Rubric for Homework 6 (36-309)

October 30, 2024

Question 1 (28 points)

Part A (16pts): Here's the point breakdown (please also read the text after this breakdown):

- 8pts: First, 4pts for showing *any* work. Then, just check that they got the correct number of positive results (470) and negative results (530); if so, award 4pts and move on. Otherwise, check for the following mistakes. First, check if they made an error when calculating the number of times they reject/fail-to-reject for the 160 null-is-true studies; if so, deduct either 0.5pts, 1pt, or 1.5pts, depending on the severity of the mistake. Then, check if they made an error when calculating the number of times they reject/fail-to-reject for the 840 null-is-false studies; if so, deduct either 0.5pts, 1pt, or 1.5pts, depending on the severity of the mistake. Finally, just check that their number of positive and negative results adds up to 1000; if it doesn't, deduct 1pt.

- First, 4pts for showing *any* work. Then, just check that they got the correct percentages for positive results ($462/470 \approx 98.3\%$) and for negative results ($92/530 \approx 17.4\%$); if so, award 4pts and move on. Note that it's fine if they answered in terms of fractions, percentages, or decimals. If they got everything right but accidentally reported the percentage of *incorrect* results instead of correct results (i.e., 1.7% and 82.6%), deduct just 1pt total. Otherwise, check for the following mistakes. When computing the total number of correct positive results, they should have taken the total number of rejections for the three null-is-false

cases (1pt) and divided by the total number of rejections (1pt). Meanwhile, when computing the total number of correct negative results, they should have taken the total number of fail-to-rejections for the null-is-true case (1pt) and divided by the total number of fail-to-rejections (1pt). For each of these 1pts, please grade based on the numbers they got in the previous part. Thus, you should grade based on these *conceptual* steps, rather than the actual numbers. For each conceptual step they are clearly getting incorrect, deduct either 0.5pts or 1pt, depending on the severity of the mistake.

For this question, students might be somewhat creative—please reach out if you think this rubric doesn't apply to how some students answered this question.

Part B (12pts): Here's the point breakdown:

- 6pts: First, 1.5pts for stating that it's only possible to make a Type 1 error for the (1) studies; if they say anything else, deduct 1pt. Meanwhile, 1.5pts for providing *any* explanation. Then, 1.5pts for stating that it's only possible to make a Type 2 error for the (2), (3), and (4) studies; if they say anything else, deduct 1pt. Meanwhile, 1.5pts for providing *any* explanation. For this part, they don't have to literally use the (1)-(4) labels, as long as it's clear which studies they're referring to.

- 6pts: First, 3pts for providing *any* explanation. Then, if they stated the correct number (446), award full credit and move on. If they accidentally stated the total number of correct results (554), deduct only 1pt total. Otherwise, check for the following mistakes. First, check if they correctly identified the number of false positives (8); if there's a mistake here, deduct 1pt. Then, check if they correctly identified the number of false negatives (438); if there's a mistake here, deduct 1pt. However, for this part, please be somewhat lenient—Part A asks them to compute the number of positive/negative results, so they may well use the same numbers they computed from Part A, and thus it's okay if their explanation here is very short. If it's obvious that their mistake is due to a mistake they made in Part A, but they're otherwise correctly interpreting incorrect numbers, do not necessarily deduct any points here.

Question 2 (12 points)

Here's the point breakdown:

- 4pts: 2.5pts for stating that Type 1 Errors will increase (deduct 1.5pts for saying anything else), and 1.5pts for providing *some* kind of reasoning for their answer. For their explanation, deduct 1pt only if they are obviously not discussing the idea of rejecting or failing to reject.

- 4pts: 2.5pts for stating that Type 2 Errors will decrease (deduct 1.5pts for saying anything else), and 1.5pts for providing *some* kind of reasoning for their answer. For their explanation, deduct 1pt only if they are obviously not discussing the idea of rejecting or failing to reject.

