Final Exam Study Guide 209

Vinny Paris

May 2026

The guide is broken into two parts. The “new” stuff and the “old” stuff which reflects the exam’s makeup. To be upfront there will be a test of a regression coefficient and I do want you to be able to interpret an ANOVA table. It is highly advisable to review the notes for mulitple linear regression (included in the outline) and the definitions from experiments as the material builds on itself.

1 P-values and Regression Errors Generally

- A large p-value is not considered a bad thing. It suggests that the evidence against the null hypothesis is lacking.

- Know what an *error* in hypothesis testing it
  - You did NOT do something wrong
  - Your sample was unlucky
  - So you came to the wrong conclusion
  - Type 1 error: Incorrectly claim strong evidence against the null hypothesis even though the null is true
  - Type 2 error: Incorrectly claim you have little/no evidence against the null even though the null is false

- The p-value is uniformly distributed between 0-1
  - So probability of getting a p-value less than .1 (weak evidence) if the null hypothesis is true is 10%
  - Similarly .05 (moderate evidence) would have a 5% probability of randomly happening if the null is true

- Effect size vs p-value

- An effect size is how strong a change in the response we expect if we change our explanatory variable
  * Increase it by one unit for numeric variables
  * Go from one group to another for categorical data
  * Often described as β in the regression equation
- A p-value is a statement about detecting statsitical differences
- Effect size and p-values are not directly related
  * A small p-value can be associated with a meaningless difference between the groups (in real life)
  * Or a large p-value could be associated with a very large $\hat{\beta}$
- p-values do NOT talk about importance or influence, those are effect sizes
- Be prepared to comment on whether a large p-value with a large effect is more meaningful than a small p-value with a small effect
  * There is no "correct" answer but I want to make you stretch your thinking
  * And to rationally discuss a position

- Multiple Comparisons and p-hacking
  - What is the multiple comparisons problem?
    * Running multiple tests is more likely to find a small p-value by accident/H_0 is true
    * Raises doubts about how reliable your results are
    * The ASA statement on p-values mentioned the number of hypothesis checked because it ties into multiple comparisons
    * Several ways to deal with this (you don't need to know them)
  - What is p-hacking? Why is it bad?
    * Forcing your model to get small p-values
    * Usually violating assumptions while you do it
    * Or you ignore the multiple comparisons problem
    * Fundamentally violates the point and ideas of science which is to move the football of human knowledge forward and instead is usually seen as a self-serving career move
      · Or it's because they are ignorant of the subject matter which is still worrisome
      · SIDE NOTE: Hanlon's Razor (wiki) states that "never attribute to malice that which is adequately explained by stupidity" and is a personal favorite
    * Originates out of "publish or perish" mentality
    * Strength of evidence in hypothesis testing is (hopefully) going to alleviate this some

2 Testing Regression Coefficients

- Explain why we test coefficients at all
  - What does it imply if our p-value is large?
  - What does it imply if our p-value is small?
- Be able to write out the null and alternative hypothesis
  - Almost always it's H_0: $\beta_{Some\ number} = 0$
- Check the assumptions
  - Same as in the MLR section below
- Interpret R's output on testing regression coefficients.
  - What does a low p-value mean?
  - What does a p-value mean generally?
- Explain when testing regression coefficients can be better than a standard t-test
  - Lurking variables aren't accounted for in t-tests if they exist
- Interpret confidence intervals for these as well.

3 Multiple Linear Regression (MLR)

- Explain why we use MLR over simple linear regression
  - SLR only allows for one explanatory variable which is too limiting
  - MLR allows us to better understand the system as a whole since we can account for differing effects
- How do we interpret the coefficient of a quantitative (numeric) explanatory variable?
  - When (explanatory variable) increases by 1 unit we expect the mean of (response variable) to increase by $\hat{\beta}_{some\ number}$ holding the other variables constant
- How do we interpret the coefficent of a indicator (eg a nominal/categorical explanatory variable)?
  - When going from (BASELINE CATEGORY) to (INDICATOR'S CATEGORY) we expect the mean of (response variable) to increase by $\hat{\beta}_{some\ number}$ holding the other variables constant
- Make a prediction uisng a MLR

- Assumptions
  - Random
  - Population is normal or n is large
    - The ”population” here is the residuals!
  - IID
    - There is where homoskedasticity assumption is (need same spread to be identically distributed)
    - Also where “linear” from SLR went....if the model doesn’t fit the reisduals won’t have the same scattering around the 0 line
    - “identically distributed” also means that the sample and the population should look the same
      - Asking how fast people run the 100m dash is a fair question
      - But using Olympic run times is not a good representation of people generally
  - NOTE: Both IID and the “pop is normal” can be check, at least in part, via the residuals!
  - Justify/explain why we color residuals by explanatory variables

4 ANOVA

- You WILL NOT be expected to fill out an ANOVA table
  - You WILL be expected to read an ANOVA table and interpret it’s output
- Identify the null and alternative hypothesis
  - Including limitations of what they indicate (doesn’t identify which mean is higher/lower/different)

5 Required Readings

- Three major papers I want you to focus on are tidy data, the box paper, and the ASA statement on p-values
- Tidy Data
  - 1 variable per column
  - 1 observation per row
  - Break data into different tables if the observation units are different for different people

    * Eg a table for students with demographic info (ie one student per row)
    * Another table where each row is one student's score on one assignment
  - Human readable is not machine readable and tidy data is for the machines

- "Science and Statistics" by Box
  - General points he was getting at
  - Statisticians needs to understand the science behind the data or their conclusions may be meaningless/ridiculous
  - All models are wrong but we can still choose a lesser wrong model
  - It's best to balance theory and practice to become a great statistician
  - "Practice" means working with real world problems and data with specific research goals in mind

- ASA's Statement on p-values
  - What is a p-vlaue? What can it do for us?
    * Measure of incompatibility between the hypothesis and the data (smaller -¿ less compatible)
    * Measure of the probability we'd see a sample statisic this extreme if the null hypothesis is true
    * Is the foundation for much of hypothesis testing in statistics and research
    * In isolation by itself it can do/say very little
  - What does a p-value NOT do?
    * It does not give the probability the null hypothesis is true (nor the alternative)
    * It does not say the model is correct and the assumptions are met
    * When small, it does not say your results matter in the real world
      · effect size vs small p-values again
    * When large, it does not say your theory is wrong.
      · We lack the evidence to detect a difference/incompatibility
    * A large, non-sig. p-value is not a bad thing. An experiment returning a p-value of .14 is not a good nor bad thing.

6 Old Stuff

- What is a p-value?
- What is the difference between standard deviation and standard error?
- What does a confidence of 95% mean?

- What are the four main challenges colorblind people face?
- How can we make graphics more accessible?
- Be sure to know how to read the different graphs we have seen early in the semester
- How does a hypothesis test work?
- Read and interpret a confidence interval for some statistic.
- Make predictions using MLR models
  - Not as scary as it sounds I promise
- Be comfortable using indicators in regression
  - Eg make a prediction that uses an indicator variable
  - Eg make indicator variables in a data set