Foundations & Explorative Analyses

Exercise 1 (Basics and Motivation)

Let us first clarify some general concepts before we start with examining concrete processes and models. Consider the discrete stochastic process $\{y_t\}_{t\in\mathbb{Z}}$, of which we observe a finite number of realizations $\{1,\dots,T\}$.

(a) What is the idea behind time series analysis and its purpose? What is the main assumption?

- The past values of a quantity (economic examples: gross domestic product, asset returns) allow conclusions and enable to predict future values.
- Implicit assumption (at least for basic time series models): The observed structure of a time-varying quantity does not change essentially.
- A theoretical concept: Even in parametric time series models, estimated parameters hardly have an (economic, biological..) interpretation.

(b) Give the definition of the *mean*, *(auto-)covariance* and *(auto-)correlation function* in general. How can these functions be interpreted?

The mean functions is defined as follows

$$\mu_t = \mathbb{E}[y_t] = \int_{-\infty}^{\infty} y \cdot f_t(y)dy$$

Recall that f represents a probability density.

The ACovF of a stochastic process $\{y_t\}$, $t \in \mathbb{Z}$ is a function $\gamma : \mathbb{Z}\times\mathbb{Z} \to \mathbb{R}$ which determines the covariance of two random variables of the process at two arbitrary points in time t_1, t_2,

$$\gamma(t_1,t_2) = \mathrm{Cov}(y_{t_1}, y_{t_2}) = E[(y_{t_1} - \mu_{t_1})(y_{t_2} - \mu_{t_2})], \qquad t_1, t_2 \in \mathbb{Z}$$

where μ_t is the mean function of the process.

The ACF of a stochastic process $\{y_t\}$, $t \in \mathbb{Z}$ is a function $\rho : \mathbb{Z}\times\mathbb{Z} \to [-1,1]$ which determines the correlation of two random variables of the process at two arbitrary points in time t_1, t_2,

$$\rho(t_1,t_2) = Cor(y_{t_1}, y_{t_2}) = \frac{\gamma(t_1,t_2)}{\sqrt{\gamma(t_1,t_1)\cdot\gamma(t_2,t_2)}} \qquad t_1, t_2 \in \mathbb{Z}$$

(c) One central concept in time series analysis is *stationarity*. Why? How is *weak* and *strict* stationarity defined? Of what form are the autocovariance and autocorrelation in context of stationary time series?

As mentioned above, in order to predict future values of time series, one must assume some kind of stability over time. However, it should be emphasized that *time series* and *time series models/ stochastic processes* have to be distinguished! We only can say whether an estimated time series model/ an underlying DGP is stationary.

There are several definitions of stationarity of a stochastic process $\{y_t\}_{t\in\mathbb{Z}}$:

- mean stationarity: $E(y_t) = \mu_t = \mu = const.\ \forall t \in \mathbb{Z}$,

$$\exists \mu \in \mathbb{R} \forall t \in \mathbb{Z} : E[y_t] = \mu$$

- variance stationarity: $\text{Var}(y_t) = const.\ \forall t \in \mathbb{Z}$,

$$\exists \sigma^2 \in \mathbb{R} \forall t \in \mathbb{Z} : \text{Var}[y_t] = \sigma^2$$

- covariance stationarity: $\gamma(t_1, t_2) = \gamma(t_1 + h, t_2 + h) \quad \forall t_1, t_2, h \in \mathbb{Z}$,
- weak stationarity: mean stationarity + covariance stationarity,
- strict stationarity: The probabilistic behavior of every collection of values

$$\{y_{t_1}, y_{t_2}, \ldots, y_{t_k}\}$$

is identical to that of the time shifted set

$$\{y_{t_{1+h}}, y_{t_{2+h}}, \ldots, y_{t_{k+h}}\},$$

that is

$$P(y_{t_1} \le c_1, \ldots, y_{t_k} \le c_k) = P(y_{t_{1+h}} \le c_1, \ldots, y_{t_{k+h}} \le c_k)$$

for all $k \in \mathbb{N}$, all time points $t_1, \ldots \in \mathbb{Z}$, all $c_1, \ldots \in \mathbb{R}$, and all time shifts $h \in \mathbb{Z}$.

For a stationary time series the ACovF becomes

$$\gamma(h) = \text{Cov}(y_t, y_{t+h}) = E[(y_t - \mu)(y_{t+h} - \mu)], \qquad h \in \mathbb{Z}$$

For a stationary time series the ACF becomes

$$\rho(h) = Cor(y_t, y_{t+h}) = \frac{\gamma(t+h,t)}{\sqrt{\gamma(t,t) \cdot \gamma(t+h,t+h)}}$$

(d) One major building block of discrete time series models are *white noise processes*. How are these processes defined? What is the purpose of this building block?

