PHP2510

RECAP WEEKS 6-9

STUDENT FEEDBACK

Thank you for identifying your challenge areas. I appreciate your candor. Below outlines the common feedback and my responses. All feedback is paraphrased and expressed without judgment.

EXAM

STUDENT QUESTIONS

- What will it be like?
- How should I study?
- Will we be given a calculator or look-up tables?
- What are the core theorems that I need to know?

PROFESSOR RESPONSE

- The exam is divided into 5 sections:
    - CONCEPTS: assesses your understanding of statistical concepts via short-form answers
    - DERIVATIONS & FORMULAE: assesses your ability to use and explain common formulae
    - END-TO-END ANALYSIS: assesses your ability to solve multi-part word problems
    - INTERPRETING R CODE: assesses your ability to interpret R code and output
    - EXTRA CREDIT: additional challenge questions
- I will give you a set of practice problems for the exam, and a study guide. I will expand the "fact fluency" Ed Discussion post before the exam to help with "concepts" section.
- As before, there will be no calculator, no cheat sheets, etc.
- I will either give you the formula, or expect you to use one that is very common (e.g. confidence interval structure, applied to a one group test; definition of a p-value in terms of a probability statement).
- I will give you critical values for the relevant distributions when necessary.

PRACTICE PROBLEMS

STUDENT QUESTIONS

- Can I have more practice problems?

PROFESSOR RESPONSE

- Please use the book's problems. I post on Ed Discussion which ones I recommend (whenever asked - I've pinned those posts). I will review any answer for accuracy - you just need to post it.
- Please review the labs and try to complete them on your own. Use the "extra problems" as well.
- Make sure that, if I showed you all the code and the output from these labs, you could explain the answers and underlying concepts.

SPECIFIC QUESTIONS / REQUESTS

HIGH-LEVEL CONCEPTS

- How do I know which test to use?
    - Is there a flow chart of decisions? Yes (just google image it) but I hate it. Instead, I want you to think from first principles. You should be asking yourself:

        - What is my outcome variable? What am I willing to assume about it (in terms of distribution)? Note that “I want to assume nothing except that the population has some mean and some variance” is a perfectly reasonable choice.
        - What is the parameter I actually want to estimate? This is often the true population mean.
        - What statistic is a good estimate for that parameter? This is often the sample mean.
        - You will notice that problems I have given you before (e.g. exam extra credit) force you to walk through this process. You should see that if you start with a 1/0 outcome, you will naturally land on a proportions test (one or two groups). If you start with “No assumptions except the population has some mean and some variance”, you will naturally land on a t-test.
        - I do NOT want you to memorize what test works for what situation. I want you to have one problem solving process that works every time.
    - Why can’t we always use exact tests? If we can, we probably should (because it avoids relying on CLT)! But typically the exact test only exists when: (1) the set of outcomes are very small (tea experiment) or (2) we are working with a few, very specific distributions (e.g. Binomial for Rx glasses example). We will discuss this more next week.
    - Should I do a z-test vs t-test? In reality, we rarely use a z-test because it requires knowing the true population variance, which is quite unusual. But when our n is very large, the difference between these two distributions is negligible, so we can use the z-test (or look-ups with z distribution) as an approximation.
    - When do I use a paired t-test? If you are studying a difference in group means, but those groups are “paired” in some manner (i.e. not independent), you should using the non-independence to your advantage. The example you have seen: calories in chocolate vs vanilla ice cream, when you have a sample of “brands” of ice cream, and 2 observations for each sample. Doing a paired test is equivalent to first taking the difference *within* each brand; then comparing *across* brands (which are now independent). It is not the point estimate that changes, but our standard error. <board work too>
        - Also see CHIHARA 7.20 (slides)
- Can you explain how to pool variances and why? <board work>
    - Also see take-home challenge problem CHIHARA 7.38 (slides)
- When am I allowed to use the Central Limit Theorem? For large n, or for modest n when your distribution is roughly symmetric to begin with. I want you to know when you are invoking the CLT, so you don’t rely on a test where that assumption is ridiculous. We will also study other options that avoid the CLT next week. But apart from exact tests, everything else we’ve done so far relies on CLT.
- Can you show how to work with more “advanced” confidence intervals: <board work>
    - Binomial / Poisson / Exponential, etc.
    - Non-symmetric
    - Also see CHIHARA 7.48 (slides)

