Final Exam Time Series Fall 2019 (75 pts):

True or False (2 pts each):

1. If a spectral density has a peak at zero then it must come from a non-stationary process.

T

1. If a spectral density has a strong peak between 0 and .5 then the data must come from a non-stationary process.

F

1. Burg estimation of parameters of models of the airline data will always yield a stationary model.

T

1. If the Ljung-Box test fails to reject the null hypothesis, then we know for sure that that set of data is white noise.

F (evidence suggests its white noise…technicality? Can never know for sure)

1. A problem with the Dickey-Fuller test is that it will indicate that there is a (1-B) in the model when the data is actually from a stationary process more often than it should.

T

1. The Cochrane-Orcutt procedure is helpful in adjusting a linear model for serial correlation.

T

1. A model must yield white noise residuals to be useful.

T

1. The model Xt = .7at-1 + at is already in GLP form.

T

1. We know that an invertible model is desirable because, it ensures that the “present is related to the past in a sensible manner.” Describe briefly what is meant by this statement. (3pts)

When a model is invertible, all the roots are outside the unit circle. In other words, another model won’t be able to reproduce the autocorrelation.

1. Consider the realization below and highlight the best answer. (3 pts)



1. This data clearly comes from a stationary process.
2. This data clearly comes from a non-stationary process.
3. We cannot tell from a single realization

Unclear if realization is population or sample. While realization may be stationary/non-stationary, the population may be otherwise. In this case it appears to be non-stationary, but we just can’t be certain.

1. Consider the realization below and highlight the best answer. (3 pts)



1. This data clearly comes from a stationary process.
2. This data clearly comes from a non-stationary process.
3. We cannot tell from a single realization.

As mentioned above, the realization in this case looks to be stationary. However, we can’t be certain if the population will behave the same. If it does, then it would be safe to assume it is stationary.

1. You have been hired by a sleep scientist to study the cycles of REM sleep. REM (Rapid Eye Movement) is the part of sleep in which people are thought to dream. The scientist’s hypothesis is that on a certain drug, these REM cycles should happen fairly regularly. In order to test this hypothesis, she measures the muscle activity in 5-minute intervals (observations are taken 5 minutes apart) and records the number of eye movements in those 5 minutes. Below is the spectral density of the resulting data set. The scientist would like to know if there is any evidence of consistent periodic behavior of the REM cycle and if so, what the period is thought to be. Write a 1 to 3 sentence response addressing if there is evidence of periodic behavior of the REM cycle and, if so, provide an estimate of the time amount of time (in minutes) between each REM cycle. (4 pts)



It appears the period is 50 minutes (10 X 5-minute intervals) based on the output provided. The frequency is at 0.1, which would support the period to be 50 minutes. 1/0.1 = 10

1. Below are the monthly retail debit card usage in Iceland from January 2000 - August 2013 and four different forecasts of the next four years. Match the forecast with the corresponding model that produced those forecasts. (2 pts each)