- 4pts: 2.5pts for stating that Power will increase (deduct 1.5pts for saying anything else), and 1.5pts for providing *some* kind of reasoning for their answer. For their explanation, deduct 1pt only if they are obviously not discussing the idea of rejecting or failing to reject.

Question 3 (60 points)

Part A (11pts): Here's the point breakdown:

- 3pts: 1pt for stating that the between-groups degrees of freedom is 2, and 1pt for stating that the within-groups degrees of freedom is 192. Then, 1pt for providing *any* explanation.

- 4pts: 0.5pts for providing *any* code, and 0.5pts for providing *any* explanation. Then, ideally, they should have used the `pf()` function. Within their code, they should have done three things: Correctly plugged in the observed $F = 1$ value as the first argument (1pt), specified the `df1` and `df2` arguments (1pt), and computed $1-$`pf()` instead of just `pf()` (1pt). For the degrees of freedom, don't deduct points if they simply plugged in their (possibly incorrect) degrees of freedom from the first part. Then, for the rare person who didn't use `pf()`, please try to apply each of these concepts to their code, but if it's difficult to apply this rubric to their submission, please let me know.

- 4pts: 1pt for providing *any* code, and 1pt for providing *any* explanation. Then, ideally, they should have used the `qf()` function. Within their code, they should

have done two things: Correctly plugged in the 1 – 0.05 = 0.95 quantile as the first argument (1pt), and specified the `df1` and `df2` arguments (1pt). For the degrees of freedom, don't deduct points if they simply plugged in their (possibly incorrect) degrees of freedom from the first part. Then, for the rare person who didn't use `qf()`, please try to apply each of these concepts to their code, but if it's difficult to apply this rubric to their submission, please let me know.

Part B (10pts): In this problem, they are asked to name three ways that we can increase the power in one-way ANOVA. They must mention these three things in some way:

- Increase (1.5pts) the sample size (1.5pts).
- Reduce (1.5pts) the error/residual variance of the experiment (1.5pts). It's okay if they just say "variance," "noise," or a similar concept.
- Increase (1.5pts) the variation among the population group means (1.5pts).

For each of these 1.5pts, if there is an obvious mistake, deduct 1pt. Please be somewhat lenient here—they don't have to say the above things verbatim, but they should mention each concept in some way. This question is more so about each of these concepts, rather than exact terminology. That said, if they don't *any* give a qualitative description for σ^2, σ_A^2, or n in their answer, deduct 1.5pts for each that they don't provide a qualitative description for. (To be clear, it's okay if they use σ^2 somewhere, as long as they also provide a qualitative description, as I do in the solutions.) The last 1pt is a "gimme point" for anyone who attempts this question.

Part C (15pts): Here's the point breakdown:

- 4pts: First, 1pt for having code that successfully defines `sigma2`. Then, 2pts for defining `sigma2` as 1010; deduct 2pts if they define it as anything else. Then, 1pt for providing *any* explanation.
- 5pts: First, 1pt for providing *any* code that successfully defines `eta.better` and `eta.worse`. Then, 1pt for ultimately having the same numbers for `eta.better` and `eta.worse`. Finally, 3pts for correctly defining `eta.better` and `eta.worse`. Their definitions should involve the following concepts in some way: the correct numerator (1pt), dividing by `sigma2` (1pt), and using `sqrt()` and `sum()` (or equivalent functions) correctly (1pt). For this part, do not deduct points if

they did everything correctly, and the only mistake is that `sigma2` isn't correctly defined.

- 6pts: Defined `groups = 3` (1pt), `n = 65` (1pt), `between.var = var(groupmeans.better)` OR approximately 33.3333 (1pt; don't worry about number of digits here), and `within.var = sigma2` OR their estimate of σ^2, whatever it is (1pt). Then, their explanation should touch on the idea that if we ran an experiment with these settings, then the chance that we correctly reject the null hypothesis is 33.6% (or whatever power value they obtain). They don't have to state this verbatim, but they should get at this interpretation in some way; if it's obvious that they aren't discussing the probability of correctly rejecting the null hypothesis (i.e., rejecting the null when it is false), deduct 1pt.