A sequence $\{y_t\} = \{\varepsilon_t\}$ is called *white noise* when its elements are uncorrelated, it has zero mean and a constant variance. This means its elements are not necessarily independent (if not explicitly stated) except when the ε's are normal (Gaussian white noise).

Exercise 2 (Linear Processes)

A linear process is defined as a linear combination of white noise variates ε_t with $t \in \mathbb{Z}$, a constant $\mu \in \mathbb{Z}$ and weights $\psi_j \in \mathbb{R}$. The process is of the form

$$y_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j \varepsilon_{t-j}. \tag{1}$$

Furthermore, the sum sum over the coefficients/weights is well-defined, i.e.

$$\sum_{j=-\infty}^{\infty} |\psi_j| < \infty. \tag{2}$$

Show that the autocovariance function is give by

$$\gamma(h) = \sigma_\varepsilon^2 \sum_{j=-\infty}^{\infty} \psi_{j+h}\psi_j \quad \text{with} \quad h \in \mathbb{Z}. \tag{3}$$

Hint: Which pairs of the white noise variates do not vanish?

$$\begin{aligned}
\gamma(h) &= \mathrm{Cov}(y_t, y_{t+h}) \\
&= \mathbb{E}[(\mu + \sum_{j=-\infty}^{\infty} \psi_j \varepsilon_{t-j} - \mu) \cdot (\mu + \sum_{k=-\infty}^{\infty} \psi_k \varepsilon_{t+h-k} - \mu)] \\
&= \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \psi_k \psi_j \mathbb{E}[\varepsilon_{t+h-k}\varepsilon_{t-j}] \\
&= \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \psi_k \psi_j \delta_{j,k-h} \sigma_\varepsilon^2 \\
&= \sigma_\varepsilon^2 \sum_{j=-\infty}^{\infty} \psi_j \psi_{j+h}
\end{aligned}$$

in this case $\delta_{j,k-h}$ is called Kronecker delta which is defined by

$$\delta_{j,k-h} = \begin{cases} 1 & j = k - h \\ 0 & \text{otherwise} \end{cases}$$

It serves as a selection function for those ε with the same time index.

Exercise 3 (Components of Time Series)

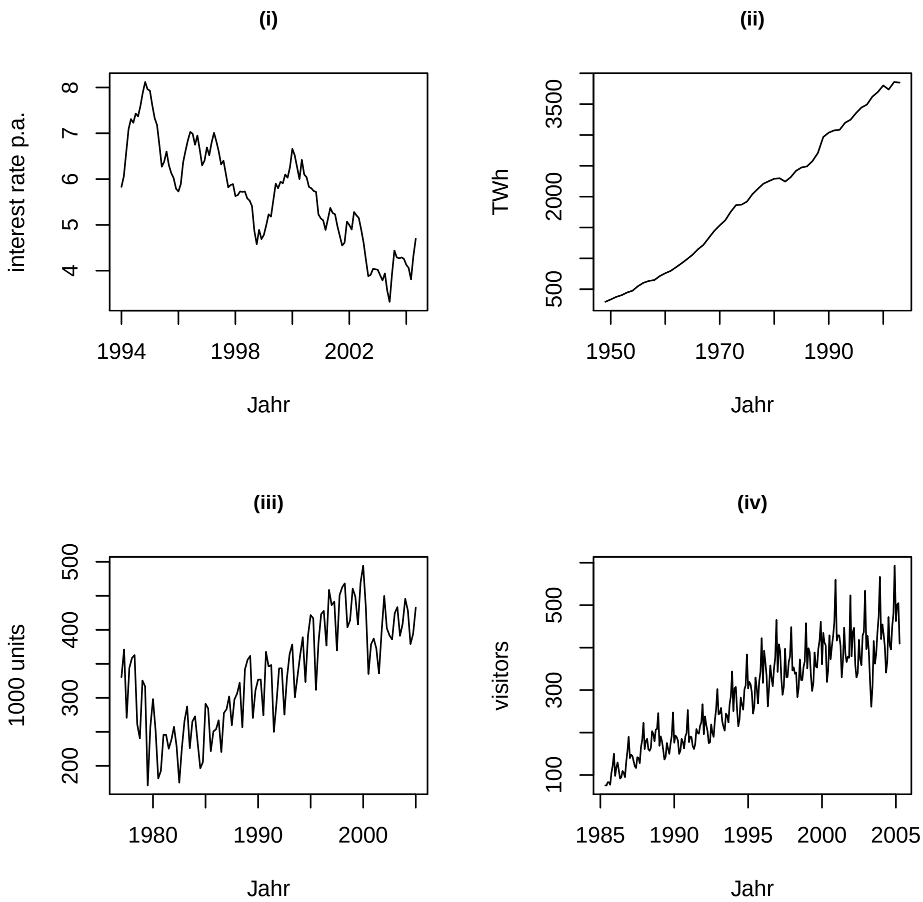
The following figures depict time series[1] of

(i) monthly interest rates of US treasury bonds with time to maturity of 10 years (January 1994 to May 2004),

(ii) yearly electricity production in the US (1949 to 2003),

(iii) quarterly UK car production (1977/1 to 2005/1),

(iv) monthly number of tourists in Australia (May 1985 to April 2005)



(a) Explain the three components of time series by means of these figures.

