, if we have a series with a large number of observations, as is typical in meteorological or financial data, in some cases LM processe can provide better fit than the short-memory ARMA processesn.

▷ However, as we will see in the next chapter, the differences in the prediction are usually small.

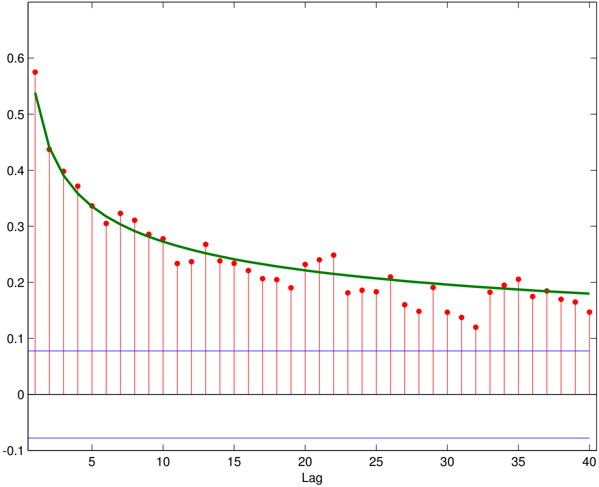
# Long-memory processes - Nile river example

## Example 55

*The figure shows a series of yearly data for minimum levels of the Nile River between the years 622 and 1284, which can be found in the file nilomin.dat.*



▷ The figure shows its autocorrelation function, and it is observed that the correlations decay very slowly.




▷ As a reference its also included the AC function of a fractional process with $d = 0.35$.

▷ The figure shows the logarithm of the AC compared to the logarithm of the lag, and we see that the relationship is approximately linear.

▷ As a comparison we have also included in this figure the Madrid Stock Exchange index data.

▷ In this integrated process for small values of $k$ the AC decay more slowly in the integrated series than in the long-memory one, but for larger $k$ the coefficients for this series, which follow an ARIMA model, decay faster than that which would correspond to a long-memory process.