Applied Time Series Analysis
FS 2012 – Week 04

Marcel Dettling

Institute for Data Analysis and Process Design

Zurich University of Applied Sciences

marcel.dettling@zhaw.ch

http://stat.ethz.ch/~dettling

ETH Zürich, March 12, 2012

Applied Time Series Analysis
FS 2012 – Week 04

Where are we?

For much of the rest of this course, we will deal with (weakly) stationary time series. They have the following properties:

- $E[X_t] = \mu$
- $Var(X_t) = \sigma^2$
- $Cov(X_t, X_{t+h}) = \gamma_h$

If a time series is non-stationary, we know how to decompose into deterministic and stationary, random part.

Our forthcoming goals are:

- understanding the dependency in a stationary series
- modeling this dependency and generate forecasts

Autocorrelation

The aim of this section is to explore the dependency structure within a time series.

Def: **Autocorrelation**

$$\rho(k) = Cor(X_{t+k}, X_t) = \frac{Cov(X_{t+k}, X_t)}{\sqrt{Var(X_{t+k}) \cdot Var(X_t)}}$$

The autocorrelation is a dimensionless measure for the amount of linear association between the random variables collinearity between the random variables X_{t+k} and X_t.

Applied Time Series Analysis
FS 2012 – Week 04

Interpretation of Autocorrelations

How to interpret autocorrelation from a practical viewpoint?

→ We e.g. assume that $\rho(k) = 0.7$.

→ Then, the square of the correlation coefficient, i.e. $\rho(k)^2 = 0.49$, , is the percentage of variability explained by the linear association between X_t and its respective predecessor X_{t-1}.

→ Here in our example, X_{t-1} accounts for roughly 49% of the variability observed in random variable X_t.

→ From this we can also conclude that any $\rho(k) < 0.4$ is not a very strong association, i.e. has small effect.

Autocorrelation Estimation: lag k

How does it work?

→ **Plug-in estimate with sample covariance**

$$\hat{\rho}(k) = \frac{\hat{\gamma}(k)}{\hat{\gamma}(0)} = \frac{Cov(X_t, X_{t+k})}{Var(X_t)}$$

where $$\hat{\gamma}(k) = \frac{1}{n}\sum_{s=1}^{n-k}(x_{s+k} - \overline{x})(x_s - \overline{x})$$

and $$\overline{x} = \frac{1}{n}\sum_{t=1}^{n} x_t$$

Application: Variance of the Arithmetic Mean

Practical problem: we need to estimate the mean of a realized/ observed time series. We would like to attach a standard error.

- If we estimate the mean of a time series without taking into account the dependency, the standard error will be flawed.
- This leads to misinterpretation of tests and confidence intervals and therefore needs to be corrected.
- The standard error of the mean can both be over-, but also underestimated. This depends on the ACF of the series.

→ **For the derivation, see the blackboard…**

Outlook to AR(p)-Models

Suppose that E_t is an i.i.d random process with zero mean and variance σ_E^2. Then a random process X_t is said to be an auto-regressive process of order p if

$$X_t = \alpha_1 X_{t-1} + ... + \alpha_p X_{t-p} + E_t$$

This is similar to a multiple regression model, but X_t is regressed not on independent variables, but on past values of itself. Hence the term auto-regressive.

We use the abbreviation **AR(p)**.

Partial Autocorrelation Function (PACF)

The k^{th} partial autocorrelation π_k is defined as the correlation between X_{t+k} and X_t, given all the values in between.

$$\pi_k = Cor(X_{t+k}, X_t \mid X_{t+1} = x_{t+1}, \ldots, X_{t+k-1} = x_{t+k-1})$$

Interpretation:

- Given a time series X_t, the partial autocorrelation of lag *k*, is the autocorrelation between X_t and X_{t+k} with the linear dependence of X_{t+1} through to X_{t+k-1} removed.
- One can draw an analogy to regression. The ACF measures the „simple“ dependence between X_t and X_{t+k}, whereas the PACF measures that dependence in a „multiple“ fashion.

Facts About the PACF and Estimation

We have:

- $\pi_1 = \rho_1$
- $\pi_2 = \dfrac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$ for AR(1) models, we have $\pi_2 = 0$, because $\rho_2 = \rho_1^2$
- For estimating the PACF, we utilize the fact that for any AR(p) model, we have: $\pi_p = \alpha_p$ and $\pi_k = 0$ for all $k > p$.

  Thus, for finding $\hat{\pi}_p$, we fit an AR(p) model to the series for various orders p and set $\hat{\pi}_p = \hat{\alpha}_p$

Facts about the PACF

- Estimation of the PACF is implemented in R.
- The first PACF coefficient is equal to the first ACF coefficient. Subsequent coefficients are not equal, but can be derived from each other.
- For a time series generated by an AR(p)-process, the p^{th} PACF coefficient is equal to the p^{th} AR-coefficient. All PACF coefficients for lags $k > p$ are equal to 0.
- Confidence bounds also exist for the PACF.

Basics of Modeling

Simulation

(Time Series) Model → Data

Estimation

Inference

Residual Analysis

Data → (Time Series) Model

A Simple Model: White Noise

A time series $(W_1, W_2, ..., W_n)$ is a White Noise series if the random variables $W_1, W_2, ...$ are independent and identically distributed with mean zero.

