Applied Time Series Analysis
SS 2013 – Week 05

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 18, 2013

Applied Time Series Analysis
SS 2013 – Week 05

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

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

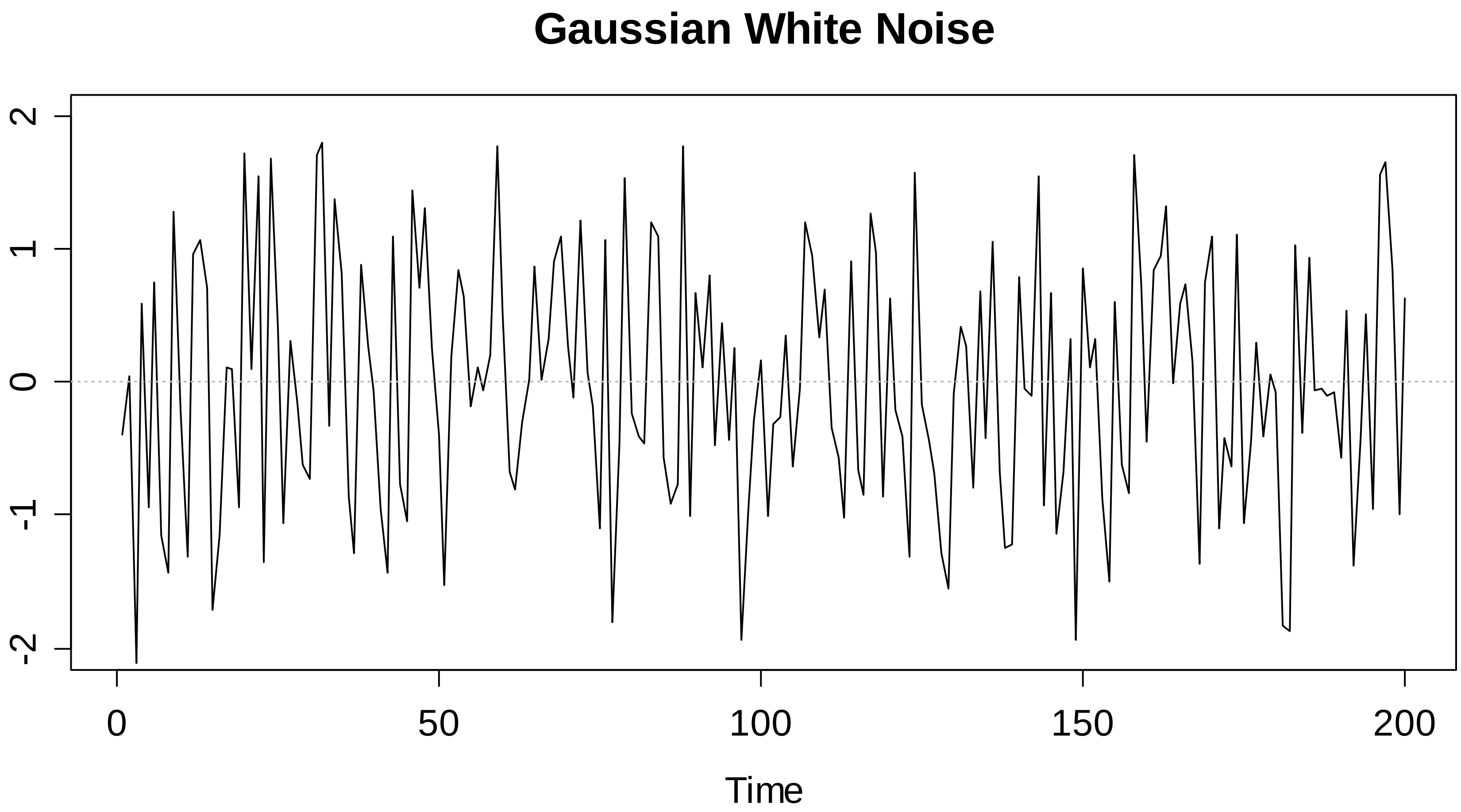
Thus, there are no autocorrelations either: $\rho_k = 0$ for all $k \neq 0$.

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**.

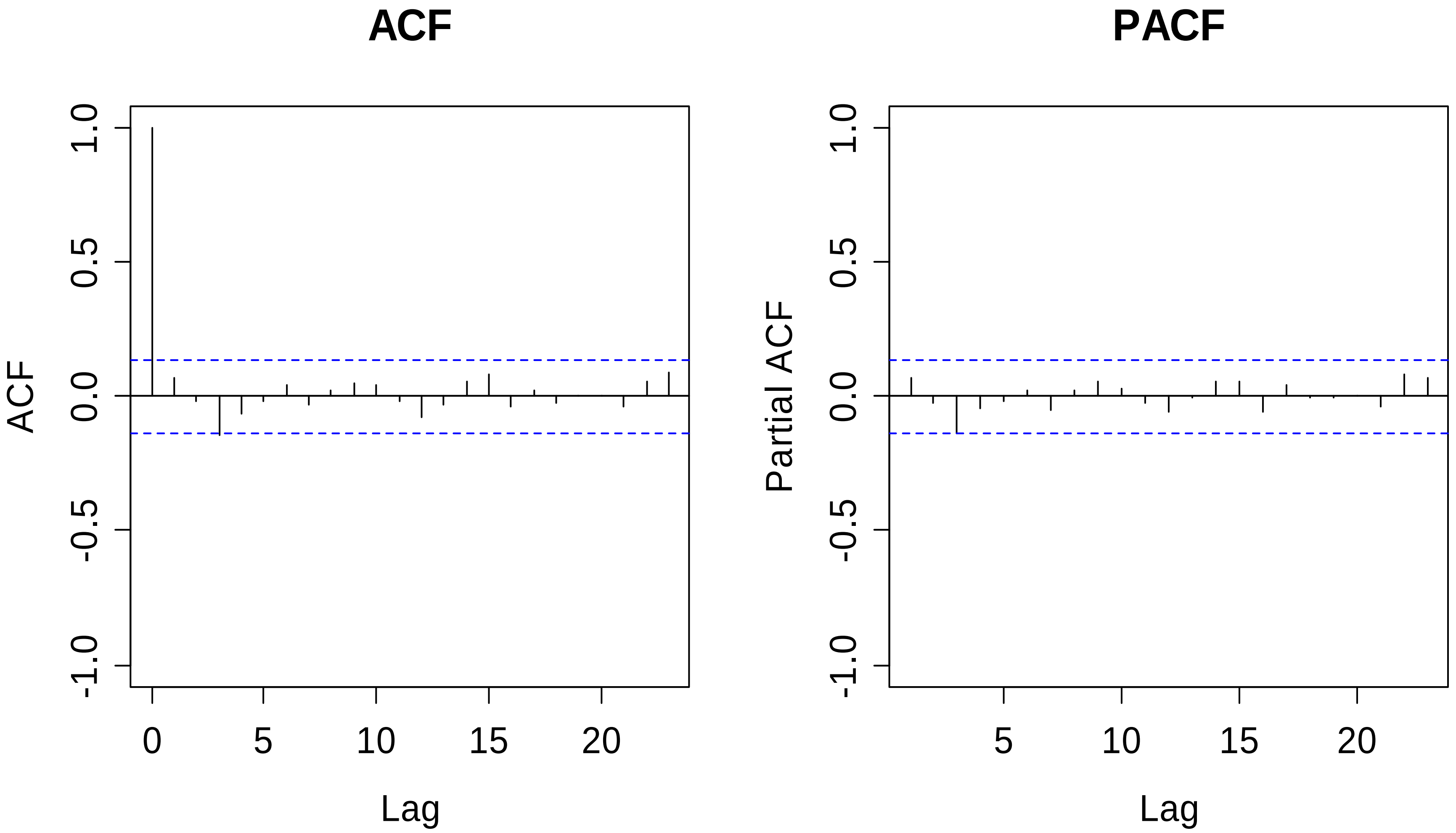
The term White Noise is due to the analogy to white light.