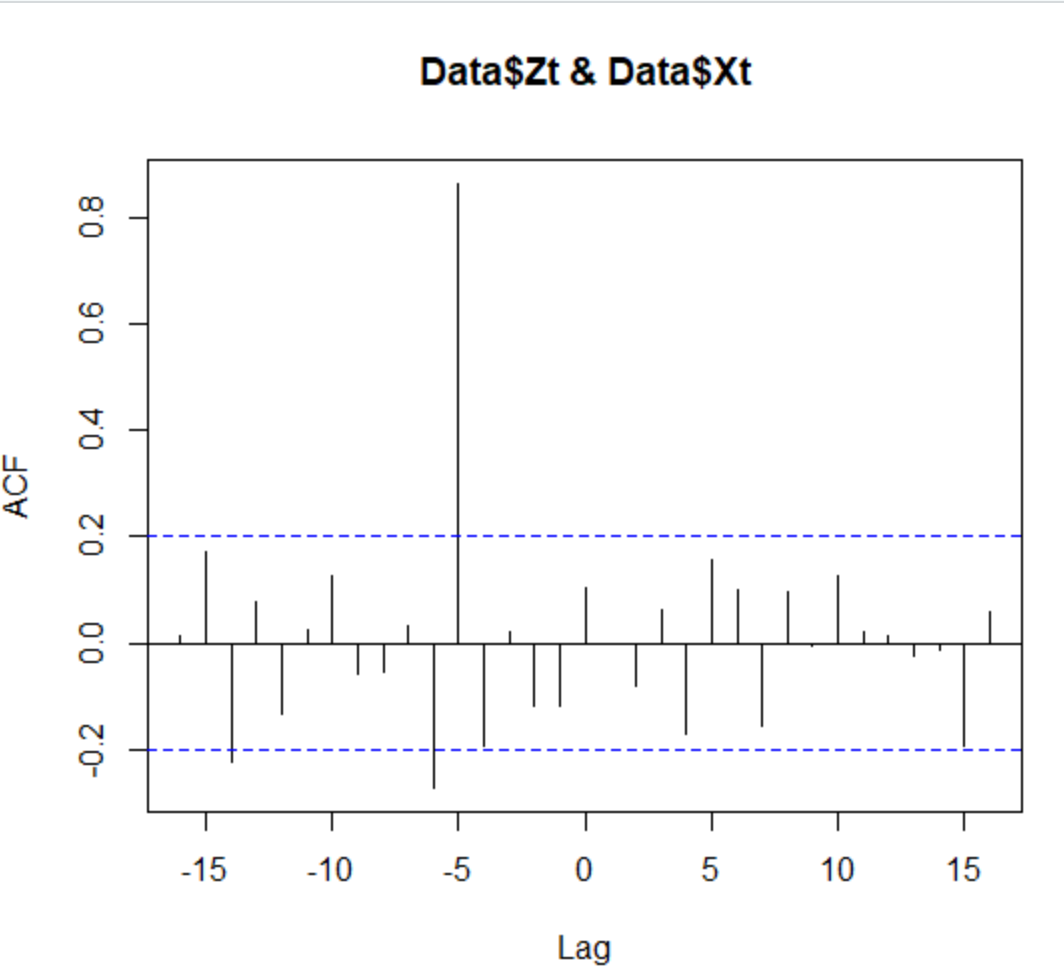
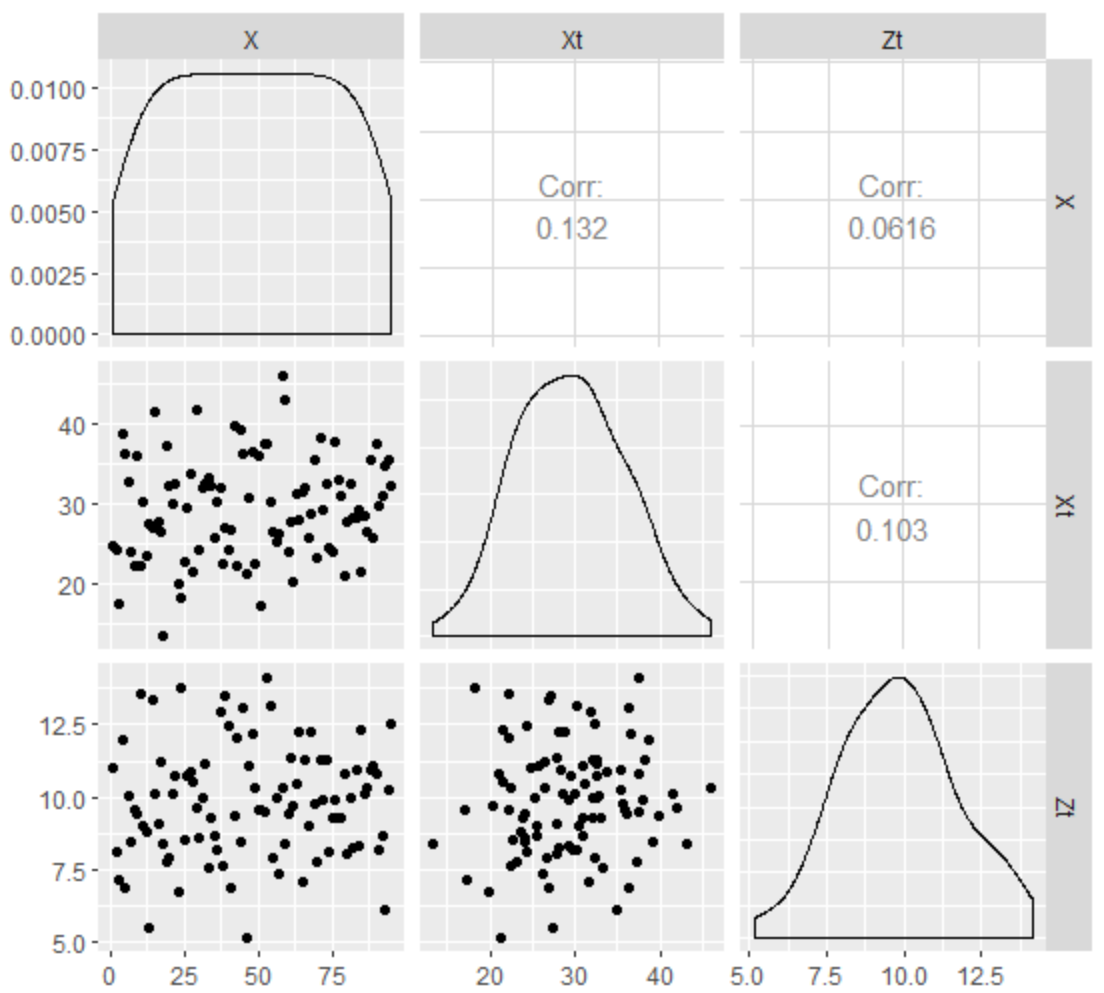
|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |

