Math 58B - Final Preparation – Spring 2018

Logistics:

• Wednesday, May 9, 9am

• Take home problem due at the final

• Two-sides of notes are allowed (**one** piece of paper)

• Bring calculator, no computer use in class

• Material from the entire semester, cumulative

• New material: Inv 5.1 – 5.14

Overview (new material):

We have analyzed three types of studies:

1. those that involve two categorical variables; the results can be organized in a rxc table (chi-square tests)

2. those that involve a categorical (2 or more groups) explanatory variable and a quantitative response variable. We looked at boxplots of the quantitative variable broken down by the explanatory variable. The boxplots allowed us to see how the between variability (between group means) compared to the within variability (within the group). We call the related analysis ANOVA.

3. those where the explanatory and response variables are both quantitative. We looked at scatterplots and calculated correlation, intercept, and slope of the least squares regression line. We were able to do inference on the slope, primarily to test the null hypothesis that the population slope is zero.

We have studied how to conduct inferences depending on whether the data were collected from:

• Independent random samples

• One random sample with two variables

• Randomized experiment

• Observational study

We have also considered how the scope of conclusions to be drawn depends on how the data were collected. More specifically:

• Random assignment allows for the possibility of drawing cause/effect conclusions.

• Random sampling allows for generalizing to a larger population.

Outline (since last exam):

• Randomiz. test for rxc tables (hypotheses on homogeneity of proportions/association)

• Chi-square test for rxc tables (hypotheses on homogeneity of proportions/association)

• statements of hypothesis for ANOVA

• the intuition of the F test statistic for ANOVA

• technical conditions for ANOVA

• the factors influencing the power of the ANOVA test

• interpretation correlation, interpretation of R2

• the process used to calculate b0 and b1 , interpreting both in context

• dangers of extrapolation

• sampling distribution of b1, factors influencing SE(b1)

• LINE assumptions (what they are and why we need them)

• reasons for transforming data

• (we did not cover prediction intervals)

Advice:

• Organize notes for efficient retrieval of information/formulas

• Don’t plan to use notes too much

* Prepare as if exam were closed book/notes
* Focus on understanding, not memorization
* Be cognizant of time constraint

• Expect similar questions to what we answer in class every day, clicker Q, on HW, on previous exams

• Be ready to interpret computer output

* Possibly exclude irrelevant output

• Be prepared to think/explain/interpret

* Do not just plug into formulas
* Be ready to explain process of how you would do calculations

• Read carefully

* Be sure to answer the question asked

• Take advantage of information provided

* Perhaps including output

• Relate conclusions to context

• Practice

* Work through solved examples at end of text chapters
* Re-work in-class investigations, examples
* Re-work HW questions
* Re-work clicker questions (posted online)
* Re-work previous exam questions
* Re-read summary sections

• Show up on time!

* Be cognizant of time constraint
* Make an attempt of all questions