Example: Gaussian White Noise

```
> plot(ts(rnorm(200, mean=0, sd=1)))
```

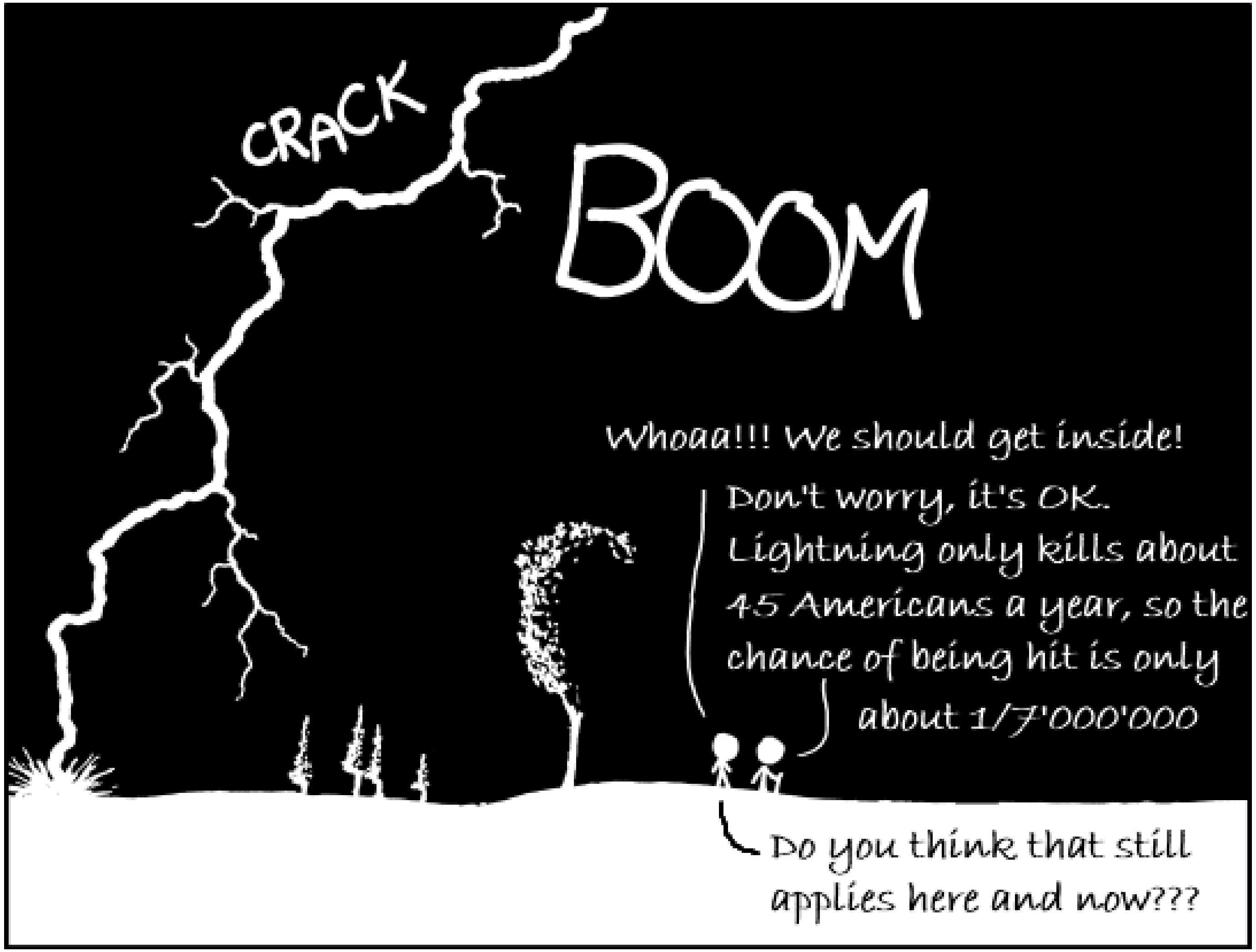


Example: Gaussian White Noise



Applied Time Series Analysis
SS 2013 – Week 05

Estimating the Conditional Mean



Applied Time Series Analysis
SS 2013 – Week 05

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
- …	

We start by discussing autoregressive models. They are perhaps the simplest and most intuitive time series models that exist.

Basic Idea for AR(p)-Models

We have a process where the random variable X_t depends on an auto-regressive linear combination of the preceding $X_{t-1},\ldots,X_{t-p}$, plus a „completely independent“ term called innovation E_t.

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

Here, *p* is called the order of the autoregressive model. Hence, we abbreviate by AR(p). An alternative notation is with the backshift operator B:

$$(1-\alpha_1 B - \alpha_2 B^2 - \ldots - \alpha_p B^p) X_t = E_t \text{ or short, } \Phi(B) X_t = E_t$$

Here, $\Phi(B)$ is called the characteristic polynomial of the AR(p). It determines most of the relevant properties of the process.

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 → E_t all are independent

All E_t are independent ~~→~~ E_t is an innovation

AR(p)-Models and Stationarity

The following is absolutely essential:

AR(p) models must only be fitted to stationary time series. Any potential trends and/or seasonal effects need to be removed first. We will also make sure that the processes are stationary.

Under which circumstances is an AR(p) stationary?

→ **see blackboard…**

Stationarity of AR(p)-Processes

We require:

