

Final Exam - Stat 313

March 11, 2021

These are the exam questions given to students for their oral midterm exams. Students answered one randomly selected question from each bank (Statistical Inference & Linear Models), and were informed of their assigned questions when they entered the “exam room.”

The numbered questions are what were given to students the Friday before exam week. The *follow-up questions* are a sample of the questions I asked students following their explanation. These questions are not exhaustive, but provide an outline of the types of questions I asked students.

Statistical Inference - Question Bank

1. Explain to someone who has never taken a statistics course what a p-value is.
2. Explain to someone who has never taken a statistics course what a confidence interval is.
3. It is very likely researchers in your field are familiar with parametric tests (e.g. t -test, F -test) but are not familiar with simulation based methods (e.g. randomization test, bootstrap confidence interval). Explain to someone in your field the similarities and differences between these two methods.
4. Large portions of the scientific community are pushing to eliminate p-value thresholds for publication of scientific research. What is the role of a p-value threshold (e.g. 0.05) in hypothesis testing? What are the benefits and costs of eliminating p-value thresholds?
5. What is a sampling distribution? How are randomization and bootstrap distributions used to approximate a sampling distribution?

Linear Models - Question Bank

1. Suppose a researcher performs a one-way ANOVA and finds strong evidence to reject the null. They are interested in reporting which groups are “significantly” different from each other and are planning to perform tests for a difference in means for all of the groups that visually appear to be different. Explain to this individual the issues with this analysis strategy and propose an alternative strategy they should take.

2. When the conditions for a linear regression are violated, how are the associated p-value and confidence interval impacted? How do these impacts differ for simulation based methods versus parametric methods?
3. A one-way ANOVA compares the means of three or more groups, and summarizes these differences with an F-statistic. What are the components of an F-statistic and what does each term symbolize in terms of the variability of the responses?
4. What are the similarities and differences between a two-way ANOVA model and a multiple linear regression? What does it mean to have “parallel lines” in a two-way ANOVA? What does it mean to have “different slopes” in a two-way ANOVA?
5. Suppose an experiment is designed to use blocking to account for a possible confounding variable. The analysis of the experiment’s results yields the following ANOVA table:

term	statistic	p-value
treatment	4.25	0.035
block	3.5	0.015
treatment:block	2.5	0.01

Are there issues if there is evidence of an interaction between the blocking variable and the treatment? Why or why not?