Class: <u>STAT</u>_____   Name _____   ID _____

Due: 11:59 PM Tuesday, Dec 8th 2020. Submit your work to Dropbox.

Instructions

Please do the final exam questions all by yourself. You can use the textbook, lecture note, previous homework, but you are not allowed to communicate with anyone else. If you do not understand the questions, please feel free to ask the instructor. Do not share any part of your answer with others. If plagiarism is detected or suspected, the instructor will have to report it to the department based on the Academic Integrity policy. Academic Dishonesty is strictly proscribed and if found may result in student discipline up to and including dismissal from the university. The grade for this test will be counted as zero. All individuals involved in plagiarism will get the same penalty.

Include the questions in your answer sheet. Arrange your answer in the same order as the questions. Mark the question number clearly before the answer. Answer the questions one by one and separate questions by two lines line spacing.

Interpret the outputs and results in plain language. No interpretation, no credits.

For R users, please submit both the .RMD file and the **.pdf** file that knitted by the .RMD file. For JMP users, please include screenshots of steps, outputs, graphs, and anything that are necessary to interpret your answer in the **.pdf** file. For all students, submit the final data set that you used for running R or JMP. Especially if you operated the original data through Excel before importing them to R or JMP.

Here is a letter written by Dr. Andrew Heiss that I want to share with you. "… Your grade on an exam says nothing about your value as a person. Exam grades are imperfect measures of what you have actually learned. You have worked hard so far this semester, and you have learned a lot, regardless of what your score might say. Grad School is hard, and just the fact that you are here in one of top programs in the country is so impressive! PLUS we're in the middle of a global pandemic, there's an election in less than one month….Just try your best…"

The data set Final_data presents data on the occurrence of violent crimes in the US from 1984 till 2005.

1. (5 points) Plot the crime rate data vs the year.
2. (10 points) Calculate and plot the sample autocorrelation function (ACF) and variogram. List the first 10 values of ACF and variogram respectively.
3. (5 points) Is there an indication of nonstationary behavior in the time series? Why or why not?
4. (10 points) Calculate and plot the first difference of the time series. Show the first 10 differences.
5. (10 points) Compute the sample autocorrelation function (ACF) and variogram of the first differences.
6. (5 points) What impact has differencing had on the time series?
7. Develop an appropriate exponential smoothing forecasting procedure for the first-differencing data by answer the questions below.
    a. (10 points) Assume the first-difference data is a constant process. For R user, use the HoltWinters() function to find the optimum value of λ to smooth the data. For JMP user, specify the λ given by the software.
    b. (10 points) Show the fitted values and corresponding SSE by using the λ obtained in part a.
    c. (5 points) Plot the fitted values and original values in a same plot.
    d. (5 points) Assume the first-difference data shows a trend. Calculate the SSE. You can get it from the HoltWinters() function. Then compare the SSE with that of obtained in part b. What can you tell from the comparison?
    e. (5 points) Suppose the first-difference is a constant process. Give the forecasts of the crime rate for years from 2006 to 2010.
8. a. (10 points) Develop an appropriate ARIMA model and a procedure for forecasting for the crime rate data. Specify the model and estimated parameters in the model. Hint: You can use the auto.arima() and forecast() functions to answer this question.
    b. (5 points) Compare the AIC obtained from part a with that of obtained from ARIMA(0,1,0) model. Which model has a smaller AIC? What can you tell by this comparison?
    c. (5 points) Show the 1- to 5- step ahead forecasts and corresponding 95% prediction intervals for the crime rate. Show only the results/outputs. Calculation process or formula are not required.