This imples that all variables W_t have the same variance σ_W^2, and

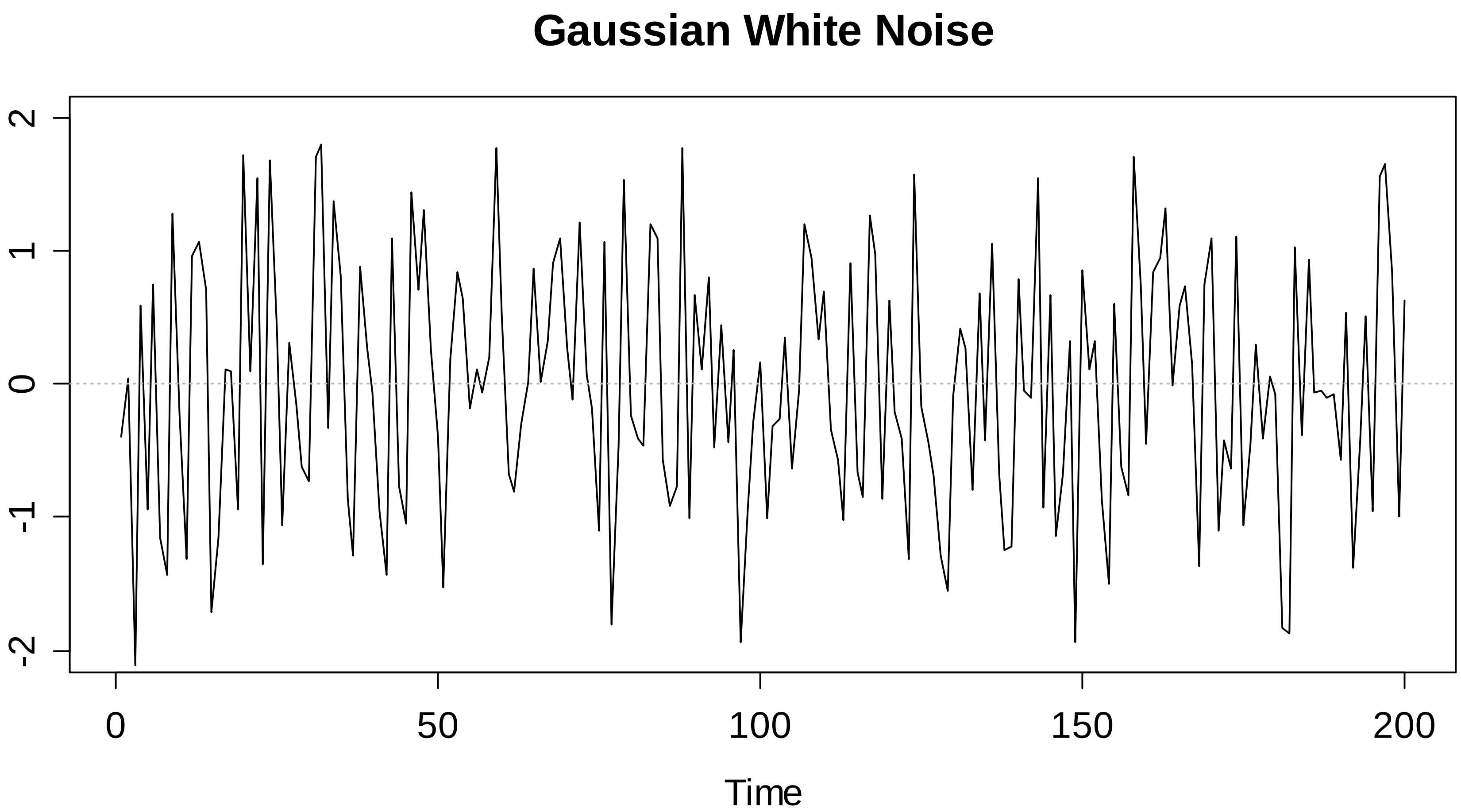
$$Cov(W_i, W_j) = 0 \text{ for all } i \neq j.$$

Thus, there are no autocorrelations either: $\rho_k = 0$ for all $k \neq 0$.

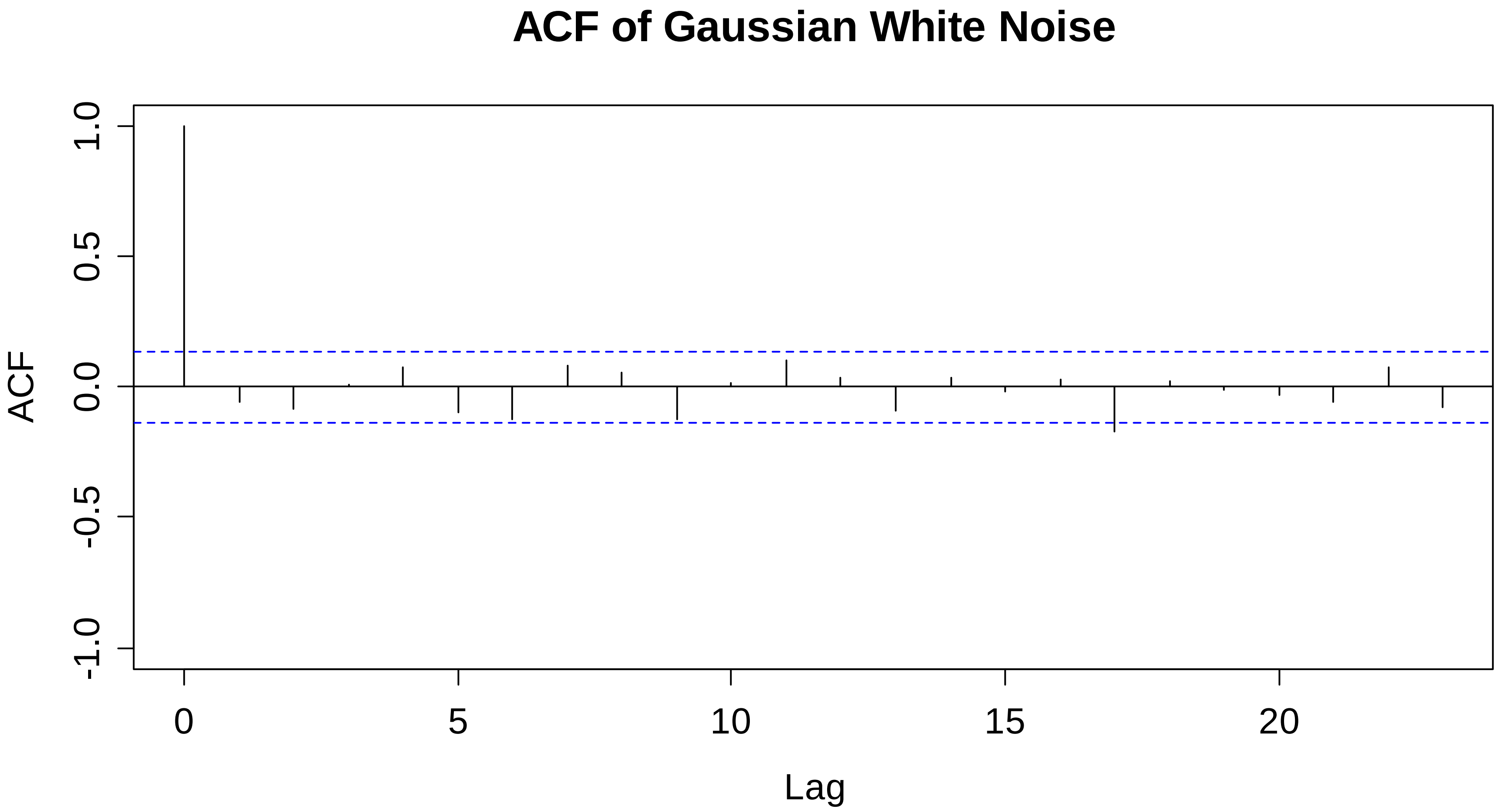
If in addition, the variables also follow a Gaussian distribution, i.e. $W_t \sim N(0, \sigma_W^2)$, the series is called Gaussian White Noise.

