Stat 470/670 Study Guide

The exam will be in class on Thursday 11/15. There will be about 10 short-answer questions, and the topics will be taken from the following list:

- Transformations: What transformations would you try to straighten out a curved relationship? If your data are skewed, what transformations are likely to reduce the skewness?
- Types of plots you should be able to read and know how to use: empirical CDF, QQ (both Q-normal and Q-Q for comparing two samples/two empirical distributions), boxplots. Given these plots, you should be able to tell things like which one has heavier tails, which one has larger variance, what indicates multiple modes, and so on.
- Reading and using coplots for modeling: If you were interested in an interaction between two variables, how would you construct a coplot to investigate (what variables would go on the axes, what variable would be the given variable)? If you are given a coplot, you should be able to tell me whether you think there is an interaction between the variables and what about the coplot indicates the interaction.
- Reading and interpreting PCA and LDA biplots.
- Using smoothers in support of modeling and visualization: What does LOESS do, when is it better than standard regression? When is robust regression helpful or necessary?
- Using residual plots for model diagnosis: What plots would you make to diagnose heteroskedasticity? Lack of model fit?

Finally, if I describe a set of variables to you and tell you what I am interested in learning about them, you should be able to describe in a couple of sentences what plots or modeling techniques you would use to answer my questions.

For example, if I say that I have two samples and I want to check whether they seem to have the same distribution, you should be able to tell me that I should make a QQ plot and that I should check that the points on the QQ plot fall on a line with slope 1 and intercept 0.