EXPLANATIONS, TAKE 2

Can you explain again :

- sigma vs s; SE vs SEM? See here. <board work>
- The relationship between confidence intervals and hypothesis testing? You fail to reject the null hypothesis (at some alpha level) for every value of your parameter that is within the 1-alpha% confidence interval
- How to write a probability statement in the context of a “power” calculation? <board work>
- The general hypothesis testing paradigm? <slide discussion>
- Experimental design and power? <slide discussion>

ONE-OFF QUESTIONS

- What is the definition of alpha and beta?

    - Alpha = probability you reject the null hypothesis when the null hypothesis is true. This is also called our "significance level". We typically set this first (we want a test that guarantees this above all else). Alpha is also the largest p-value for which I will still reject the null hypothesis.
    - Beta = probability you fail to reject the null hypothesis when the null hypothesis is false. This is typically of secondary importance, but important to determine to know whether your test is useful. A high beta means your test is not useful at actually detecting an effect.
    - Each of these can be written as probability statements about the test statistic (which has a distribution, either under the null hypothesis or the alternative hypothesis).
- What does it mean when you say "distribution of"? <board work>
- What is the test statistic (for a one-sample t-test)?
    - A "statistic" is a function of your data. A "test statistic" is a function of your data that allows you to perform an hypothesis test. That means you must know its distribution under the null; and it must give you evidence away from the null when the alternative is true.
    - For a z-test, you are welcome to say either of the following are your test statistic, because they are logically equivalent.
        - X_bar under the null: ~ N(mu0, sigma^2/n)
        - (X_bar - mu0) / (sigma/sqrt(n)) under the null: ~ N(0,1)
    - For a t-test, you change sigma to s above, N to t, and specify your df. I prefer the second writing – because you likely need that result to continue anyways, and it's better aligned with R syntax. But it's not wrong to say that X_bar (unstandardized) is your test statistic.
- What is the relationship between "hypothesis testing" and "t-test"?
    - A t-test is a specific hypothesis test (see above for when the t-test is appropriate to use).
- How do I look up critical values, especially for a t-distribution or non-symmetric distributions?
    - Before you get started, write out your test statistic and the distribution it follows. I suggest standardizing it (I think it's always easier to work with that one).
    - Decide how much mass you want in your tail(s) (alpha? alpha/2)? Let's say it's 5%.
    - Decide which tail you care about (Lower one? Upper one? Both?). Let's say it's the upper one.
    - Determine any other necessary parameters. For a t-distribution, you need to know the degrees of freedom (which = n-1). Let's say you have 100 subjects.
    - Then use the distribution that your standardized test statistic follows:
        - qt(0.05, df = 99, lower.tail = F); output = 1.66. How to interpret this? At the value 1.66, there is 5% mass to the right for a standardized t-distribution with 99 degrees of freedom.
        - The above is the cut-off value for a 1-sided (alternative is that the mean is bigger) alpha = 0.05 t-test; Or for a two-sided alpha = 0.1 t-test (the cutoff values would be both +1.66 and -1.66 because t-distributions are symmetric).
    - Use the same process for any distribution. If it's non-symmetric and you are doing a 2-sided test, you need to use R to look up both critical values (they aren't +/-).

MISC (OTHER)

- Do I need to know bootstrap/permutation tests? Yes - that is next week's content.

MISC (POSITIVE)

- I like the visualizations and simulations. 🙂
- I am feeling more confident and comfortable with the pace of the class recently. 🎉