Applied Time Series Analysis
FS 2012 – Week 04

Example: Gaussian White Noise



Example: Gaussian White Noise



Time Series Modeling

There is a wealth of time series models

- AR	autoregressive model
- MA	moving average model
- ARMA	combination of AR & MA
- ARIMA	non-stationary ARMAs
- SARIMA	seasonal ARIMAs
- …	

Autoregressive models are among the simplest and most intuitive time series models that exist.

Basic Idea for AR-Models

We have a time series where, resp. we model a time series such that the random variable X_t depends on a linear combination of the preceding ones $X_{t-1},...,X_{t-p}$, plus a „completely independent“ term called innovation E_t.

$$X_t = \alpha_1 X_{t-1} + ... + \alpha_p X_{t-p} + E_t$$

p is called the order of the AR-model. We write AR(p). Note that there are some restrictions to E_t.

AR(1)-Model

The simplest model is the AR(1)-model

$$X_t = \alpha_1 X_{t-1} + E_t$$

where

E_t is i.i.d with $E[E_t] = 0$ and $Var(E_t) = \sigma_E^2$

Under these conditions, E_t is a white noise process, and we additionally require **causality**, i.e. E_t being an **innovation**:

E_t is independent of $X_s, s < t$

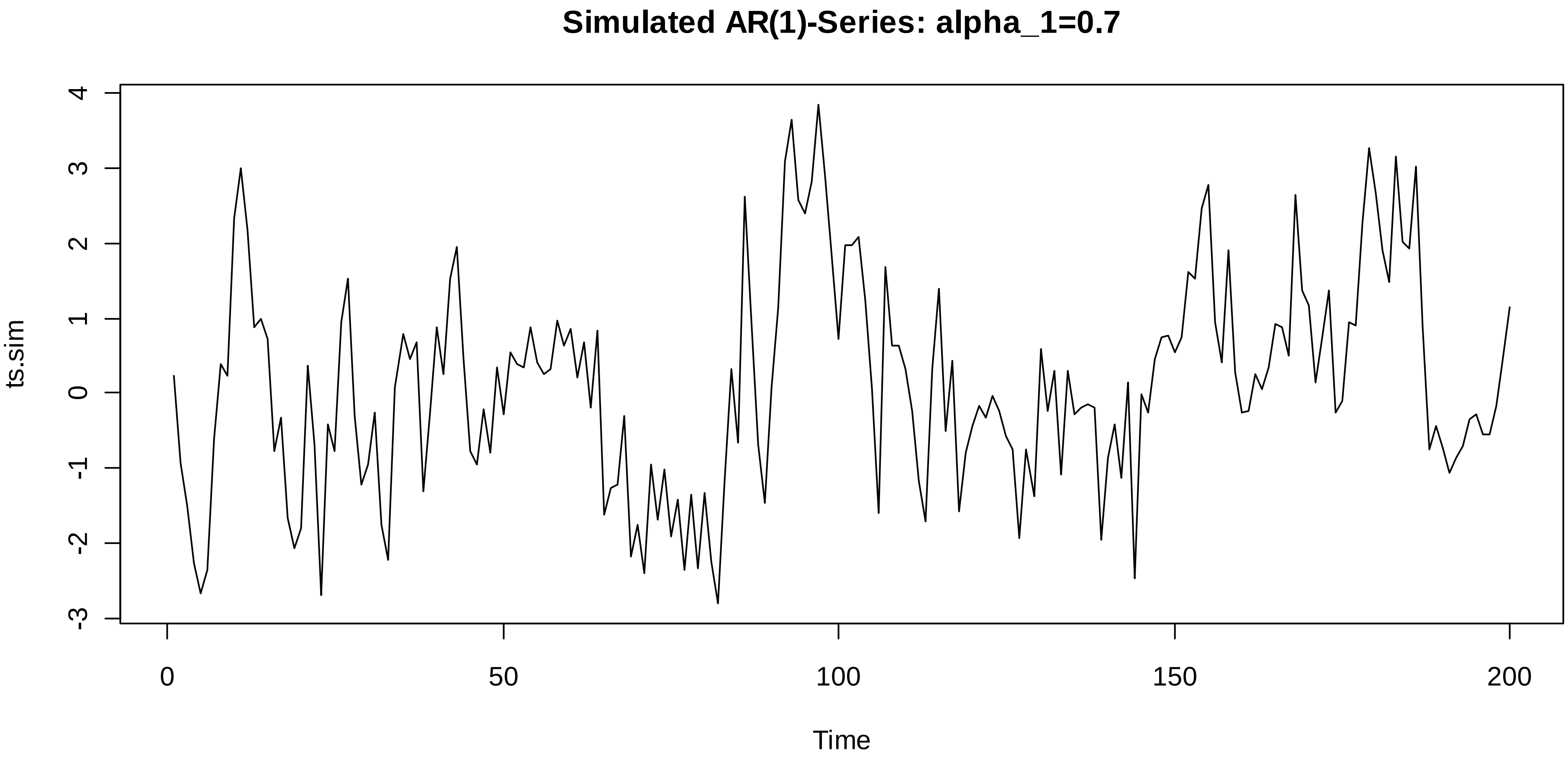
Causality

Note that causality is an important property that, despite the fact that it‘s missing in much of the literature, is necessary in the context of AR-modeling:

