coursera (/)

Like this course? Become an expert by joining the Methods and Statistics in Social Sciences Specialization (/specializations/social-science).

Upgrade

(https://accounts.coursera.org/i/zendesk/courserahelp?return_to=https://learner.coursera.help/hc)

Course Forums

- All Course Discussions (/learn/infer...
- General Discussion (/learn/inferenti...
- Meet and Greet (/learn/inferential-s...

Module Forums

- Before we get started... (/learn/infer...
- Comparing two groups (/learn/infer...
- Categorical association (/learn/infer...
- Simple regression (/learn/inferentia...
- Multiple regression (/learn/inferenti...
- Analysis of variance (/learn/inferent...
- Non-parametric tests (/learn/infere...
- Exam time! (/learn/inferential-statis...

Help Center

← All Course Discussions

(/learn/inferential-statistics/discussions?sort=lastActivityAtDesc&page=1&q=)

DS Exam Q29 - very many

(/learn/inferential-statistics/profiles/1f7f042688de1cdf42d40ee185779133)

David Scott Exam time? (/learn/inferential-statistics/module/jyCOE/dis days ago (/learn/inferential-statistics/discussions/Zw4w

Question 29 seems to have meaningful ambiguity in its formulation. Does anyone have a good understanding of the answer/rationale to any of the below questions:

- Can we assume equal variance? My default is to say "no" and go through the more complex calculations for pooled SE and associated df.
- If assuming unequal variance, should df be truncated (as per Excel) or rounded or applied as-is (as per R)? The truncated and as-is df produce the same p-value in this specific case, while rounded df produces a p-value that is 0.001 lower. There is really no good case for rounding-up df, and I prefer the R as-is approach
- Do we assume a two-tailed test? My default is to say "yes" due to the

issues below, but that is ambiguous based on the problem statement

- If a one-tailed test, which direction? My default Ha is that procrastination score is higher in the graduated group based on the precise phrasing of the problem set-up. The issue is that it is obvious from the data that this is not the case, plus it is a weird hypothesis that higher procrastination will be positively associated with graduating. Perhaps it is one-sided test for not graduated > graduated?
- And, after all that, what p-value do you enter? If you got p=0.1234, do you enter 0.123 (round to 3 digits) or 0.124 (to match the strict "p<" from the instructions)? To make it worse, if you use the strict p-hurdle approach, then the rounded df and as-is df agree on p-value, but truncated df does not (it will send p 0.001 higher)

Below are simulated data to match the provided parameters. I am not sure what to use for var.equal and alternative nor how to truncate or ceiling or round or format the resulting p-value. My defaults FALSE (use resulting df "as is" for full precision) and "two.sided" and round to 0.xxx do not match the auto-grader.

```
ngT <- c(5, 8, 5, 8, 6, 8, 5, 5, 5, 7, 7.523214, 3.770386)
gT <- c(5, 6, 5, 4, 4, 4, 6, 5, 7, 7, 6, 5.049329, 2.657571)

mean(ngT) ; sd(ngT) ; length(ngT) ## 6.1078 1.51213 12
mean(gT) ; sd(gT) ; length(gT) ## 5.1313 1.26706 13

## This p-value rounded to 0.xxx is not an auto-graded correct answer
t.test(ngT, gT, paired=FALSE, var.equal=FALSE, alternative="two.sided")
```

Any ideas?

▲ 0 · Follow · Reply to David Scott

Earliest (/learn/inferential-statistics/discussions/Zw4w1vgwEeWPcBKgw3iPNw?sort=createdAtAsc&page=1)

Top (/learn/inferential-statistics/discussions/Zw4w1vgwEeWPcBKgw3iPNw?sort=upvotesDesc&page=1)

Most Recent (/learn/inferential-statistics/discussions/Zw4w1vgwEeWPcBKgw3iPNw?sort=createdAtDesc&page=1)

ZB (/learn/inferential-statistics/profiles/3cdb3ac214a020ebfbe64f1e50b03eac)

Zara Bamdad, Teaching Staff · 3 days ago (/learn/inferential-statistics/discussions/Zw4w1vgwEeWPcBKgw3iPNw/replies/mlE0MPsmEeWW0grgb0UTkQ) · Edited

Thanks for posting :) This is a tough post to respond to without giving too much away!

The question expects you to figure out the equal/unequal variance assumption based on values provided in the question. The tails/direction of the test should be determined from the question itself, rather than the values. The question also relies on you using the table to get the threshold, rather than R or Excel (will make this explicit in the question!).

Does this help at all?

▲ 0 Upvote · Reply

1

SD (/learn/inferential-statistics/profiles/f34069ce8df6de7dbebfedfb7e760d9f)

Reply

Reply