Series 4

1. Look at two time series `ts1` and `ts2` – possibly AR processes – which you can load using the following commands in R:

```
> d.ts <- read.table("http://stat.ethz.ch/Teaching/Datasets/WBL/ts_S3_A2.dat",
                     header = TRUE)
> ts1 <- ts(d.ts$ts1)
> ts2 <- ts(d.ts$ts2)
```

a) Plot both of these time series, and check whether they could have been generated by an AR process. Are these stationary time series? What is their mean?

b) Regard the autocorrelations and partial autocorrelations of the time series and decide whether they can be generated by an AR process. If yes, what is the order of the respective AR process?

2. Consider the AR(3) model with coefficients $\alpha_1 = 0.5$, $\alpha_2 = -0.4$ and $\alpha_3 = 0.6$:

$$X_t = 0.5 \cdot X_{t-1} - 0.4 \cdot X_{t-2} + 0.6 \cdot X_{t-3}$$

a) Simulate 1 realization of length 100 of this time series and plot it. Does the time series look stationary?

R hint:

```
> set.seed(3)
> ar.coef <- c(0.5, -0.4, 0.6)
> ts.sim <- arima.sim(list(ar = ar.coef), n = 100)
```

b) Compute whether or not this process is stationary by calculating the roots of the polynomial $\Phi(z) := 1 - \sum_{i=1}^{p} \alpha_i z^i$ with the R function `polyroot`.

c) Have a look at (partial) autocorrelations. Do they look as you expect? Comment. Also compare the estimated autocorrelations to the true ones (`armaACF()`).

d) Simulate realizations of this time series of different lengths (shorter and longer than 100) and compare each of their estimated autocorrelations with the real ones. What do you observe?

3. In this exercise we look at the yield of a chemical process. The relevant data from 70 successive experiments can be found in the dataset `yields.dat`. The aim of this exercise is to estimate the mean yield and construct a 95% confidence interval.

R hint: Load the dataset and create a time series as follows:

```
> d.yields <- read.table("http://stat.ethz.ch/Teaching/Datasets/WBL/yields.dat",
  header = FALSE)
> t.yields <- ts(d.yields[, 1])
```

a) Make a time series plot, estimate the mean yield and mark this in the plot.
R hint: Use `mean()` to estimate the mean yield. You can then draw a horizontal line with intercept a using the command `abline(h = a)`.

b) Investigate the dependence structure of this time series. Look at its autocorrelations. Compare with lagged scatterplots, and characterise the dependence structure.

R hints:

```
> acf(...)
> lag.plot(t.yields, lag = ..., layout = c(..., ...), do.lines = FALSE)
```

c) Construct a 95% confidence interval for μ by estimating each of the autocorrelations that differ from 0.
How large would this confidence interval be if independence were falsely assumed?
R hint: You can compute $\widehat{\gamma}(0)$ with either of the following commands:

```
> var(t.yields) * (length(t.yields) - 1) / length(t.yields)
> acf(t.yields, type = "covariance", plot = F)$acf[1]
```

d) Look at the partial autocorrelations. Would you use an AR model to fit this series? Which order would you take? Comment.

e) Use the Yule-Walker equations to estimate by hands the parameters $\alpha_1, \ldots, \alpha_p$ of the AR(p) model that you would use to fit the time series; p is the order you determined in Part d). Compute the estimate $\hat{\sigma}^2$ of the variance of the innovations $\mathrm{Var}(E_t)$. Check your results using R .
R hint:

```
> r.yw <- ar(yields, method = "yw", order.max = 1)
> str(r.yw)
```

Exercise hour: Monday, March 10.