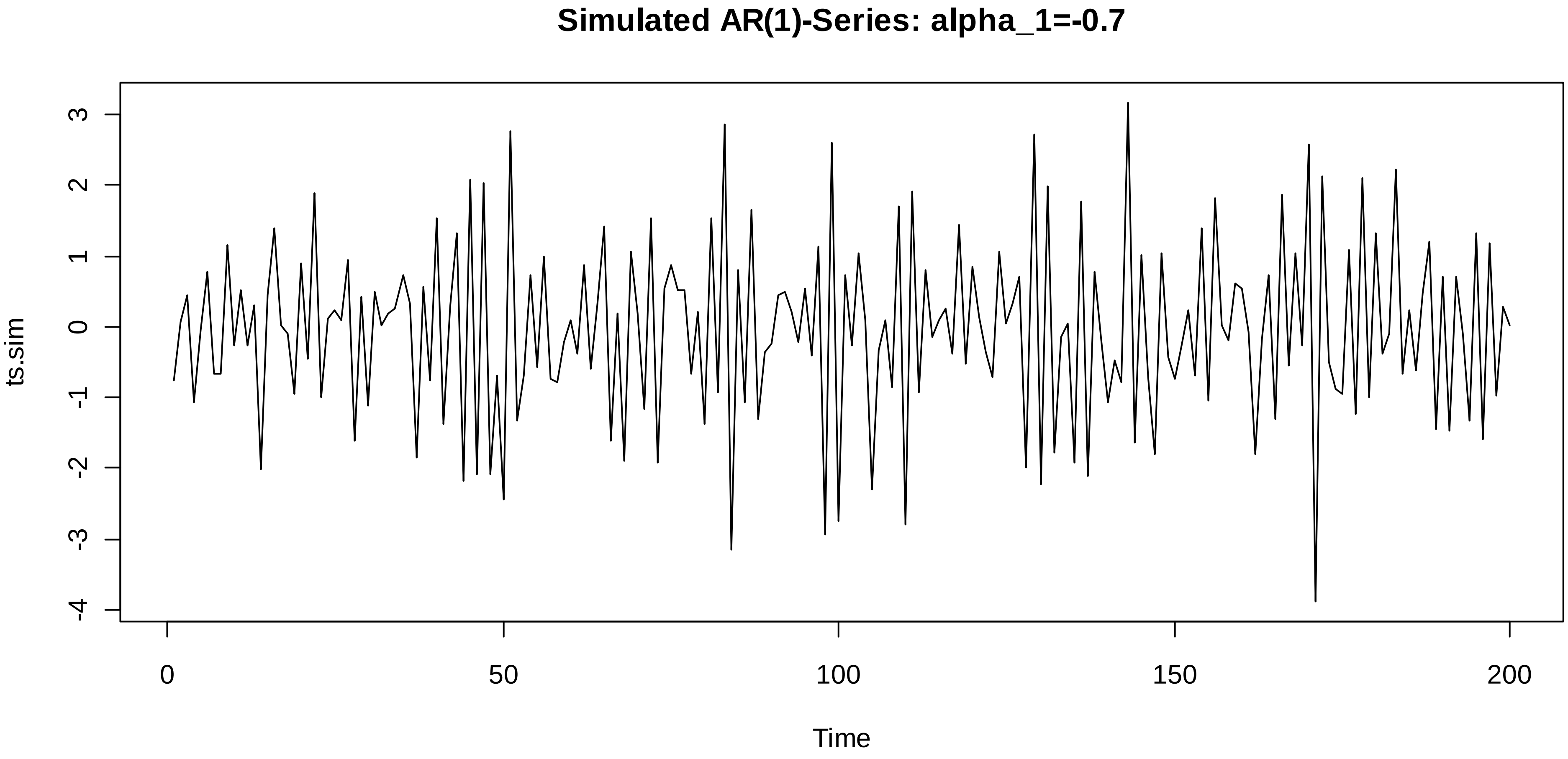
E_t is an innovation process $\rightarrow$ E_t all are independent