For this question, you don't necessarily have to check everyone's code carefully. For everyone who got the correct output, you can award them full credit (assuming they also gave correct explanations). You only have to go through the rubric for code when someone obviously does not have the correct output (but note that we shouldn't deduct someone points multiple times if their only mistake is that `sigma2` isn't defined correctly).

Part D (10pts): Here's the point breakdown for this question:

- 1pt: Do they have `power.anova.test()` code that runs, such that `power.seq` is successfully defined?

- 3pts: Defined `groups = 3` (1pt), `between.var = var(groupmeans.better)` OR approximately 75 (1pt), and `within.var = sigma2` OR their estimate of σ^2, whatever it is (1pt). Note that they defined `sigma2` in Part C, and we don't need to deduct points here if they did not define that correctly.

- 3pts: Made a scatterplot (1pt) with `n.seq` on the x-axis (1pt) and `power.seq` on the y-axis (1pt).

- 3pts: Wrote an interpretation for the plot. For those who made the correct plot, just check that they mention in some way that power increases as sample size increases. If they obviously don't get at this idea, deduct either 1pt, 1.5pts, or 2pts, depending on the severity of the mistake. Meanwhile, for anyone who

doesn't get the correct plot, please grade according to their plot. Ultimately, they just need to discuss the relationship between sample size and power based on their plot; so, deduct either 1pt, 1.5pts, or 2pts if they make an incorrect interpretation about this relationship from their plot.

Part E (14pts): Here's the point breakdown:

- 6pts: First, used `anova.power.test()` (0.5pts) and defined `groups = 3` (0.5pts), `between.var = var(groupmeans.better)` OR 75 (0.5pts), and `within.var = sigma2` OR their estimate of σ^2, whatever it is (0.5pts), and set `power = 0.8` (1pt). Note that they defined `sigma2` in Part C, and we don't need to deduct points here if they did not define that correctly. Then, 2pts for stating the rounded-up version of the n they obtain; if they only report the decimal-version or round down, deduct 1pt, but if they report a number that isn't close to the n they obtain, deduct 2pts. Finally, 1pt for providing *any* explanation.

- 8pts: First, used `anova.power.test()` (0.5pts) and defined `groups = 3` (0.5pts), `n = 75` (1pt), `within.var = sigma2` OR their estimate of σ^2, whatever it is (0.5pts), and set `power = 0.8` (1pt). Note that they defined `sigma2` in Part C, and we don't need to deduct points here if they did not define that correctly. Then, 1.5pts for reporting the `between.var` they receive from their output; if it's obvious they aren't reporting the correct number from their output, deduct 1.5pts (but don't necessarily deduct points if they correctly interpret incorrect output). Then, 2pts for saying that this is smaller than the variance of `groupmeans.better`; if they say anything else, deduct 1.5pts (however, do not necessarily deduct points if they correctly interpret incorrect output). Finally, 1pt for providing *any* explanation.

Question 4 (5 points)

Here's the point breakdown:

- 1pt: Correctly defined `sigma2.seq` as the numbers $25, 50, 75, \ldots, 975, 1000$. They don't necessarily have to use `seq()`, but it should be defined correctly.

- 1.5pts: Within `anova.power.test()`, defined `groups = 3`, `n = 50`, and `between.var = var(groupmeans.better)` OR 75 (0.5pts). Furthermore, defined `within.var = sigma2.seq` (1pt).

- 1pt: Made a scatterplot with `sigma2.seq` on the x-axis and `power.seq.var` on the y-axis.
- 1.5pts: Stated in some way that power decreases as variance increases. They don't have to say this verbatim, but if they obviously don't get at this idea, deduct either 1pt or 1.5pts, depending on the severity of the mistake.