Problem Sheet 3

1. Consider the MA(2) process, where all the e_t values are independent white noise with variance σ_e^2:

$$Y_t = e_t - 0.5e_{t-1} - 0.3e_{t-2}$$

   (a) Find the autocovariance function $\gamma_k = \mathrm{cov}(Y_t, Y_{t-k})$ and autocorrelation ρ_k.

2. Consider the AR(2) process, where all the e_t values are independent white noise with variance σ_e^2:

$$Y_t = \mu + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + e_t$$

   (a) Show that the expectation of the process Y_t,

$$E(Y_t) = \frac{\mu}{1 - \phi_1 - \phi_2}.$$

   (b) Define a new process $\tilde{y}_t$ that represents the deviation from the mean at time t, $\tilde{y}_t = y_t - \mu$ E[yt]. thus we get

$$\tilde{y}_t = \phi_1 \tilde{y}_{t-1} + \phi_2 \tilde{y}_{t-2} + \epsilon_t,$$

   For the process $\tilde{y}_t$, find recursive relationships between the autocovariances: γ_0, γ_1 and γ_2, and thus deduce their expressions in terms of σ_e^2, ϕ_1 and ϕ_2.

3. For each of the following ARIMA(p,d,q) models, what are the values of p,d and q. Furthermore, state whether the model is stationary and/or invertible. Below, $\{a_t\}$ denotes a sequence of white noise with zero mean and variance $\sigma^2 = 1$.

   (a) $Z_t = 5 + a_t + 1.5a_{t-1} + 0.5a_{t-2}$.

   (b) $Z_t - \frac{5}{6}Z_{t-1} + \frac{1}{6}Z_{t-2} = 1 + a_t + 2a_{t-1}$.

   (c) $Z_t - 0.5Z_{t-1} - 0.5Z_{t-2} = a_t + 0.5a_{t-1}$.

4. We described hypothesis testing procedures to select MA and AR models using sample autocorrelations and sample partial autocorrelations, respectively. Use these testing procedures in each of the following situations.

   (a) From a time series of n = 150 observations, we compute the following sample auto-correlation coefficients: $r_1 = 0.37$, $r_2 = 0.28$, $r_3 = 0.31$, $r_4 = 0.13$, $r_5 = 0.04$, $and\, r_6 = 0.15$. The remaining sample autocorrelations appear to be negligible. Find the MA(q) process most consistent with this information, that is, determine the value of q.

   (b) From a time series of n = 150 observations, we compute the following sample partial autocorrelation coefficients: $\hat{\phi}_{11} = 0.24$, $\hat{\phi}_{22} = 0.81$, $\hat{\phi}_{33} = 0.31$, $\hat{\phi}_{44} = 0.09$, $\hat{\phi}_{55} = 0.04$, $and\, \hat{\phi}_{66} = 0.02$. The remaining sample partial autocorrelations appear to be negligible. Find the AR(p) process most consistent with this information, that is,determine the value of p.

5. Generate three time series data sets, each of length $n = 200$, including (i) an AR(1) with $\phi = 0.6$, (ii) an MA(1) with $\theta = 0.8$, and (iii) an ARMA(1,1) with $\phi = 0.6$ and $\theta = 0.8$. For each one,

   (a) plot the observed time series.

   (b) plot the sample ACF, the sample PACF.

   (c) use the armasubsets function in R to identify the best model in terms of the BIC. Do the plots in part (b) agree with what you know to be true? Remember, you know the correct models! That is, you are assessing here whether the sample identification functions agree with the truth. Does the BIC identify the correct model as the best model in each case? If not, where is the correct model ranked, if at all?

6. For the following data sets:

   - ibm: daily closing IBM stock prices (dates not given).
   - internet: number of users logged on to an Internet server each minute.
   - robot: final horizontal position of an industrial robot put through a series of planned exercises which is in the TSA package.

   Identify a small set of candidate ARIMA(p, d, q) models for each data set. There may be a single model that emerges as a clear favorite.