All E_t are independent $\nrightarrow$ E_t is an innovation

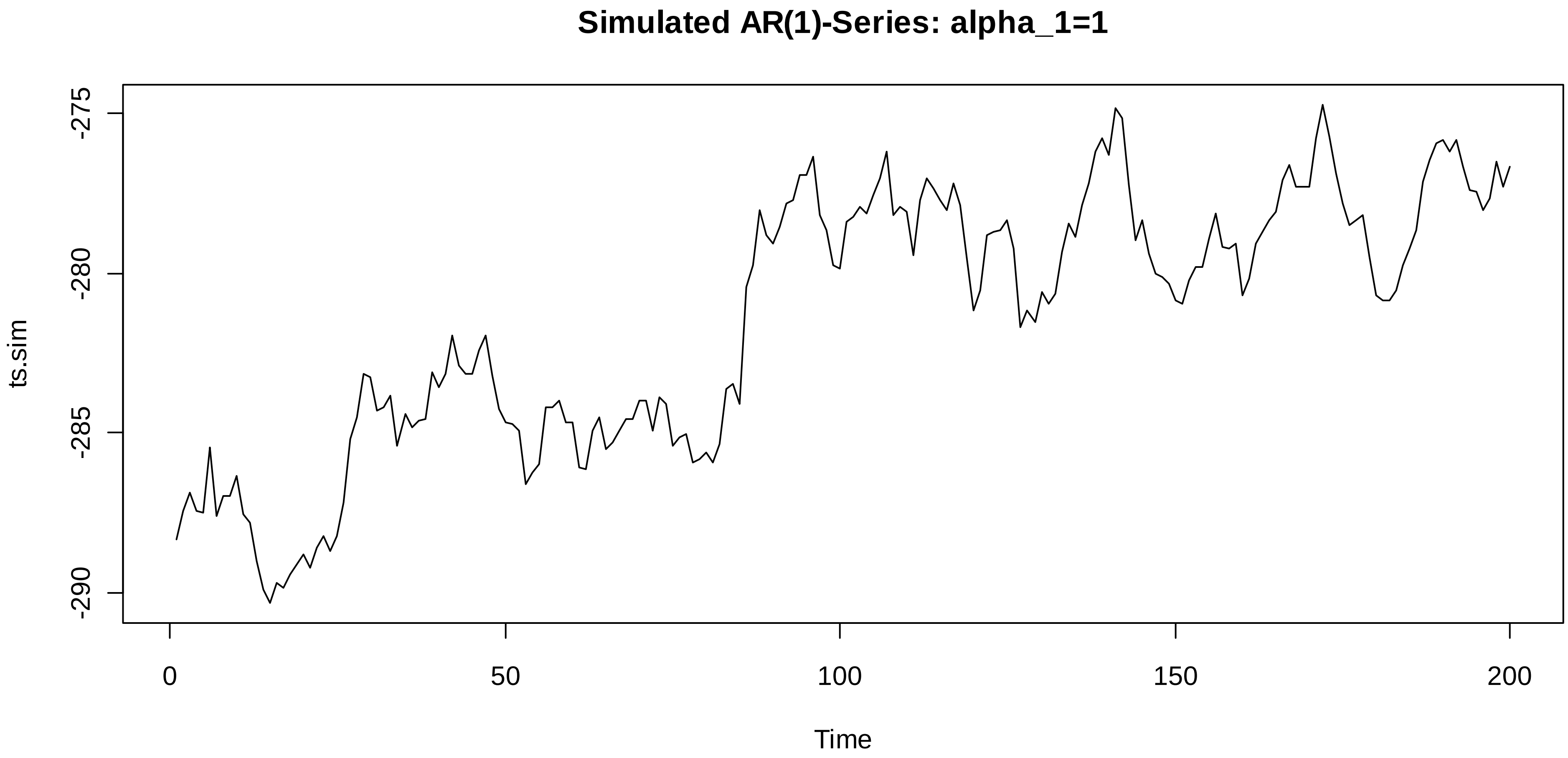
Simulated AR(1)-Series



Simulated AR(1)-Series



Simulated AR(1)-Series

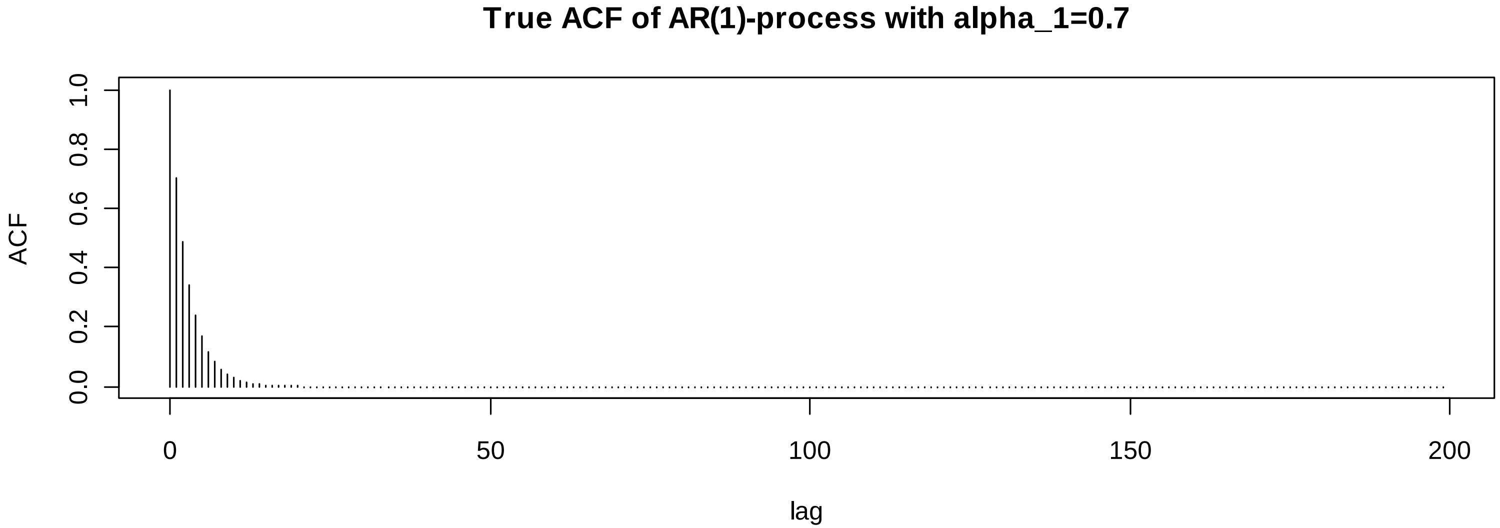


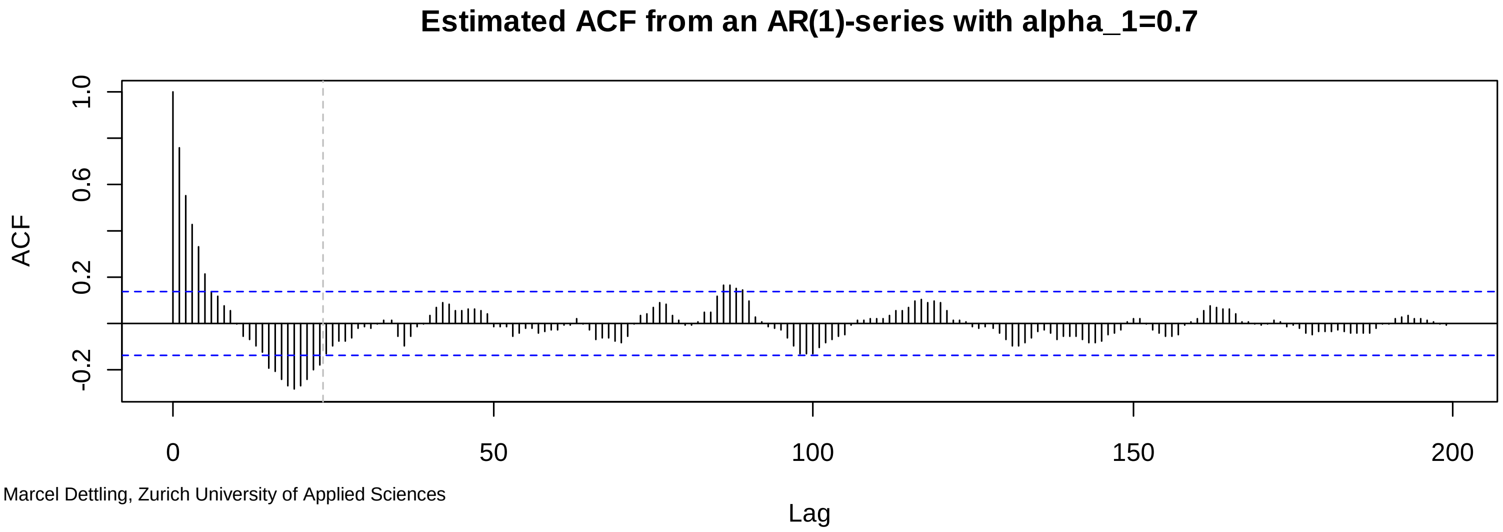
Moments of the AR(1)-Process

Some calculations with the moments of the AR(1)-process give insight into stationarity and causality