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

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

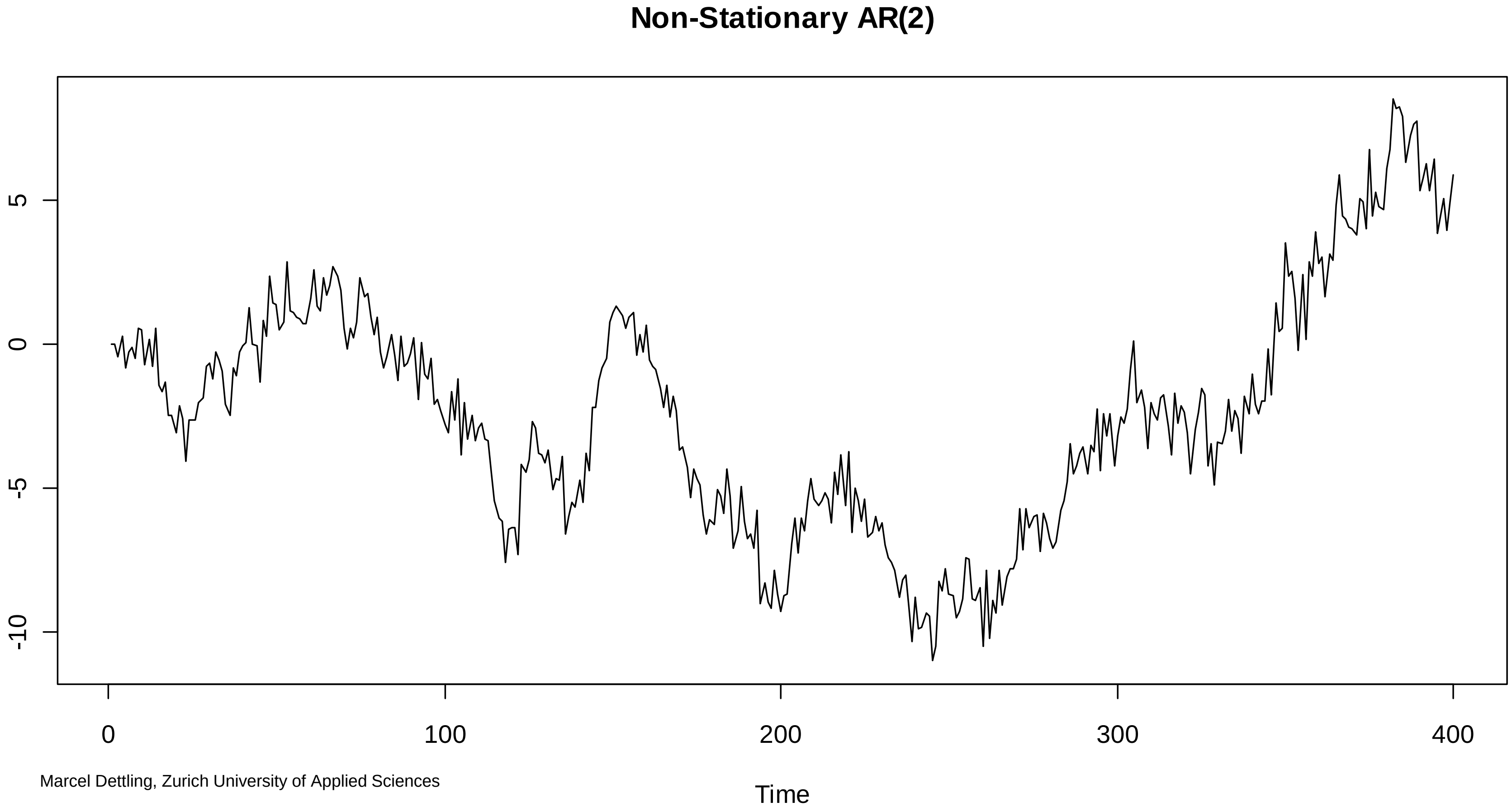
All (complex) roots of the characteristic polynom

