

STAT 520 – Homework 2 (HWs 2A and 2B)

For Homework 2A, do Problem 2 and Problem 3(a)-3(f) below.

For Homework 2B, do Problem 1, Problem 3(g), and Problem 4 below.

[For these homework assignments, all students should do all the assigned problems.]

Note: For questions that ask you to create a plot, you should INCLUDE the plot within your submitted homework.

1) Consider the regression involving the cardiac mortality time series that we studied in class. Note that the cardiac mortality time series data are in the `cmort` object in the `astsa` package. Type `library(astsa); data(cmort); print(cmort)` in R to see the data set. Type `temp = tempr-mean(tempr) # center temperature`
`temp2 = temp^2`
`trend = time(cmort) # time`

to obtain the centered temperature time series, the squared centered temperature data, and the time data. Type `print(part)` to see the Pollution (particulate) time series.

In class, we considered the regression model with `cmort` as the response, and `temp`, `temp2`, `part`, and the time trend as predictors. Consider adding the particulate count four weeks prior, denoted P_{t-4} , as another predictor. That is, add the lag-4 particulate count to the model. The R code:

```
card.mort = ts.intersect(cmort, temp, temp2, trend, part, partL4=lag(part,-4))
```

may be useful. Write the equation of the estimated regression model you have fit. Based on AIC, is this model better than the model we fit in class? How about based on BIC? Clearly explain your answer.

2) The monthly values of the average hourly wages for U.S. apparel and textile workers for July 1981 to June 1987 are in the `wages` object in the `TSA` package. Type `library(TSA); data(wages); print(wages)` in R to see the data set.

(a) Plot the time series. What basic pattern do you see from the plot?

(b) Fit a linear time trend model using least squares. Give the plot of the linear trend overlain on the data, and give the estimated regression equation.

(c) Plot the standardized residuals from the linear regression over time. Comment on any notable pattern.

(d) Fit a quadratic time trend model using least squares. Give the plot of the quadratic trend overlain on the data, and give the estimated regression equation.

(e) Plot the standardized residuals from the quadratic regression over time. Comment on any notable pattern.

(f) Perform a runs test on the standardized residuals from the quadratic regression. What is your conclusion?

(g) Plot the autocorrelation function for the standardized residuals from the quadratic regression. What do you conclude about the standardized residuals?

(h) Investigate the normality of the standardized residuals (error terms) from the quadratic regression. What is your conclusion?

3) The monthly U.S. prescription costs per claim for August 1986 to March 1992 are in the `prescrip` object in the `TSA` package. Type `library(TSA); data(prescrip); print(prescrip)` in R to see the data set.

(a) Plot the time series, using plotting symbols that allow you to check for seasonality. What basic pattern do you see from the plot?

- (b) Fit a harmonic regression model using least squares, including one pair of harmonic functions AND a linear time trend as predictors. Give the plot of the harmonic regression overlain on the data, and give the estimated regression equation.
- (c) Plot the standardized residuals from the harmonic regression over time. Comment on any notable pattern.
- (d) Perform a runs test on the standardized residuals from the harmonic regression. What is your conclusion?
- (e) Plot the autocorrelation function for the standardized residuals from the harmonic regression. What do you conclude about the standardized residuals?
- (f) Investigate the normality of the standardized residuals (error terms) from the harmonic regression. What is your conclusion?
- (g) Fit a seasonal means model to the *detrended* prescription data. That is, first fit a simple linear time trend model to the prescription data, and then fit a (monthly) seasonal means model to the *residuals* of this linear trend model. Based on the AIC, which model is better, the harmonic regression from part (b) or the seasonal means model for the detrended data? What about based on BIC? Explain your answer.

4) Suppose $\{e_t\}$ is a normal white noise process with mean zero and variance σ_e^2 . Let $\{Y_t\}$ be a process with a linear trend over time, defined as:

$$Y_t = \beta_0 + \beta_1 t + e_t$$

- a) Explain why the time series $\{Y_t\}$ is not stationary.
- b) Consider the differenced time series $\{\nabla Y_t\}$, where $\nabla Y_t = Y_t - Y_{t-1}$. Show that the differenced time series $\{\nabla Y_t\}$ is stationary. [Hint: Note that β_0 , β_1 , and t are constants, not random variables, here.]