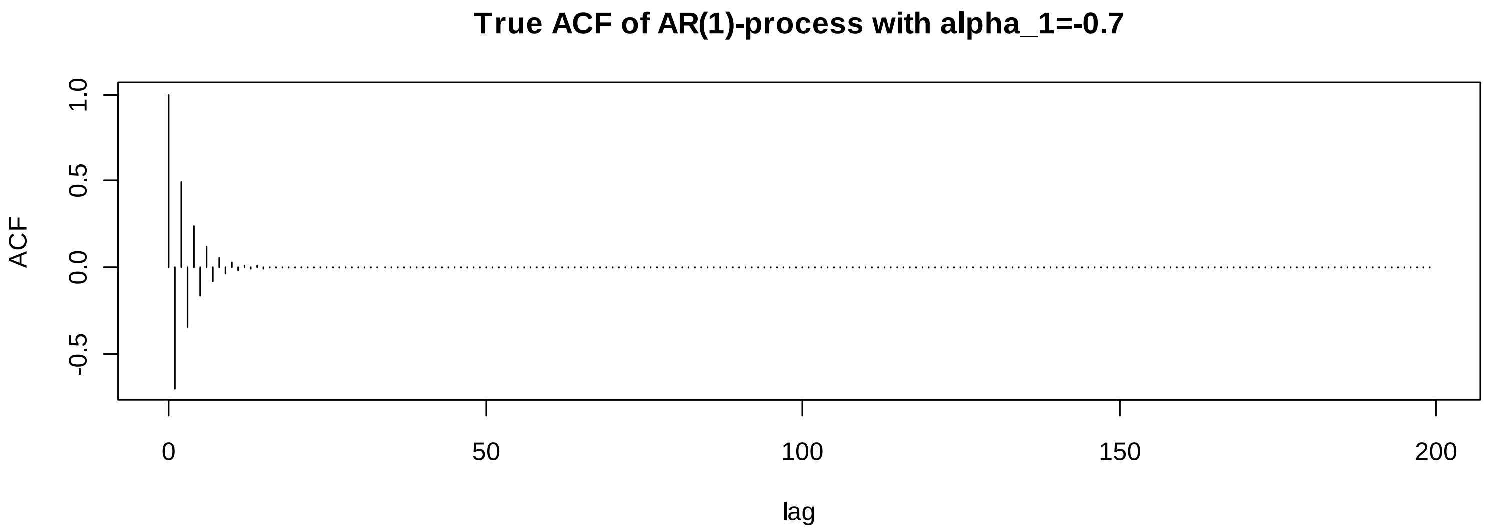
Proof: **See blackboard…**

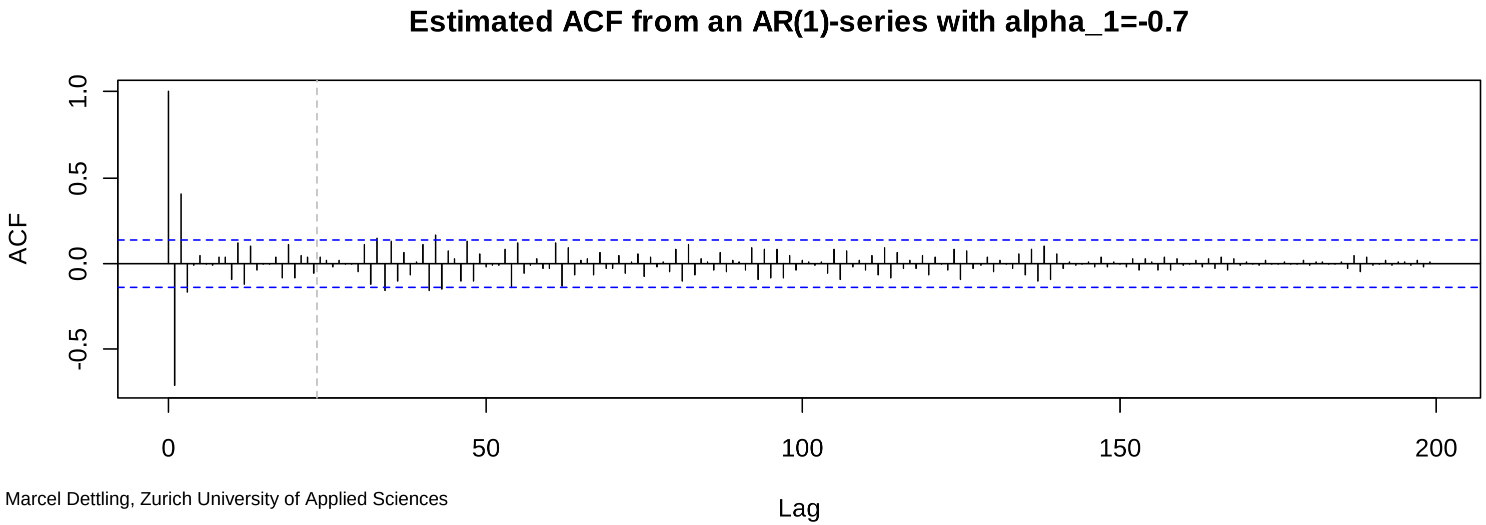
Theoretical vs. Estimated ACF





Theoretical vs. Estimated ACF





AR(p)-Model

We here introduce the AR(p)-model

$$X_t = \alpha_1 X_{t-1} + \ldots + \alpha_p X_{t-p} + E_t$$

where again

E_t is i.i.d with $E[E_t] = 0$ and $Var(E_t) = \sigma_E^2$

Under these conditions, E_t is a white noise process, and we additionally require **causality**, i.e. E_t being an **innovation**:

E_t is independent of $X_s, s < t$

Mean of AR(p)-Processes

As for AR(1)-processes, we also have that:

$(X_t)_{t \in T}$ is from a stationary AR(p) => $E[X_t] = 0$

Thus: If we observe a time series with $E[X_t] = \mu \neq 0$, it cannot be, due to the above property, generated by an AR(p)-process

