STAT 479 - Assignment 2 - due date is on course outline

NOTE: On questions involving data analysis, such as #7 and #8, there is no one 'right' answer. Just do something sensible, and explain - clearly and in proper English - why it is sensible.

1. 3.2 in the text.
2. 3.3 in the text.
3. Study examples 3.11 and 3.12. Then consider the AR(2) model given by $X_t = -.9X_{t-2} + w_t$.
   (a) Determine the roots of the characteristic equation. Is the series stationary? Why or why not?
   (b) Mimic the example in the text to show that
   $$\rho(h) = \begin{cases} (-.9)^{h/2}, & \text{if } h \text{ is even,} \\ 0, & \text{if } h \text{ is odd.} \end{cases}$$
   (c) Check this by using the R code on page 70 of the text to plot the ACF.
4. Recall that, if we wish to find a prediction function $g(x)$ that minimizes $\text{MSE} = E\left[(Y - g(X))^2\right]$, then the minimizer is $g(x) = E[Y|x]$.
   (a) Show that $Y - E[Y|X]$ is uncorrelated with X, and hence (using the usual notation) that the prediction error $X_{t+l} - X_{t+l}^t$ is uncorrelated with the predictors X_s for $s \leq t$. (Recall that in linear regression the residuals are uncorrelated with the independent variables.)
   (b) Consider the model $Y = X^2 + Z$, when X and Z are i.i.d. $N(0,1)$ random variables. Evaluate the minimizing predictor and show that the minimum MSE is 1.
   (c) Suppose we restrict our choices for the function $g(x)$ to linear ones: $g(x) = a + bx$, and then choose the coefficients to minimize MSE. Show that then
   $$a = E[Y] - bE[X], \ b = \frac{\text{COV}[X,Y]}{\text{VAR}[X]}.$$
   (d) Apply (c) to (b), to get $g(x) = 1$, with a minimum MSE of 3. How do you interpret this result?

5. Let M_t represent the cardiovascular mortality series discussed in Chapter 2, Example 2.2, and called '`cmort`' in `astsa`.

   (a) Fit an AR(2) by linear regression. For this, first subtract the mean from the series, then use `dynlm` to fit a regression model without an intercept. (See the R code from Lecture 4 for an example - the command `fit = dynlm(x ~L(x,1) + L(x,2) -1))` will fit a model without an intercept.) Write out the estimated model for M_t.

   (b) Use sarima to fit an AR(2) model to M_t directly, i.e. without first subtracting the mean. Write out the estimated model for M_t; compare this, and the estimated variance of the white noise, with your results from part (a).

6. 3.10 in the text. [This will illustrate a remark made in class. The Gauss-Newton procedure was developed to handle cases in which w_t is a *nonlinear* function of the parameters, by making a sequence of linear approximations. In this AR(1) case the w_t are *linear* in the parameter, with the consequence - which you will find - that the final estimate is obtained in just one step.]

7. 3.15 in the text. This entire question can be carried out in R. [Note: Here is how the data seem to have arisen. Environmental conditions were measured daily, except for Sundays, for 10 years. These approximately 3000 records were then reduced to 508 weekly averages, with a 'week' being one of these 6 day periods. Thus four periods constitute about 1 month.]

8. 3.21 in the text. This entire question can be carried out in R.

9. For the FRB series discussed in class, obtain expressions (rather than numbers) for the 13-month ahead forecast and prediction interval. You should write things out using the notation $\alpha_1, ...\alpha_{39}$ adopted in class. Although you're not being asked to compute anything here, you should clearly explain what *would* be computed. Refer liberally to the results obtained in class for this model - there's not much left, after that.