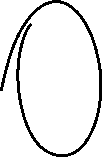
\_D\_\_ \_C\_\_ \_\_F\_ \_\_A\_

1. ~~Airline Model~~
2. φ(Β)(1-B)Xt = θ(Β)at
3. ~~ΑR(2) Complex Roots~~
4. ~~AR(2) Real Roots~~
5. φ(Β)(1-Bs)Xt = θ(Β)at
6. ~~(1-B)Xt = at~~
7. Match the ACF on the left with the corresponding spectral density or realization on the right. Simply place the letter next to the corresponding number below the plots. (2 pts each)

|  |  |  |
| --- | --- | --- |
| 1. |  |  |
| 2. | b. |
| 3. | c. |
| 4. |  | d. |
| 5. |  | e. |

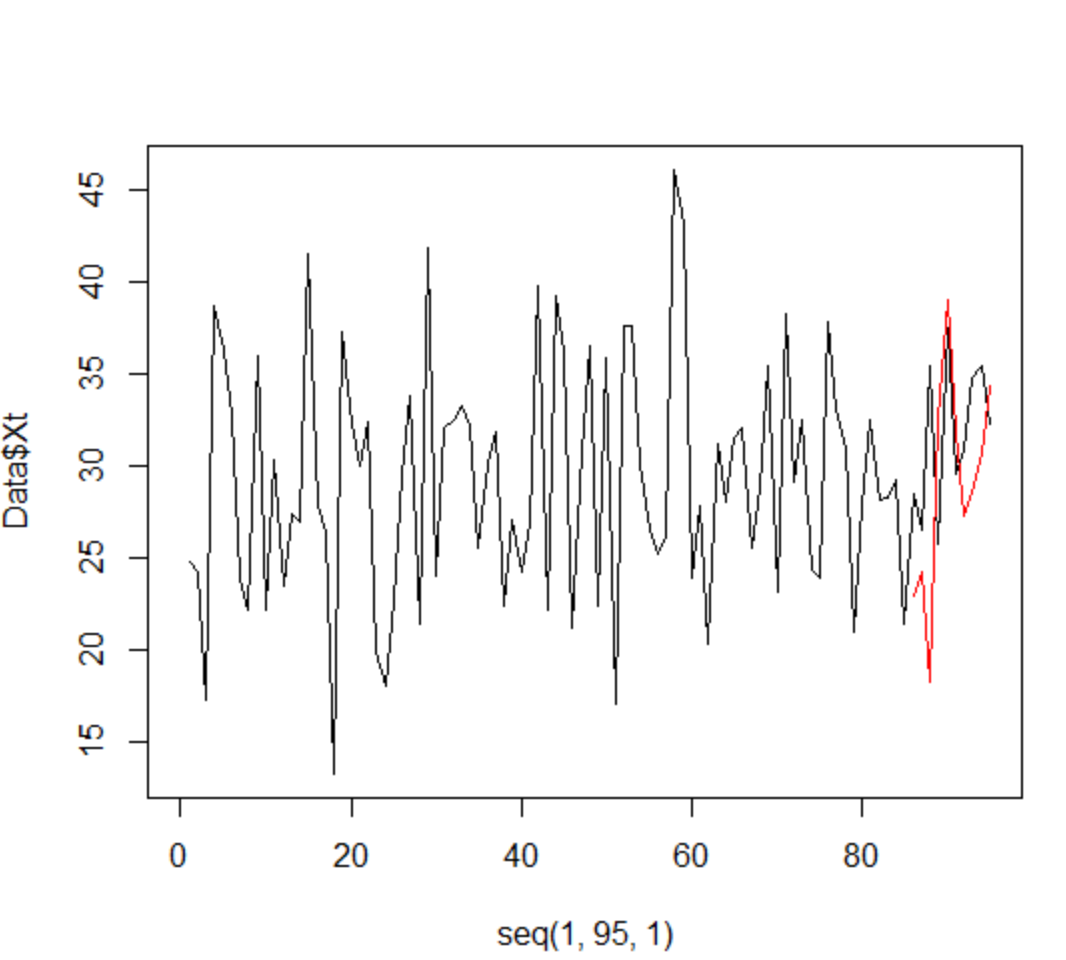
1. \_E\_\_ 2.\_B\_\_ 3. \_A\_\_ 4. \_\_D\_ 5. \_C\_\_\_
2. Consider the data in the file: FinalExamData.csv. This file contains a column Xt and a column Zt. Your goal is to simply model this data the using a vector autoregressive model (VAR) and a multi-layer perceptron model (MLP). You ultimately want to forecast Xt with a horizon of 10 using Zt if it is useful. Provide the following information in your response:
   1. Identify the relationship/association between Xt and Zt. Specifically, is there evidence of a relationship/association between Xt and a lagged Zt? … if so, what is the lag?) What evidence do you have to support this relationship/association? (4 pts)





Based on the output from the ccf() function, it appears there is a significant correlation at lag -5 for Zt.

* 1. For each model (VAR / MLP) (10 pts each model):
     1. Provide the code you used to fit the model and make the predictions.
     2. Provide a plot of the 10 forecasts (time points 96 – 105). You do not need prediction intervals for this question.
     3. Find the ASE for each model using the last 10 observations of the dataset. Include your code for this as well.



* 1. Make a quick statement about which model you feel is most useful and why. (4 pts)

library(vars)

library(tswge)

library(GGally)

library(dplyr)

library(nnfor)

plot(Data$Xt, type = 'l')

plot(Data$Zt, type = 'l')

ggpairs(Data)

ccf(Data$Zt,Data$Xt)

# Lag -5 for Zt

Zt\_5 = dplyr::lag(Data$Zt,5)

#VARS

Var1 = VAR(cbind(Data$Xt,Data$Zt5), type = "both", lag.max = 10)

AIC(Var1)

preds = predict(Var1,n.ahead = 10)

ASE = mean((Data$Xt[86:95] - preds$fcst$y1[,1])^2)

ASE

plot(seq(1,95,1), Data$Xt, type = "l",xlim = c(0,95))

lines(seq(86,95,1),preds$fcst$y1[,1], type = "l", col = "red")

#MLP

Xt85 = ts(Data$Xt[1:85])

Zt1 = data.frame(ts(Data$Zt))

fit = mlp(Xt85,xreg = Zt1)

f = forecast(fit, h = 10, xreg = Zt1)

plot(Data$Xt[1:85],type = "l")

lines(seq(86,95),f$mean, col = "blue")

ASE = mean((Data$Xt[86:95]-f$mean)^2)

ASE

Considering the short on time, it looks like the MLP model performs better. It produced a lower ASE then the other model. However, I was not able to prodict the 10 future values.