But: In practice, we can always de-“mean“ (i.e. center) a stationary series and fit an AR(p) model to it.

Yule-Walker-Equations

On the blackboard…

We observe that there exists a linear equation system built up from the AR(p)-coefficients and the ACF-coefficients of up to lag p. These are called Yule-Walker-Equations.

We can use these equations for fitting an AR(p)-model:

1) Estimate the ACF from a time series
2) Plug-in the estimates into the Yule-Walker-Equations
3) The solution are the AR(p)-coefficients

Stationarity of AR(p)-Processes

We require:

1) $E[X_t] = \mu = 0$

2) Conditions on $(\alpha_1, ..., \alpha_p)$

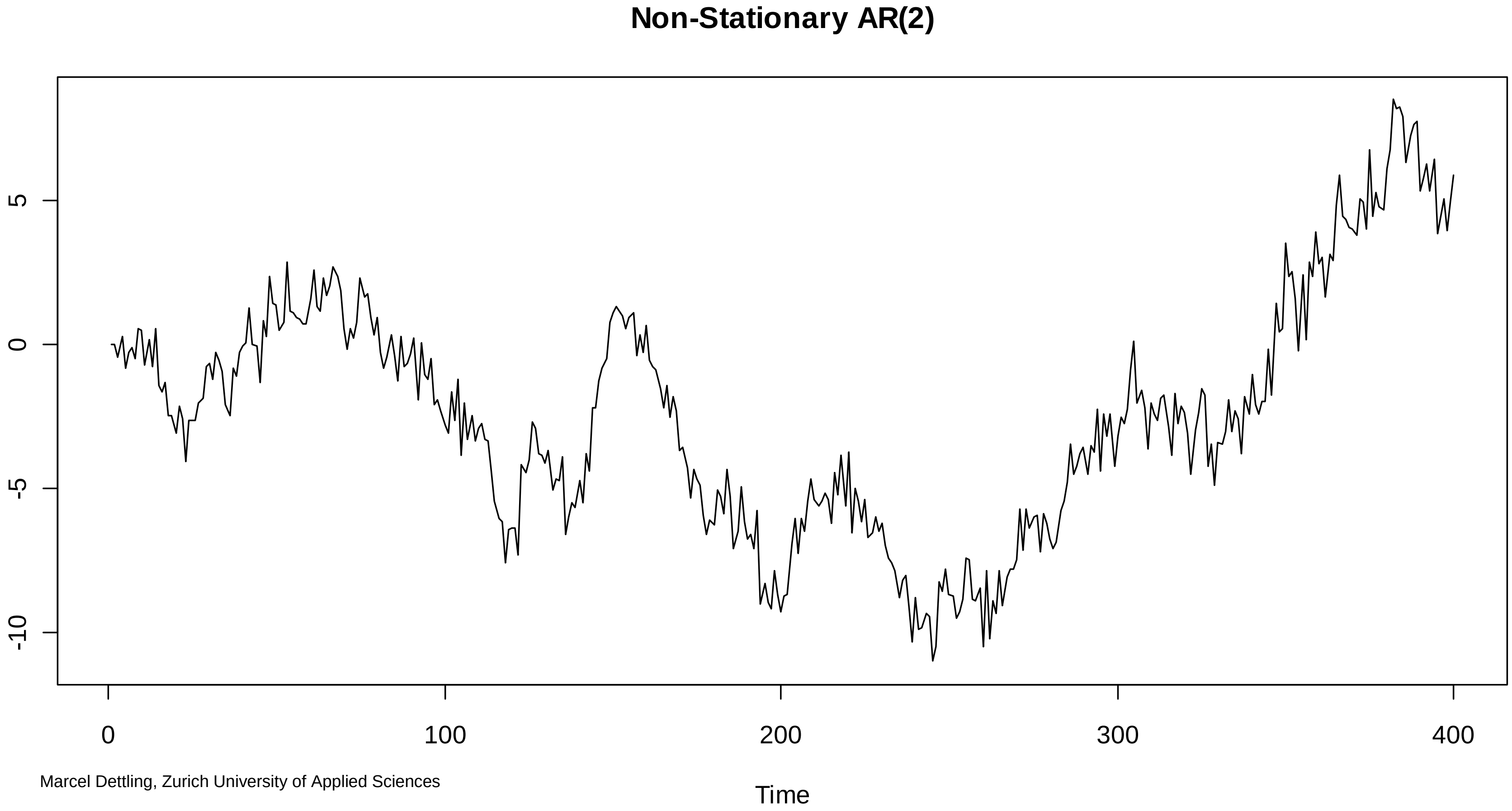
All (complex) roots of the characteristic polynom

$$1 - \alpha_1 z - \alpha_2 z^2 - \alpha_p z^p = 0$$

need to lie outside of the unit circle. This can be checked with R-function `polyroot()`

