Statistics 110/201 Midterm Review, Fall 2017

The midterm exam will cover Chs 0, 1, 2, and some of 3, including all material covered in class up to Fri, Oct 27. You are expected to know how to read and interpret R commands and output from R, but you don't need to know thow to write exact R code. The exam is open notes, and bring a (basic) calculator.

Here is what you should know how to do and/or interpret:

1. Understand the difference between a deterministic (functional) relationship and a statistical relationship.
2. Understand the steps in hypothesis testing, and what they mean. For instance, be able to interpret a *p*-value and understand the various conclusions that can be made for a hypothesis test based on the *p*-value.
3. Understand how to interpret a confidence interval, a prediction interval, and the confidence (or prediction) *level* that accompanies them.
4. Understand when it can be concluded that a change in X *causes* a change in Y, and when it can only be concluded that there is a relationship (experiments versus observational studies; also see Chapter 0 of the textbook and the 1st lecture).
5. Write the *population model* for regression using correct notation, including the assumptions that go with it.
6. Interpret each term in the population model (slope, intercept, error, population standard deviation). Be able to do this for multiple regression as well as simple linear regression.
7. Understand the concept of least squares estimation.
8. Understand why you shouldn't extrapolate after getting a regression equation.
9. Write the *sample equation* in notation, and numerically once you have computer output.
10. Know what sample values are used to estimate each of the terms in the population model, including coefficients and standard deviation.
11. Use the sample equation to find a predicted value and a residual for an individual in the sample.
12. Interpret everything we've covered in the computer output given by R.
13. Understand how to check for outliers, lack of normality and non-constant variance, including what plots should be examined.
14. Understand what causes outliers and what to do about them.
15. Understand how transformations can be used to correct deficiencies in the assumptions, including when to transform Y instead of X and vice versa, and what to do after using a transformation to get results in the original units.
16. Using the computer output, test individual coefficients (using a t-test), get confidence intervals for them, and interpret the results.
17. Understand the relationship between the t-test and the F-test for slope, and the t-test for correlation in the simple linear regression situation.
18. Understand the difference in what is being tested with the t-test and F-test for slope, and the t-test for correlation in the multiple regression situation.
19. Construct an ANOVA table, including how the pieces fit together (e.g. SSModel + SSE = SSTotal, and so on).
20. Understand how to estimate and predict for a new case with values x^*, including how to interpret a confidence interval for μ_Y and a prediction interval for an individual.
21. Find and interpret R^2 and know why we would sometimes use R^2-adjusted instead.
22. Understand how to compare the intercepts of two regression lines using indicator variables.
23. Understand how to compare the slopes of two regression lines using indicator variables.
24. Understand what the sequential sums of square in the ANOVA table represent.
25. Understand how to use the nested F-test to simultaneously compare intercepts and slopes.