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

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

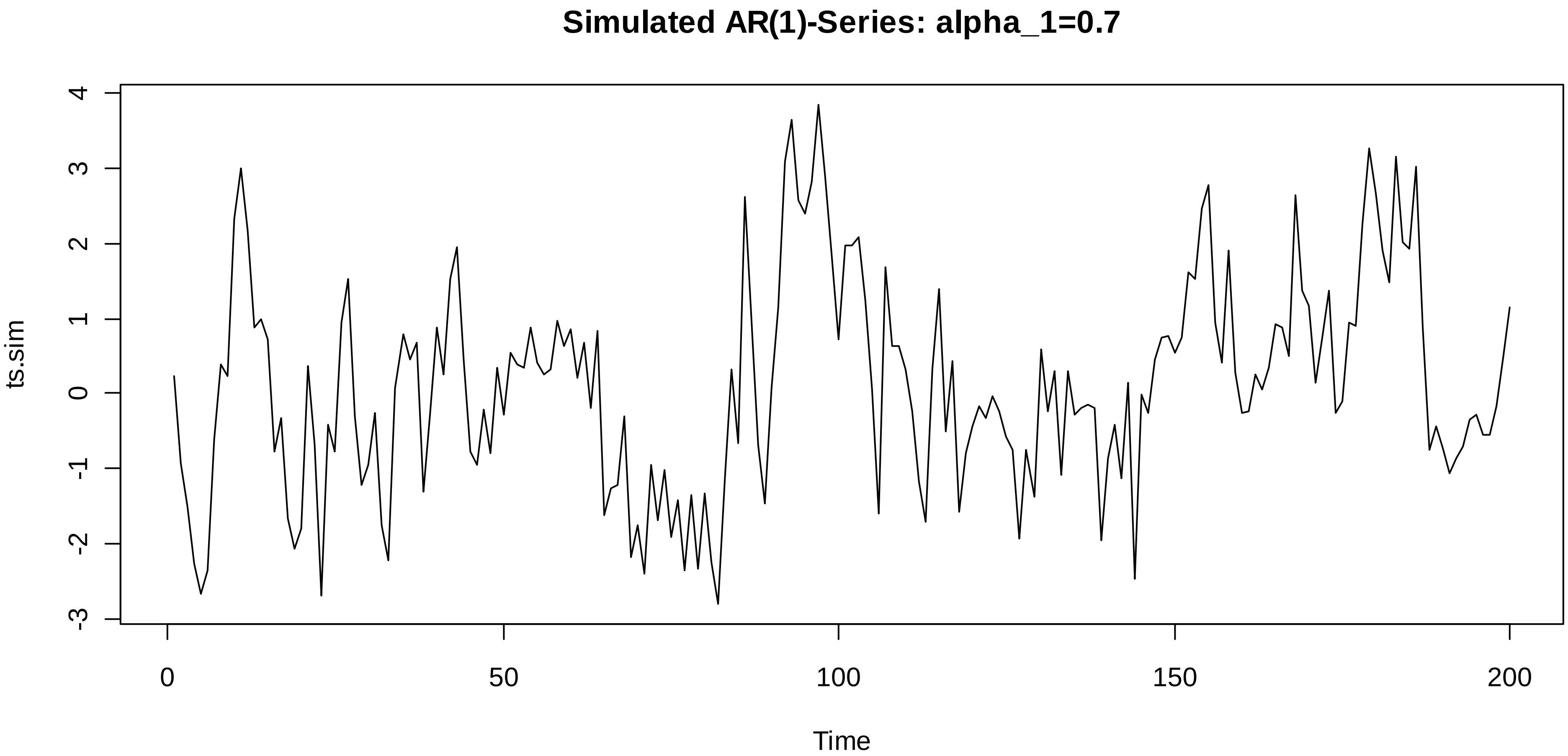
A Non-Stationary AR(2)-Process

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

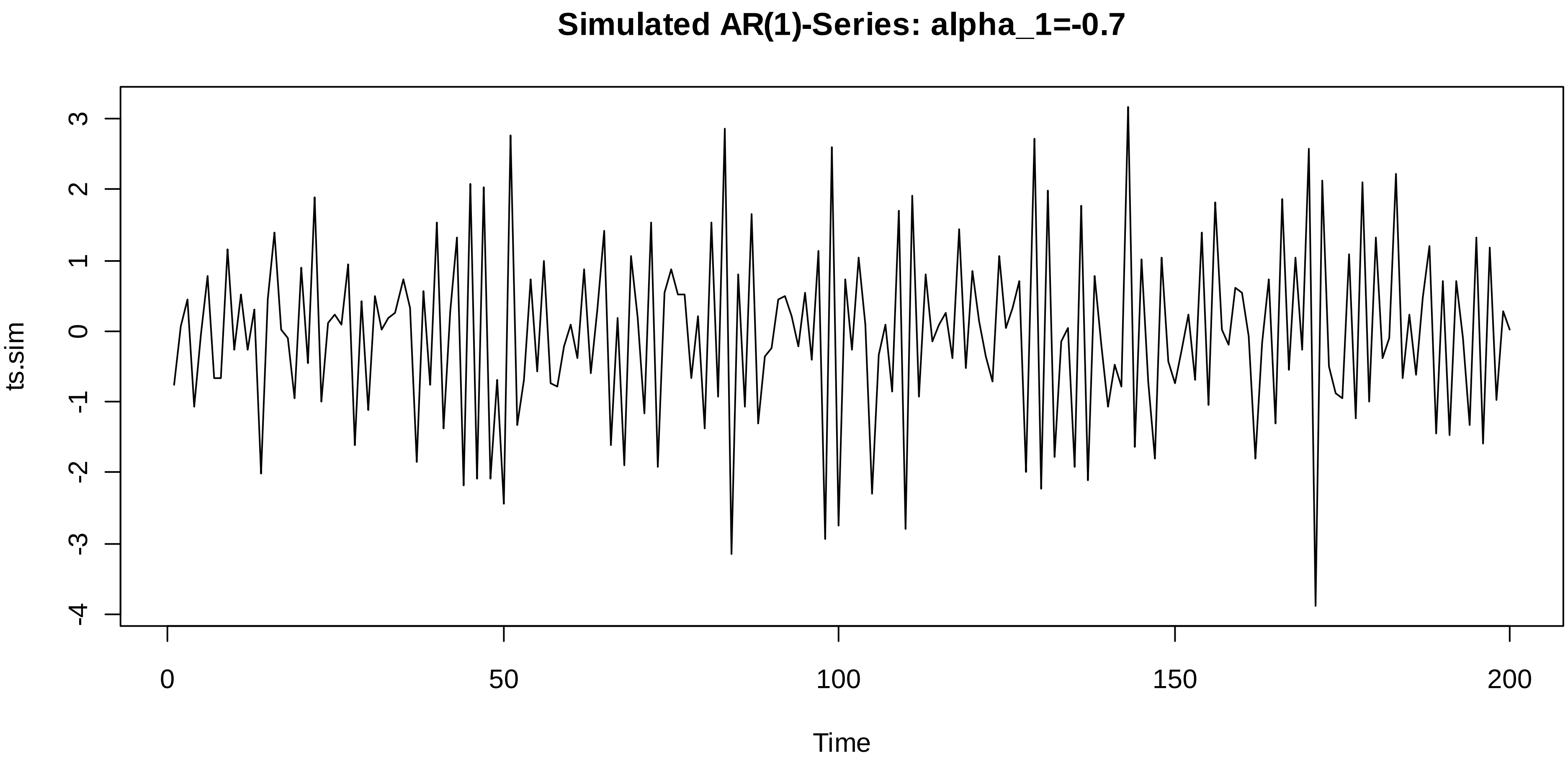


Applied Time Series Analysis
SS 2013 – Week 05

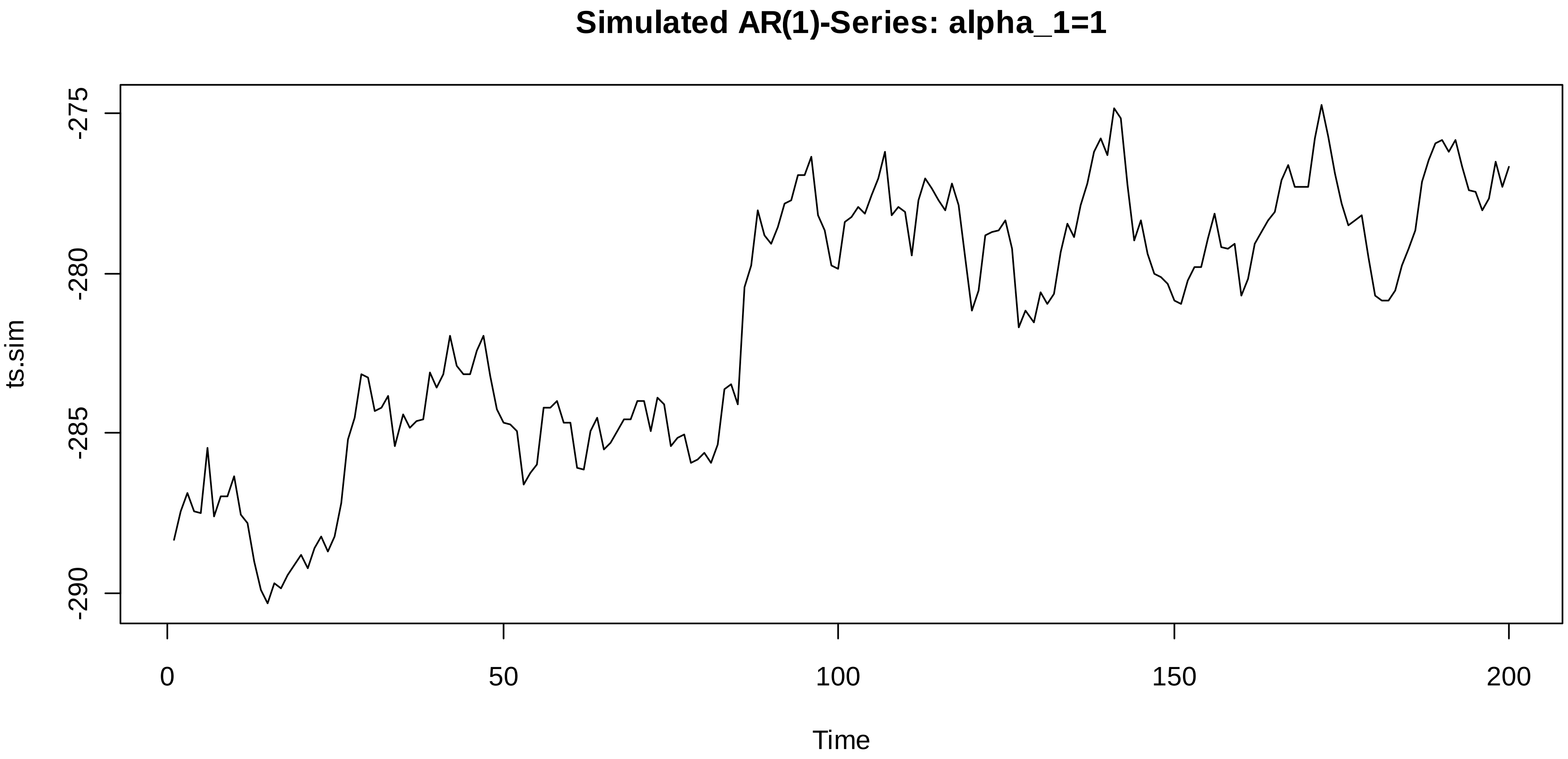
Simulated AR(1)-Series



Simulated AR(1)-Series



Simulated AR(1)-Series



Autocorrelation of AR(p) Processes

