Final Exam - Stat 313

November 24, 2020

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.

Possible follow-up questions:

- A p-value is often reported as a percentage. What does that percentage p-value mean?
- How is a p-value calculated?
- What does a small / large p-value suggest?

2. Explain to someone who has never taken a statistics course what a confidence interval is.

Possible follow-up questions:

- What value does a confidence interval hope to capture?
- What does it mean to say you're "95% confident"?
- How is a confidence interval calculated?

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.

Possible follow-up questions:

- How does a t-test calculate a p-value / confidence interval?
- Why would a researcher choose to use one method over the other?
- How are the data "simulated" to obtain a permutation distribution?
- What are similarities and differences between a permutation distribution and a t-distribution?

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?

Possible follow-up questions:

- What does a p-value of 0.05 mean?
- Does a p-value less than 0.05 demonstrate that the findings of a study are "significant"?
- What could replace a p-value threshold in determining the "significance" of a study?

5. In the last lab assignment (Inference for Linear Regression), you used a bootstrap distribution to estimate a plausible range of values for the population slope. What is the key assumption behind a bootstrap distribution? Why was it not necessary to create a bootstrap distribution to estimate a range of values for the population slope?

Possible follow-up questions:

- What were the observations in the data from the lab assignment?
- What value was recorded in each row of the data?
- What population were we interested in?

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 appear to be different. Explain to this individual the issues with this analysis strategy and propose an alternative strategy they should take.

Possible follow-up questions:

- What issue is there will performing lots of tests for a difference in means?
- How does the post-hoc test control the Type I error rate?
- How do you determine if a difference in means is "significant" with a post-hoc test?

2. When the conditions for a linear regression are violated, how is the associated p-value and confidence interval impacted? How do these impacts differ for simulation based methods versus parametric methods?

Possible follow-up questions:

- What are the conditions of a linear regression?
- If these conditions are violated, what happens to the p-value? The confidence interval?
- What conditions do simulation based methods have? How does this differ from a t-test?
- In the lab, we saw that with violated conditions we obtained contradictory findings from a randomization distribution and a bootstrap distribution. Would this same thing have happened if we had used a t-test instead?

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?

Possible follow-up questions:

- What does the mean square of the groups represent? How is it calculated?
- What does the mean square of the errors represent? How is it calculated?
- If I were to obtain a large F-statistic, what does that indicate about the mean square of our groups relative to the mean square of the error?
- What would you expect to see in a data visualization for a one-way ANOVA with a large F-statistic?

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?

Possible follow-up questions:

- How are the "slopes" in a two-way ANOVA model formed?

- What do parallel slopes / different slopes indicate about the relationship between the explanatory variables and the response?
- How does a two-way ANOVA model summarize the relationships between variables? How does this differ from a multiple linear regression?

5. Prior to this year, it was fairly common for medical studies to have a lack of representation of women and people of color (and women of color). Suppose a researcher was interested in assessing the relationship between race and age on the metabolic rates of women. The researcher categorized race and age as follows:

- Black, Hispanic, American Indian or Alaska Native, Asian, Native Hawaiian or Pacific Islander, White
- 20 - 30, 30 - 40, 40 - 50, 50 - 60, 60 - 70, 70 - 80

The researcher recruited one woman from each intersecting group (e.g. Black and 30-40). When testing for an interaction between race and age on metabolic rates, the researcher obtained the following degrees of freedom for their test. What is the issue? What can the researcher do to assess the relationship between these variables and the metabolic rate?

term	degrees of freedom
race	5
age	5
race:age	25
error	0

Possible follow-up questions:

- Why does the ANOVA table have 0 degrees of freedom in the error?
- What issues does this cause for the hypothesis tests?
- What are ways the researcher could add degrees of freedom in the error?
- If the researcher could not add subjects in their study, what type of statistical model could they fit?