A Non-Stationary AR(2)-Process

$$X_t = \frac{1}{2}X_{t-1} + \frac{1}{2}X_{t-2} + E_t$$ is not stationary…



Fitting AR(p)-Models

This involves 3 crucial steps:

1) **Is an AR(p) suitable, and what is p?**
   - will be based on ACF/PACF-Analysis

2) **Estimation of the AR(p)-coefficients**
   - Regression approach
   - Yule-Walker-Equations
   - and more (MLE, Burg-Algorithm)

3) **Residual Analysis**
   - to be discussed

Applied Time Series Analysis
FS 2012 – Week 04

AR-Modelling

1

Identification of the Order p

- ACF/PACF
- AIC/BIC

2

Parameter Estimation

- Regression
- Yule-Walker
- MLE
- Burg

3

Model Diagnostics

- Residual Analysis
- Simulation

Is an AR(p) suitable, and what is p?

- For all AR(p)-models, the **ACF** decays exponentially quickly, or is an exponentially damped sinusoid.

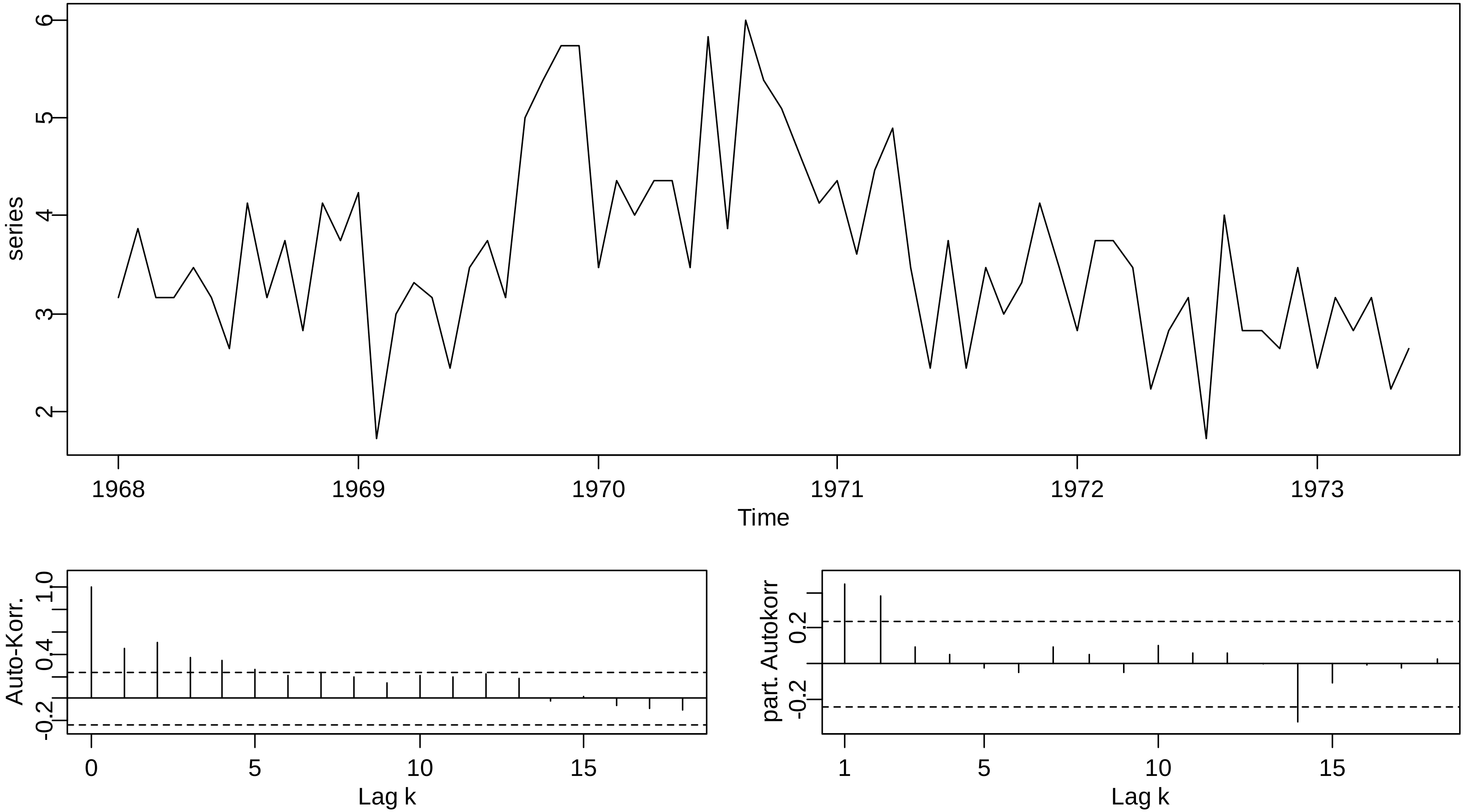
- For all AR(p)-models, the **PACF** is equal to zero for all lags k>p.

If what we observe is fundamentally different from the above, it is unlikely that the series was generated from an AR(p)-process. We thus need other models, maybe more sophisticated ones.

Remember that the sample ACF has a few peculiarities and is tricky to interpret!!!

Applied Time Series Analysis
FS 2012 – Week 04

Model Order for sqrt(purses)



Applied Time Series Analysis
FS 2012 – Week 04

Model Order for log(lynx)