On the blackboard…

Yule-Walker Equations

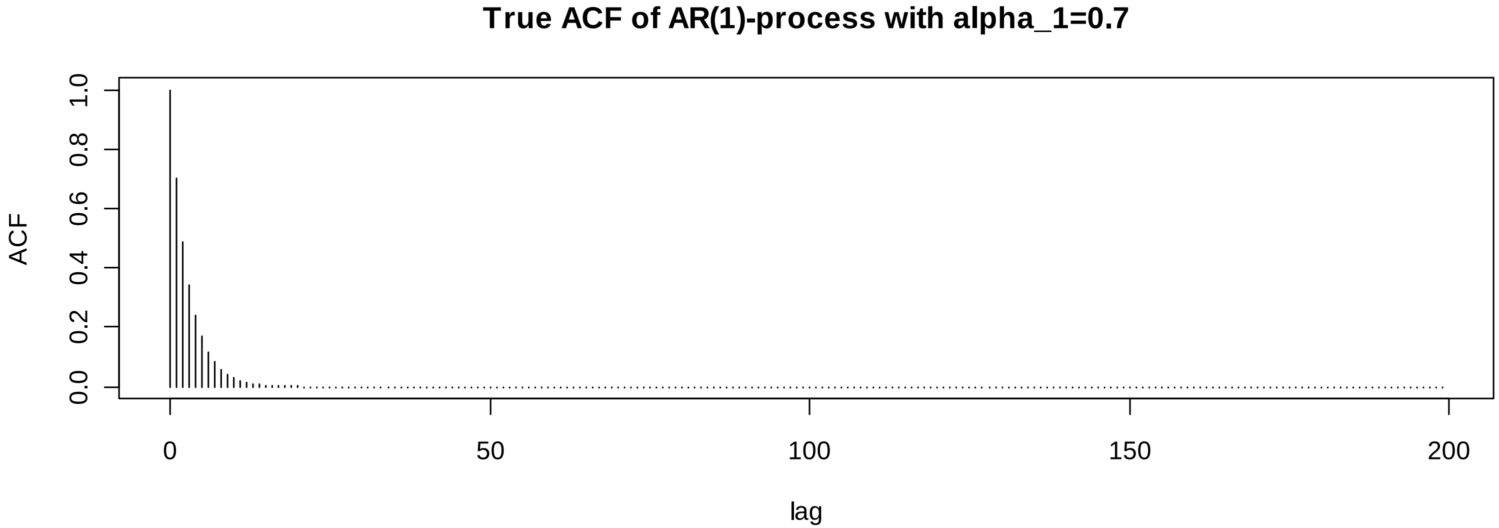
We observe that there exists a linear equation system built up from the AR(p)-coefficients and the ACF-coefficients of up to lag p. These are called Yule-Walker-Equations.

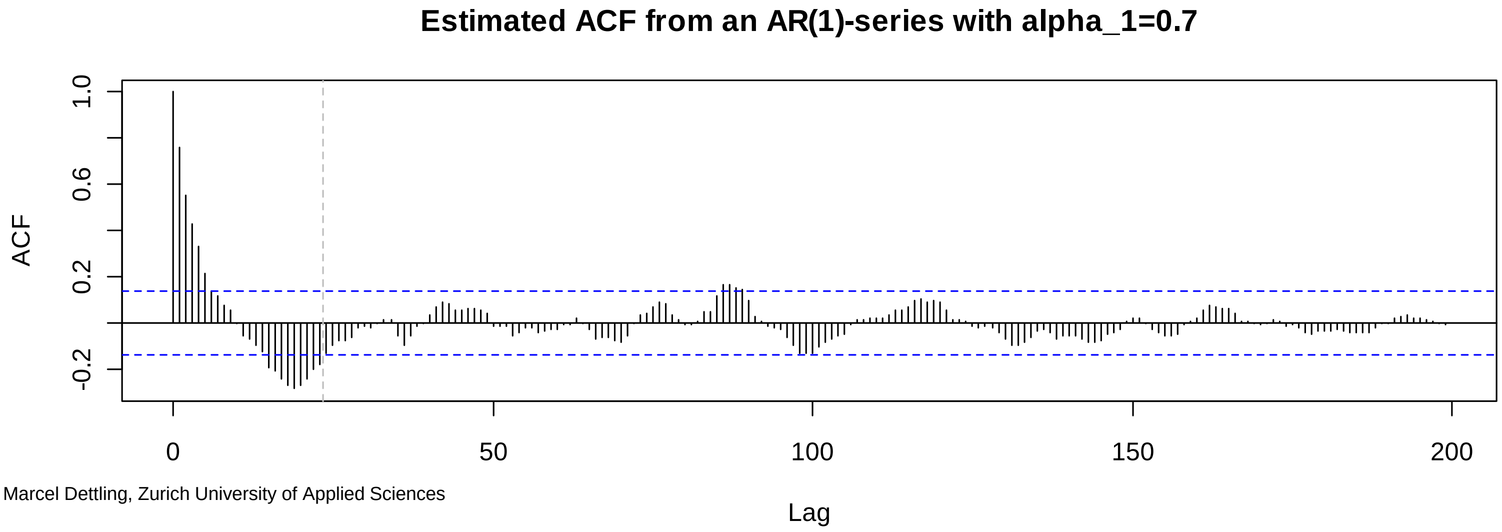
We can use these equations for fitting an AR(p)-model:

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*

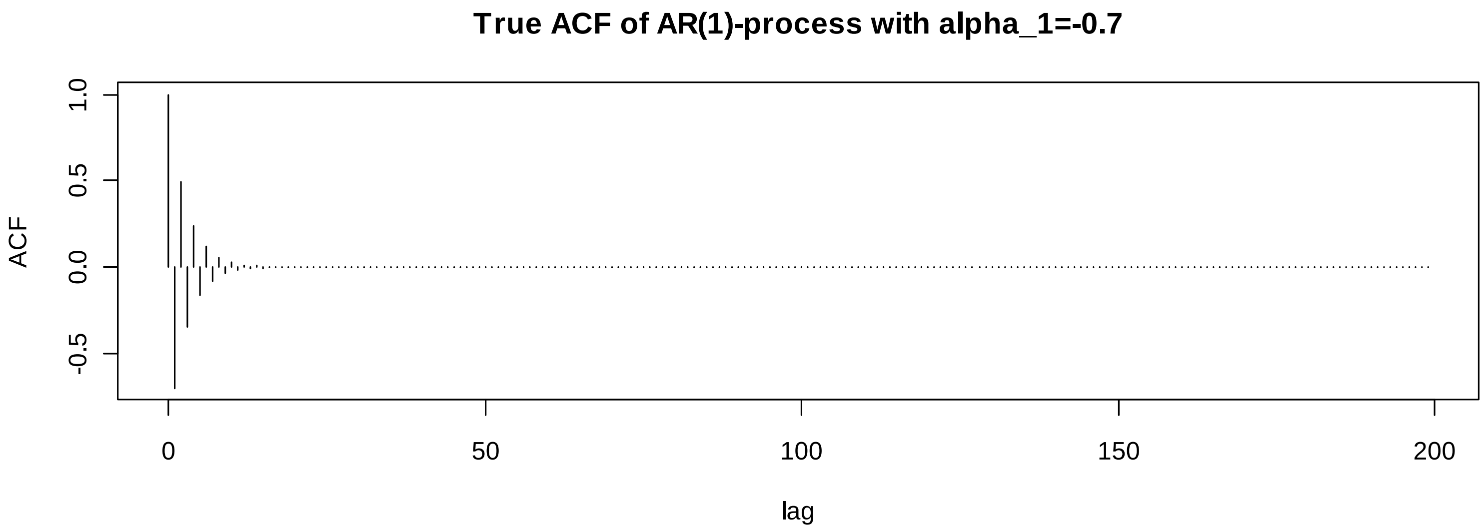
Applied Time Series Analysis
SS 2013 – Week 05

