Stat 510 Take Home Exam

Due Date

Sunday February 19, 2017, by midnight U.S. Eastern Standard Time.

Scoring

There are two data analysis questions and three other questions. Each data analysis question counts 40%. The other questions count a total of 20%.

Instructions

Please submit a MS Word document or PDF file to the designated Drop Box in the Drop Boxes folder. Collaboration with other students is NOT allowed. Please do NOT discuss exam questions on the message board.

Data Analysis Questions

Tasks

For each of two datasets, the general task is to develop an ARIMA model for the data and then, for each model, write a brief discussion that includes the following:

1. A discussion of the initial model identification process. This should include a time series plot of the raw data with a discussion of important features, a discussion of any necessary differencing or transformations, and a presentation and interpretation of all relevant ACFs and PACFs used to identify an ARIMA model.

2. The final model, including output for this model only. Provide evidence that the terms in your final model are significant.

3. If necessary, a brief (one or two paragraphs) discussion of any iterations you may have made to get from the end of part 1 to your answer for part 2. By this, I mean that sometimes your first guess at a model doesn't work, so you try a few other possibilities before arriving at the end. If that happened, briefly describe the process. (DO NOT provide all of the output).

4. Explain and provide evidence that the errors (residuals) for your model are well-behaved.

Datasets

Analysis 1: Use the e1x1.dat dataset in the Dataset folders. It gives n = 150 simulated monthly data points.

Analysis 2: Use the e1x2.dat dataset in the Dataset folders. It gives n = 180 simulated annual data points.

Other Questions

1. Evaluate each of the following expressions and determine whether the model is a MA model, an AR model, or an ARMA model.

   (a) $x_t = \mu + .79Bx_t + .64Bw_t + w_t$

   (b) $x_t = \mu + (1 - .67B + .53B^2)Bx_t + w_t$

2. Consider the AR(1) model $x_t = 2 + .58x_{t-1} + w_t$, where $w_t \sim N(0, \sigma_w^2)$.

   (a) Give a numerical value for the first and second lag autocorrelations.

   (b) Simulate 200 observations of an ARMA$(1, 0)$ model with $\phi = .58$. Here is how you can simulate this data in R:

   x=arima.sim(list(ar = .58), 200)

   (1) Provide the ACF and PACF of your simulation.

   (2) Comment on how the ACF and PACF in part 1 either supports or does not support the true model that you simulated.

   (c) Now assume that this is an estimated AR model for n=200 observations with $\hat{\sigma}_w^2 = .7$ and $x_{200} = 1.22$. Psi-weights, ψ_1 to ψ_5, as given by R are:

   0.58000000 0.33640000 0.19511200 0.11316496 0.06563568

   Forecast the value of x_{201} and determine a 95% prediction interval for x_{201}.