1. (35pt) The Johnson and Johnson dataset contains quarterly earnings per share for the U.S. company Johnson & Johnson. There are 84 quarters (21 years) measured from the first quarter of 1960 to the last quarter of 1980. To load the dataset, run the following: install.packages("astsa"); library(astsa). The dataset is under the name jj. **Do a log transformation of the original time series before answering the following.**

(a) Construct a time series plot for the logged data. Comment on overall trend and seasonality variation.

(b) Fit the a regression model on the logged data

$$y_t = \beta_0 + \beta_1 t + \alpha_1 Q_2(t) + \alpha_2 Q_3(t) + \alpha_3 Q_4(t) + e_t,$$

where $Q_i(t) = 1$ if time t corresponds to quarter $i = 1, 2, 3$, and zero otherwise. Assume e_t is a normal white noise sequence. Report model coefficients estimates. Superimpose the fitted values on the time plot in part (a). **Note:** you will need to first create a variable for time and quarter. To do that, you may use: t=1:84; qt=as.factor(rep(1:4,21)).

(c) Calculate the MSE. Make a time plot, a ACF plot and a histogram for the residuals. Does the residuals look like a normal white noise process?

(d) Make predictions for the first quarter in 1981. Construct the 95% prediction interval.

(e) Fit a additive model using the Holt-Winters method. Let the function choose the optimal smoothing parameters automatically. Report the smoothing parameters and coefficients. Superimpose the fitted values on the time plot in part (a)

(f) Calculate the MSE for the Holt-Winters additive model and compare it with par (c). Make a time plot, a ACF plot and a histogram for the residuals. Does the residuals look like a normal white noise process?

(g) Make forecast using the Holt-Winters model for the first quarter in 1981. Also report the 95% prediction interval. Compare the result with part (d).

2. (15pt) Suppose that $\{e_t\}$ is a zero mean white noise process with variance σ^2. Let B denote the backshift operator. Consider the processes:

(i) $(1 + 0.4B)Y_t = e_t$
(ii) $(1 - 0.9B)(1 - B)Y_t = (1 - 0.5B)(1 + 0.4B)e_t$
(iii) $(1 - 0.4B - 0.45B^2)Y_t = (1 + B + 0.25B^2)e_t$

(a) Identify each model as an ARMA(p, q) process; that is, specify p, and q. (look out for parameter redundancy.)

(b) Give the autocorrelation function ρ_k for those processes which are stationary. If you want, you can use the ARMAacf function in R to see the first dozen or so correlations.

(c) Simulate a data set from each process identified above show the data set in a time series plot and also show the sample ACF for it. In each case, does the time plot agree with the stationarity? Does the sample ACF agree with what we know to be true from the theory? You may pick your favorite sample size (anything larger than 100) and the white noise variance.