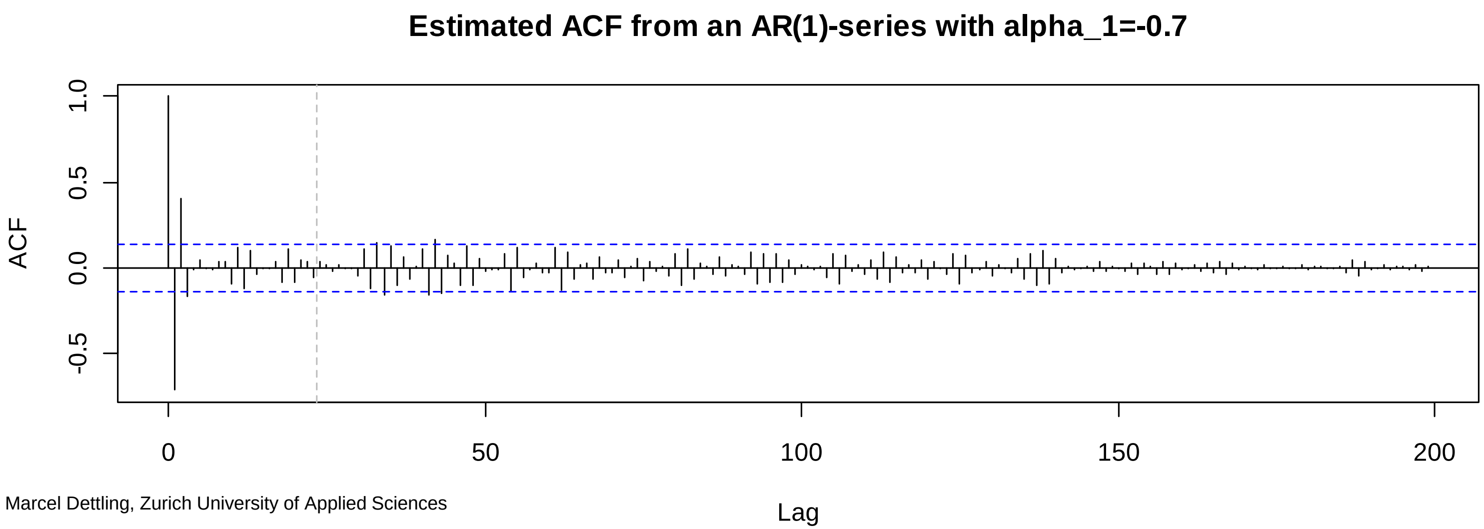
Theoretical vs. Estimated ACF





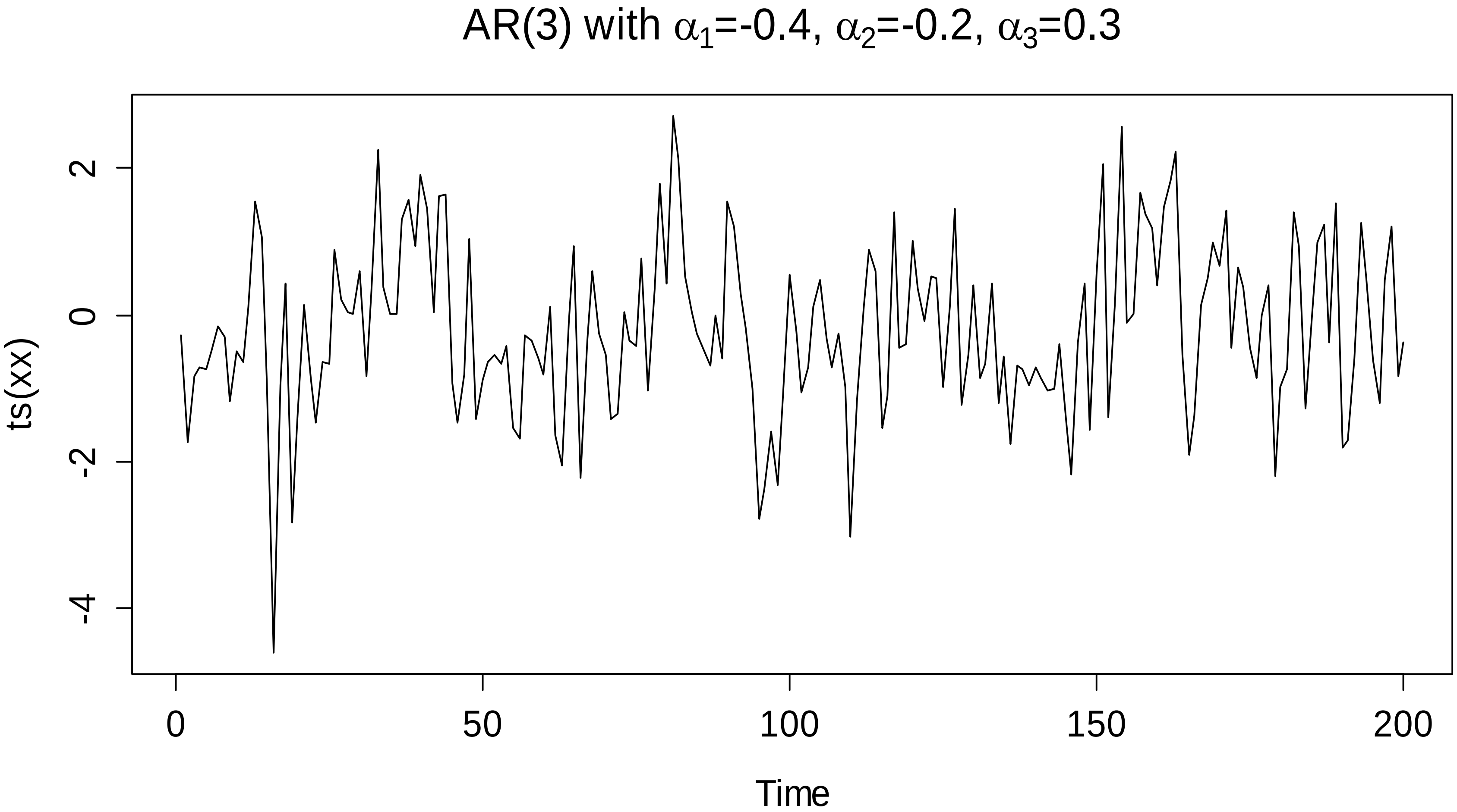
Theoretical vs. Estimated ACF





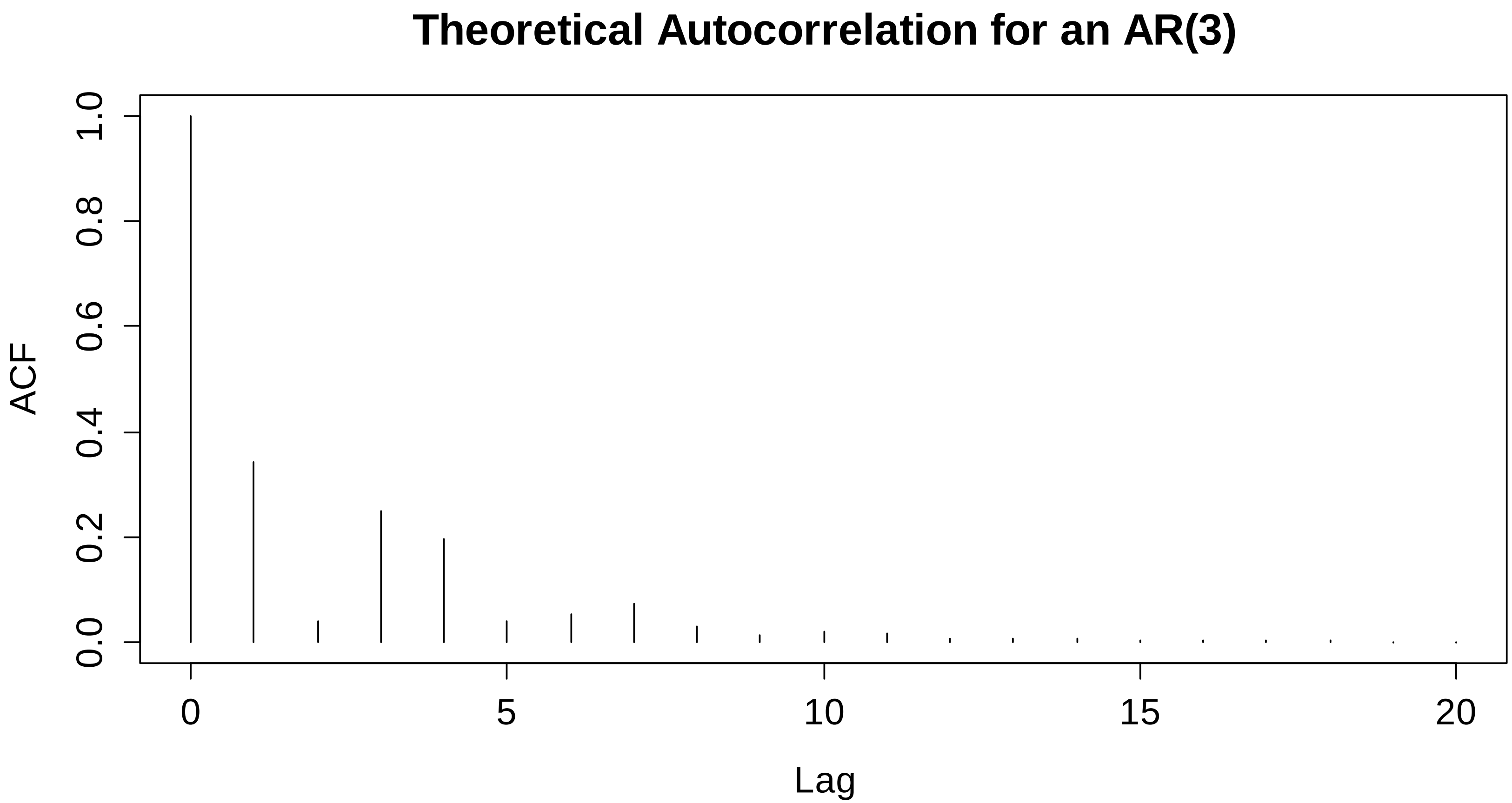
AR(3): Simulation and Properties

```
> xx <- arima.sim(list(ar=c(0.4, -0.2, 0.3)),
```



