STAT 520 – Homework 2B – Fall 2023

[For this homework assignment, all students should do all the assigned problems.]

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:

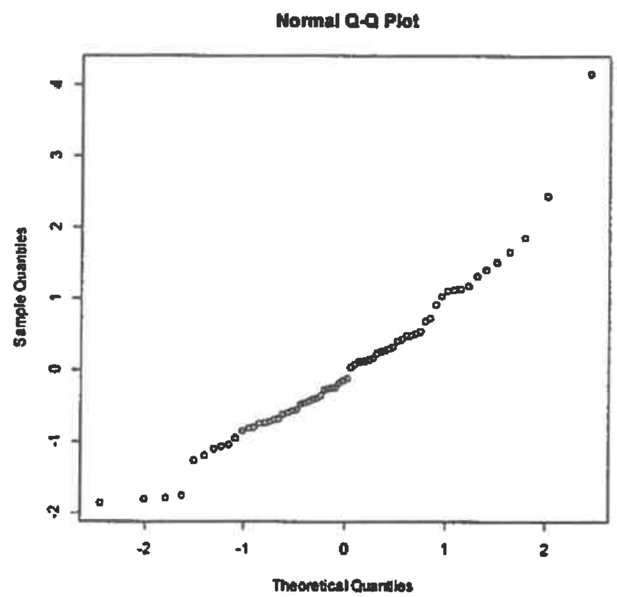
```
card.mort = ts.intersect(cmort, temp, temp2, trend, part, partL4=lag(part,-4))
```

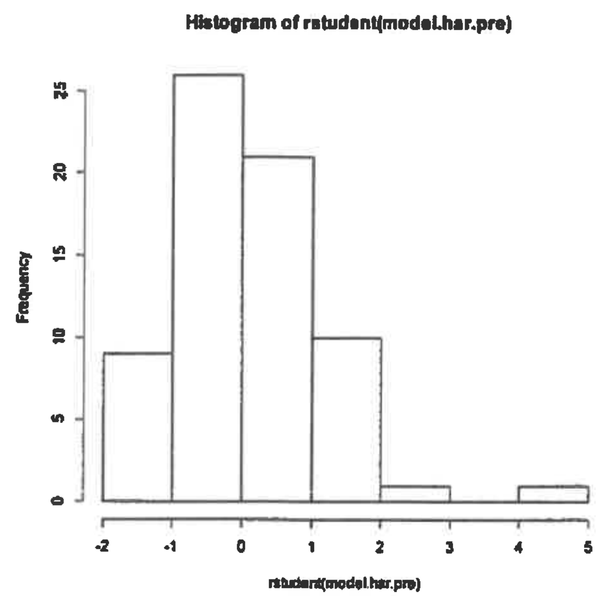
may be useful. Write the equation of the estimated regression model you have fit. Based on AIC, is this model better than the model we fit in class? How about based on BIC? Clearly explain your answer.

Fitted model equation:

$$m_t = 2808.33 - 0.4058079 \text{ Temp}_t + 0.0215466 \text{ Temp}_t^2 - 1.3853388\, t + 0.2028823 \text{ P}_t + 0.1030370 \text{ P}_{t-4}$$

The AIC of this model is 3291.52, which is lower (better) than the AIC of 3332.28 from the model we fit in class. The BIC of this model is 3321.08, which is lower (better) than the BIC of 3357.664 from the model we fit in class.





(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.

The harmonic regression model for (b) has AIC 127.29, compared to an AIC of 139.91 for the seasonal means model of the detrended data, so the harmonic regression model is judged as better based on AIC. The harmonic regression model for (b) has BIC 138.38, compared to an BIC of 168.77 for the seasonal means model of the detrended data, so the harmonic regression model is judged as better based on BIC.

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.

$E(Y_t) = \beta_0 + \beta_1 t$ which is NOT constant over time, so $\{Y_t\}$ 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.]

$E(\nabla Y_t) = E(Y_t) - E(Y_{t-1}) = \beta_0 + \beta_1 t - \beta_0 - \beta_1(t-1) = \beta_1$, which is constant with respect to t.

Find $cov(\nabla Y_t, \nabla Y_{t-k})$ for all k:

$\underline{k=0}$: $cov(\nabla Y_t, \nabla Y_t) = var(\nabla Y_t) = var(\beta_0 + \beta_1 t + e_t - \beta_0 - \beta_1 t + \beta_1 - e_{t-1}) = var(\beta_1 + e_t - e_{t-1}) = var(e_t) + var(e_{t-1}) = 2\sigma_e^2$

$\underline{k=1}$: $cov(\nabla Y_t, \nabla Y_{t-1}) = cov(\beta_0 + \beta_1 t + e_t - \beta_0 - \beta_1 t + \beta_1 - e_{t-1}, \beta_0 + \beta_1(t-1) + e_{t-1} - \beta_0 - \beta_1(t-2) - e_{t-2}) = cov(e_t - e_{t-1}, e_{t-1} - e_{t-2}) = 0 - 0 - \sigma_e^2 + 0 = -\sigma_e^2$

$\underline{k>1}$: $\text{cov}(\nabla Y_t, \nabla Y_{t-k}) = 0$, since there are no overlapping terms. So $\text{cov}(\nabla Y_t, \nabla Y_{t-k})$ depends only on the lag k, not the time t.
And since the process is normal, it is stationary.

$$\text{cov}(\nabla Y_t, \nabla Y_{t-k}) = \begin{cases} 2\sigma_e^2 & \text{if } k=0 \\ -\sigma_e^2 & \text{if } k=1 \\ 0 & \text{if } k>1 \end{cases}$$