Series 2

1. The performance of a machine that analyses the creatine content of human muscular tissue is to be investigated using 157 known samples. These samples are given to the machine one at a time for it to determine their creatin content.

The data are from an investigation into the correct functioning of automated analysis machines. You can find them in the dataset

`http://stat.ethz.ch/Teaching/Datasets/WBL/kreatin.dat` .

In this exercise, we will focus on one of the variables in this dataset, namely `gehalt` (content).

a) Which stochastic model should this series of data follow if the machine is working correctly?

b) Use the time series plot, the autocorrelations and the partial autocorrelations to determine whether or not these data fit the ideal model found in Part a).

R hints:

Converting the data frame (`d.creatine`) to a time series:

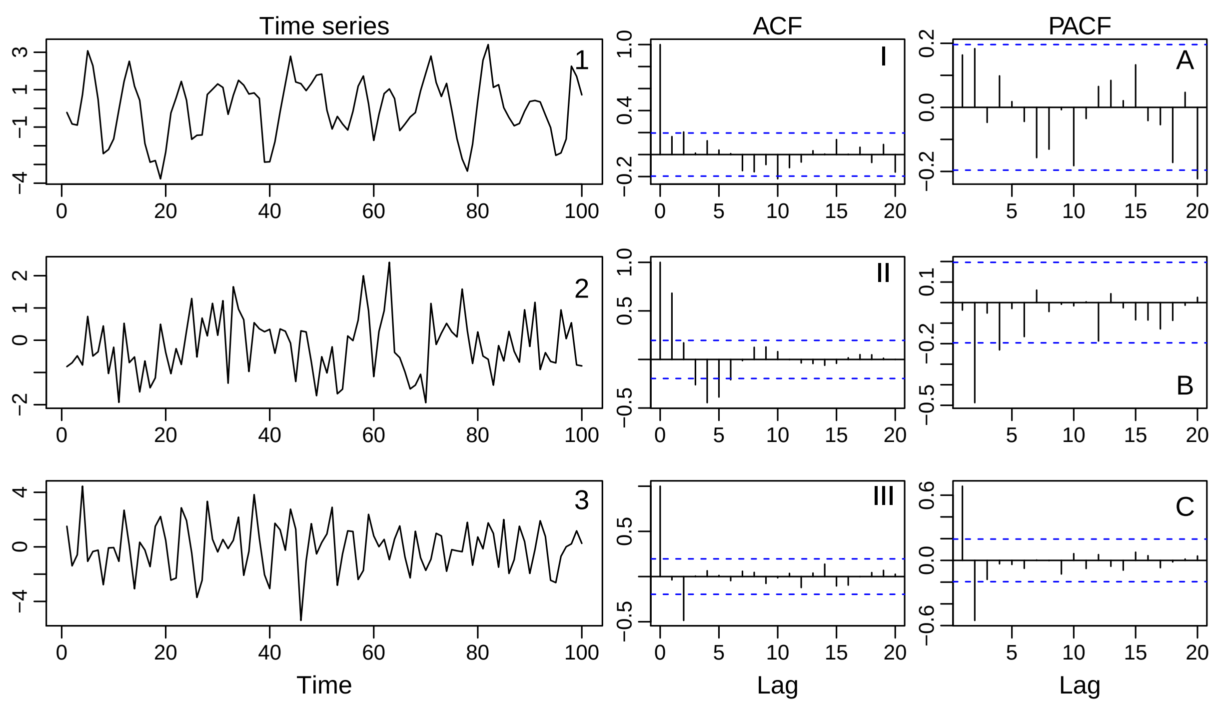
```
> t.creatine <- ts(d.creatine[, 2], start = 1, frequency = 1)
```

Plotting ACF and PACF:

```
> acf(t.creatine, plot = TRUE)
> acf(t.creatine, type = "partial", plot = TRUE)
```

2. In the figure below, three time series and their correlograms for usual and partial autocorrelations can be seen. Unfortunately, we have forgotten which correlogram belongs to which time series. Can you help?

Hint: First match up the usual autocorrelations with the time series they come from.



3. In this exercise, we consider the `AirPassengers` dataset, a time series indicating the monthly numbers of international airline passengers departing from the USA in the years 1949 to 1960. We use different methods to decompose the time series into trend, seasonal effect and remainder (cf. Series 1, Exercise 1) and compare the remainders by looking at their correlogram.

a) The `AirPassengers` dataset is a sample time series provided by R ; you don't have to read it in. Look at a plot of the time series:

```
> plot(AirPassengers)
```

Why is the correlogram of this time series not meaningful? Explain in a few sentences.

b) Decompose the time series into trend, seasonal component and remainder using the R function `decompose()`; plot the remainder and its correlogram. Interpret the plots in few sentences.

R hints:

```
> airpass.decomp <- decompose(AirPassengers, type = "multiplicative")
> plot(...)
> acf(..., na.action = na.pass, plot = TRUE)
```

See Series 1, Exercise 1 for hints on extracting the remainder from the object `airpass.decomp`, or use the R-help: `?decompose`. The function uses a filter to estimate the trend; therefore, the first and the last few entries of the decomposition are not defined (value `NA` in R). For this reason we have to use the parameter `na.action = na.pass`, otherwise R complains about missing values.

c) Decompose the log-transformed time series using the R function `stl()`. Estimate the seasonal effect once by averaging (parameter `s.window = "periodic"`) and once by choosing an appropriate smoothing window (parameter `s.window = ...`, where you have to choose an odd integer; cf. Series 1, Exercise 1.c). To determine an appropriate smoothing window, you can look at the monthplot, cf. Series 1, Exercise 2.d.

For both estimation approaches (averaging and smoothing window), plot the remainder and its correlogram, and comment on the plots.

R hints:

```
> airpass.stl <- stl(log(AirPassengers), s.window = ...)
> plot(...)
> acf(..., plot = TRUE)
```

d) Explain why you used the parameter `type = "multiplicative"` in Task b), and why you log-transformed the time series before performing an `stl()` decomposition in Task c).

Preliminary discussion: Monday, March 14.

Deadline: Monday, March 21.