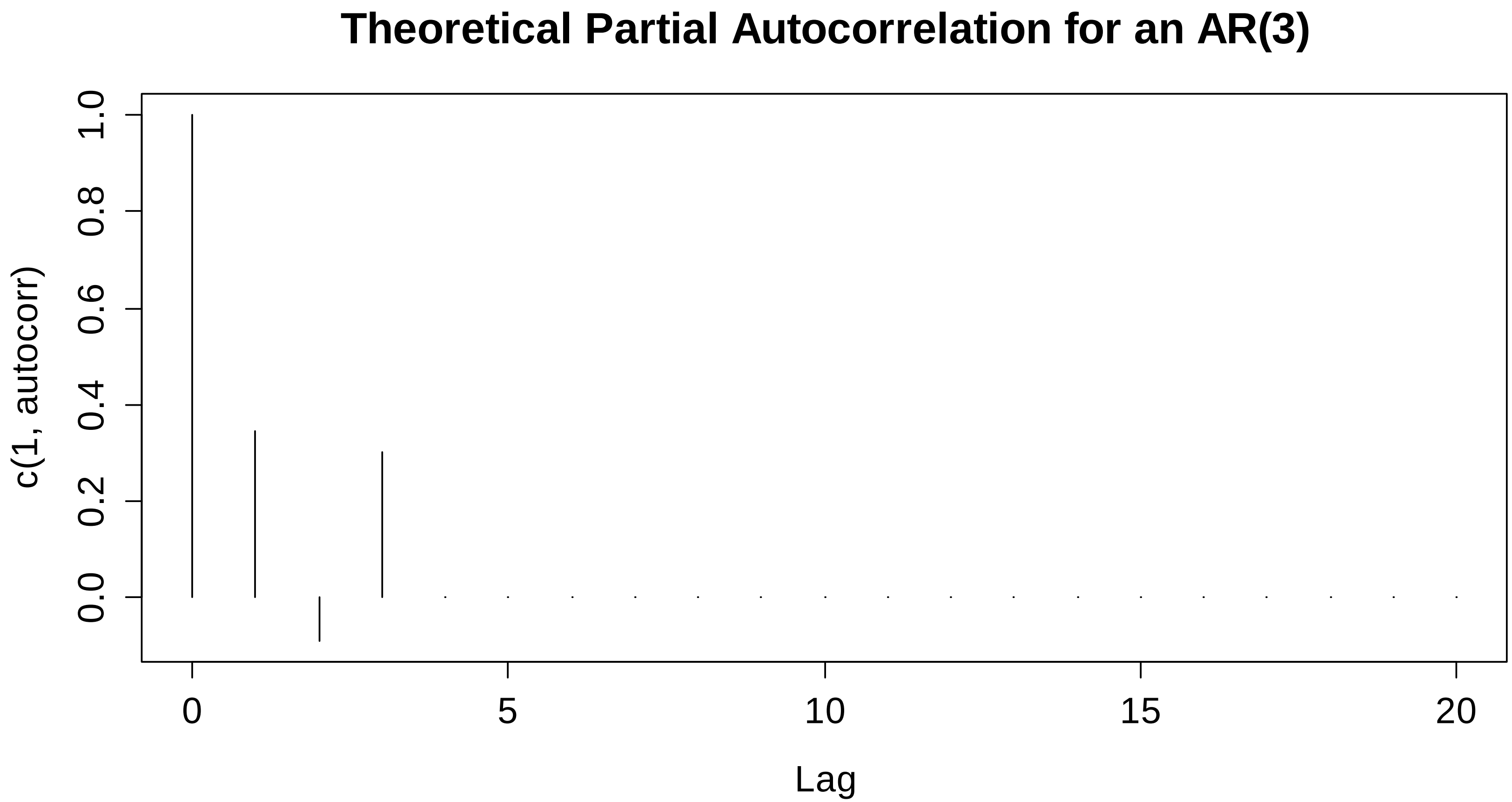
AR(3): Simulation and Properties

```
> autocorr <- ARMAacf(ar=c(0.4, -0.2, 0.3),...)
> plot(0:20, autocorr, type="h", xlab="Lag")
```



AR(3): Simulation and Properties

```
> autocorr <- ARMAacf(ar=..., pacf=TRUE, ...)
> plot(0:20, autocorr, type="h", xlab="Lag")
```