Four components of time series:
$y_1, \ldots, y_T$ (observed time series) can be decomposed into

$$y_t = d_t + c_t + s_t + \epsilon_t$$

- The **deterministic trend** d_t captures the development of the series in the long run.
- The **cyclical component** c_t collects all medium- to long-run regular deterministic behavior (e.g. business cycles which span over several years).
- The **seasonal component** s_t is a deterministic periodically pattern, which is repeated with a fixed period like daily, weekly, monthly, quarterly, etc.
- The difference between $d_t + c_t + s_t$ and y_t is the so-called **residual component** ϵ_t. It's assumed to be stochastic and stationary.

[1]The datasets are taken from the R package `expsmooth`: `bonds`, `usnetelec`, `ukcars` and `visitors`.

(b) One can distinguish between additive and multiplicative conjunction of the components. Which kind of conjunction do you think is appropriate for time series (i) to (iv)? What are the consequences for further analyses?

Additive model:

$$y_t = d_t + c_t + s_t + \epsilon_t$$

Multiplicative model:

$$y_t = d_t \cdot c_t \cdot s_t \cdot \epsilon_t$$

$\rightarrow$ Within the multiplicative model, the size of the cyclical component depends on the trend, i.e. it grows or shrinks with time evolving. Within the additive model, the amplitude of the cyclical component is constant.

Taking logs of the multiplicative model yields the additive model,

$$\begin{aligned} y_t &= d_t \cdot c_t \cdot s_t \cdot \epsilon_t \\ log(y_t) &= log(d_t \cdot c_t \cdot s_t \cdot \epsilon_t) \\ &= log(d_t) + log(c_t) + log(s_t) + log(\epsilon_t) \end{aligned}$$

Our examples:

1. **Yield data**: Trend plus business cycles with constant amplitude; additive model.
2. **Electricity data**: linear trend, no cyclicality; additive model.
3. **Car production data**: Yearly cyclicality, trend between 1980 and 2000, observations before and after this period contradict this trend; rather additive model.
4. **Number of tourists data**: Trend and seasonality (yearly and sub-yearly); multiplicative model because of growing amplitude(s).

Exercise 4 (Deterministic Trend)

Consider the process

$$y_t = \beta_1 + \beta_2 t + \varepsilon_t \tag{4}$$

where β_1, $\beta_2 \in \mathbb{R}$ are known and non-zero weights. Let $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ be a white noise process with variance σ^2.

(a) Determine whether the process y_t is stationary.

The process y_t is not stationary. To see this, it suffices to compute the mean, which is time-dependent:

$$E(y_t) = \beta_1 + \beta_2 t + \underbrace{E(\varepsilon_t)}_{=0} = \beta_1 + \beta_2 t.$$

(b) Show that the process $x_t = y_t - y_{t-1}$ is stationary.

$$\begin{aligned} x_t &= \beta_1 + \beta_2 t + \varepsilon_t - (\beta_1 + \beta_2(t-1) + \varepsilon_{t-1}) \\ &= \beta_2 + \varepsilon_t - \varepsilon_{t-1}. \end{aligned}$$

Hence:

$$\begin{aligned} E(x_t) &= \beta_2 \\ \mathrm{Cov}(x_t, x_{t-h}) &= \mathrm{Cov}(\beta_2 + \varepsilon_t - \varepsilon_{t-1}, \beta_2 + \varepsilon_{t-h} - \varepsilon_{t-1-h}) \\ &= \mathrm{Cov}(\varepsilon_t, \varepsilon_{t-h}) + \mathrm{Cov}(\varepsilon_t, -\varepsilon_{t-h-1}) + \mathrm{Cov}(-\varepsilon_{t-1}, \varepsilon_{t-h}) + \mathrm{Cov}(-\varepsilon_{t-1}, -\varepsilon_{t-h-1}) \\ &= \begin{cases} 2\sigma^2, & h = 0, \\ -\sigma^2, & |h| = 1 \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

Neither the mean nor the covariance are time-dependent, the latter particularly depends on h, not on t. Therefore the process is shown to be weakly stationary.