Fitting AR(p)-Models

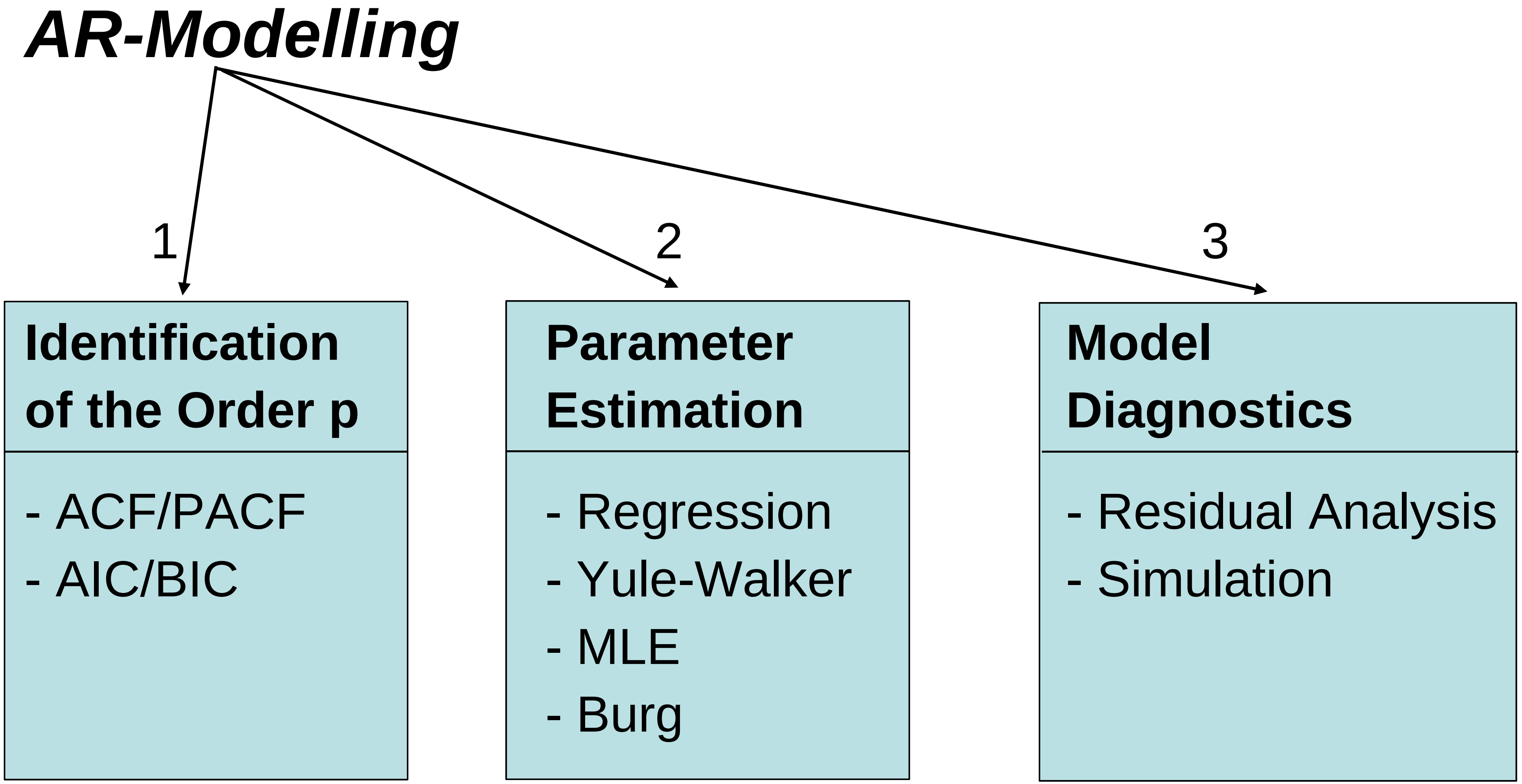
This involves 3 crucial steps:

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
SS 2013 – Week 05



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.

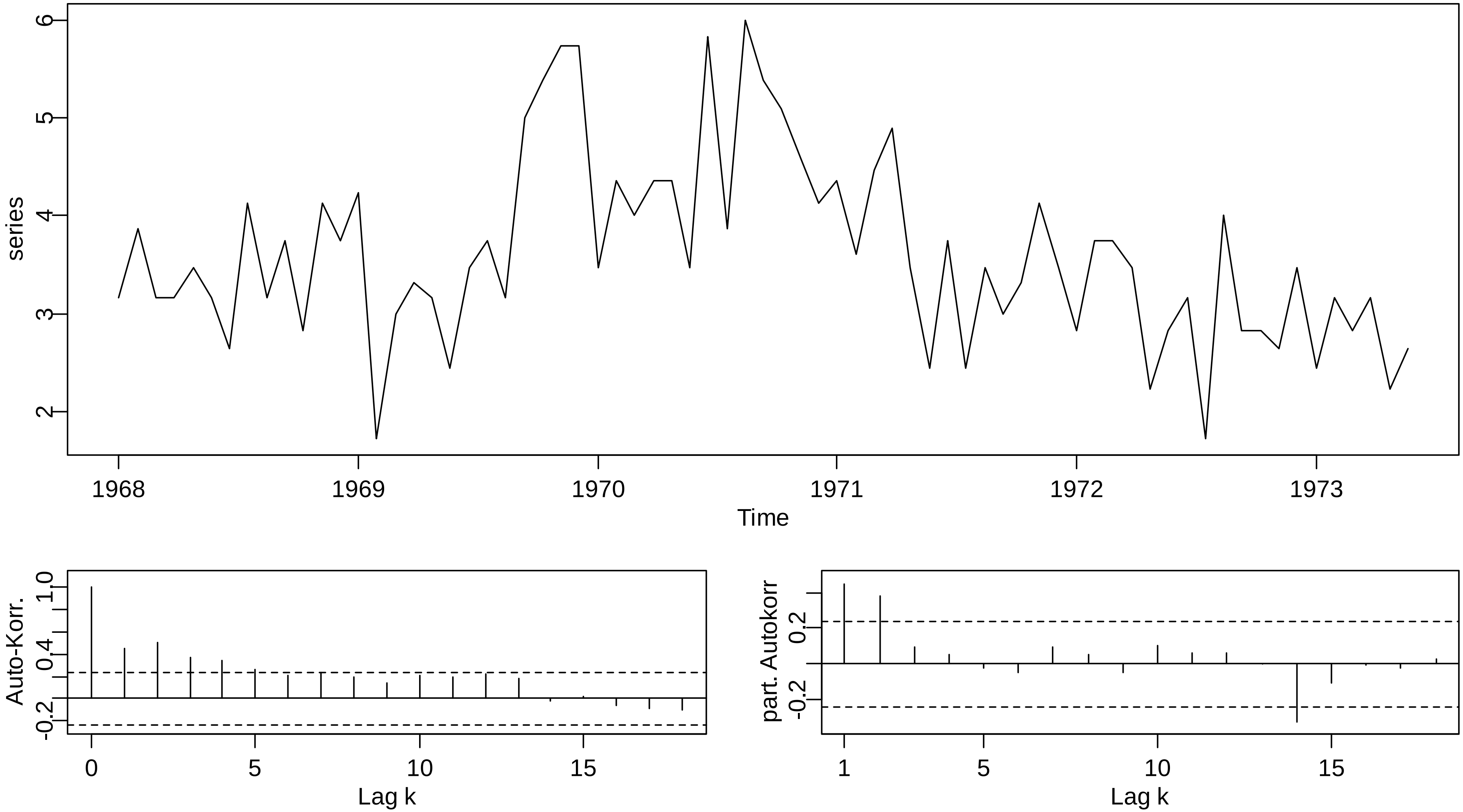
- For all AR(p)-models, the **PACF** is equal to zero for all lags k>p.

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
SS 2013 – Week 05

Model Order for sqrt(purses)



Applied Time Series Analysis
SS 2013 – Week 05

Model Order for log(lynx)