Exercise 5 (Smoothing: Linear Filter)

A general linear filter M_t of the t-th value of a time series is defined by

$$M_t = \sum_{i=-\infty}^{\infty} \lambda_i y_{t+i}. \tag{5}$$

Let $b \in \mathbb{Z}$. One specific linear filter, the running mean, can be defined as

$$D_t = \sum_{i=-b}^{b} \frac{1}{2b+1} y_{t-i}. \tag{6}$$

(a) Calculate the mean of the filter D_t with x_t as defined in exercise 4, i.e. $y_t = \beta_1 + \beta_2 t + \varepsilon_t$.

$$v_t = \frac{1}{2b+1} \sum_{j=-b}^{b} (\beta_1 + \beta_2(t-j) + \varepsilon_{t-j})$$

Hence

$$\begin{aligned}
E(v_t) &= \frac{1}{2b+1} \sum_{j=-b}^{b} (\beta_1 + \beta_2(t-j) + \underbrace{E(\varepsilon_{t-j})}_{=0}) \\
&= \frac{1}{2b+1} [(2b+1)\beta_1 + \beta_2 \sum_{j=-b}^{b} (t-j)] \\
&= \frac{1}{2b+1} [(2b+1)\beta_1 + \beta_2(2b+1)t + 0] \\
&= \beta_1 + \beta_2 t.
\end{aligned}$$

(b) Calculate the corresponding autocovariance function $\gamma(h)$ of (**??**) and derive a simplified expression.

$$\text{Cov}(v_t, v_{t-h}) = \text{Cov}\left(\frac{1}{2b+1} \sum_{j=-b}^{b} (\beta_1 + \beta_2(t-j) + \varepsilon_{t-j}), \frac{1}{2b+1} \sum_{j=-b}^{b} (\beta_1 + \beta_2(t-j-h) + \varepsilon_{t-j-h}) \right).$$

The deterministic parts of v_t and v_{t-h} can be dropped and we obtain

$$\begin{aligned}
&= \frac{1}{(2b+1)^2} \text{Cov}\left(\sum_{j=-b}^{b} \varepsilon_{t-j}, \sum_{j=-b}^{b} \varepsilon_{t-j-h} \right) \\
&= \frac{\sigma^2(2b+1-|h|)}{(2b+1)^2},
\end{aligned}$$

for $h = 0, \pm 1, \pm 2, \dots, \pm 2b$, and zero for $|h| > 2b$.

Exercise 6 (Random Walk)

Consider the random walk with drift model

$$y_t = \delta + y_{t-1} + \varepsilon_t \tag{7}$$

for $t \in \mathbb{N}$, with $y_0 = 0$, where $\{\varepsilon_t\}_{t\in\mathbb{N}}$ is white noise with variance σ^2.

(a) Show that the model can be written as

$$y_t = \delta t + \sum_{k=1}^{t} \varepsilon_k. \tag{8}$$

$$\begin{aligned} y_t &= \delta + y_{t-1} + \varepsilon_t \\ &= \delta + (\delta + y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\ &= \delta + (\delta + (\delta + y_{t-3} + \varepsilon_{t-2}) + \varepsilon_{t-1}) + \varepsilon_t \\ &= \dots \\ &= \delta t + \varepsilon_t + \dots + \varepsilon_1 \\ &= \delta t + \sum_{k=1}^{t} \varepsilon_k. \end{aligned}$$

(b) Find a representation of the mean function and the autocovariance function of y_t.

$$E(y_t) = \delta t + 0,$$

$$\begin{aligned} \mathrm{Cov}(y_t, y_{t-h}) &= \mathrm{Cov}\left(\delta t + \sum_{k=1}^{t} \varepsilon_k, \delta(t-h) + \sum_{l=1}^{t-h} \varepsilon_l\right) \\ &= \mathrm{Cov}\left(\sum_{k=1}^{t} \varepsilon_k, \sum_{l=1}^{t-h} \varepsilon_l\right) \\ &= \min(t, t-h)\sigma^2. \end{aligned}$$

Note that $h \in \mathbb{Z}$.

(c) Show

$$\rho(t-1,t) = \sqrt{\frac{t-1}{t}} \to 1 \text{ as } t \to \infty. \tag{9}$$

What is the implication of this result?

$$\begin{aligned}
\rho_y(t-1,t) &= \frac{\mathrm{Cov}(y_{t-1}, y_t)}{\sqrt{\mathrm{Var}(y_{t-1})\mathrm{Var}(y_t)}} \\
&= \frac{(t-1)\sigma^2}{\sqrt{(t-1)t\sigma^2}} \\
&= \sqrt{\frac{t-1}{t}}.
\end{aligned}$$

Obviously

$$\lim_{t\to\infty} \sqrt{\frac{t-1}{t}} = 1.$$

This implies that in the far future new ε's do not influence y_t very much. In other words — in terms of prediction — the best one-step-ahead forecast is given by its own realization in the present.