

BIOM/SYSC5405 – Pattern Classification and Experiment Design

Assignment 2— Peer Evaluation Guide

Overall Guidance

- Only evaluate “presentation” once (in the first rubric criterion). Please don’t penalize poor presentation quality repeatedly throughout the rubric.
- Code can be either inline or in an appendix.
- Given word limits for answers are just a guide. Please allow ~150% of that limit before considering a response to be unduly long.
- The peer review has five phases:
 - o 12 Oct: submit your own solution
 - o 16 Oct: evaluate three of your peers’ solutions
 - o 18 Oct: deadline to review the evaluations that you have received and provide feedback by replying to your reviewers’ comments
 - o 20 Oct: deadline to double-check all three of the evaluations that you gave to see if you received any replies to your original grading.
 - o **For each reply, you must respond with a new comment justifying your grade (whether you agree to change it or not).** Respond by adding a ‘reply’ to each of the ‘replies’ that you received.
 - o If you fail to respond to any replies, your own grade may be impacted.
 - o 22 Oct: deadline to let me know if there are unresolved issues with the grading that you have received.
- Please keep in mind that changing a single grade by 15% (e.g., “QA goes from 75% to 90% for one of your graders”) will have a negligible impact on your final grade:
$$+15\% / (3 \text{ graders}) * (14\% \text{ per question}) / (4 \text{ assignments}) * (30\% \text{ of final grade}) = +0.0005$$

Question A

For Question A, please grade Q1, where students were asked to provide a contingency table for the healthy temp & RR data (converted to categorical and ordinal variables). A chi-squared test is required.

- 1) A contingency table is required. It should look like the following, but can be transposed:

	RR_ord			
T_cat	RR-low	RR-med	RR-high	All
t-normal	26	45	35	106
t-fever	33	31	30	94
All	59	76	65	200

- 2) The null hypothesis, H_0 , must be stated. It should be something like “**T_cat and RR_ord are independent**”.
- 3) An alternate hypothesis H_1 must be stated, such as: “**T_cat and RR_ord are correlated**” or “RR_ord depends on T_cat” (or vice versa).
- 4) χ^2 value should be: **3.085** ($\pm 5\%$)
- 5) Degrees of freedom should be: **2**
- 6) p-value should be: **0.213** ($\pm 5\%$)
- 7) Students must state their conclusion, which should be something like: “**Since the p-value is above our alpha threshold of 0.05, we fail to reject our null hypothesis, H_0** ” or “**since the p-value is above our alpha threshold of 0.05, there is insufficient evidence to reject our null hypothesis, H_0 , and therefore cannot accept H_1 .**”

- It is not necessary to show the 'expected' contingency table values, but they are:

[[31.27 40.28 34.45], [27.73 35.72 30.55]]

Common errors:

- One of the seven elements above are missing (each missing element counts as a minor error)
 - Note that the contingency table does not have to include row and column marginals
- The student arrives at the incorrect conclusion
- The chi-squared value or p-value are off by more than 5%
- The DOF is not 2
- Note: If one error causes other errors, just count it as one error. If everything else looks correct (peak at code), but one error (e.g., incorrect DOF) causes downstream errors, that should be treated as a single error.

Question B

For Question B, please grade Questions 2 & 3.

Question 2:

- IQR of T_healthy = **0.36425** ($\pm 5\%$)
- The 10% trimmed mean of T_healthy = **36.786** ($\pm 5\%$)

Question 3:

- Student uses bootstrapping to compute 90% confidence interval on trimmed mean of T-healthy
- The 90% confidence interval on the 10% trimmed mean of T-healthy is **[36.756, 36.816]**

Common Errors

- Any of the answers is wrong by more than 5% (one minor error per erroneous answer)
- The 10% trimmed mean from Q2 falls outside of the 90% confidence interval from Q3
 - o This would be a separate error from having an incorrect trimmed mean of interval bounds, since conceptually, the observed trimmed mean should fall inside the interval.
- Confidence interval was computed using an approach other than bootstrapping.

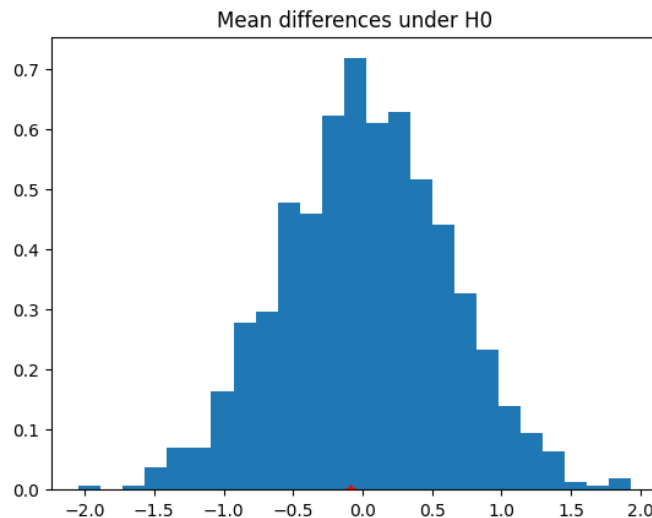
Question C

For Question C, please grade Question 5.

Student should describe how they used a permutation test to test whether RR_covid has significantly greater mean than RR_healthy. The description should include:

1. The null hypothesis, H_0 , which should be something like: "There is no difference between the means of RR_covid and RR_healthy"
 - Will accept 'Same' or 'not different'; not 'less than'
2. The alternative hypothesis, H_1 , which should be something like: "The mean of RR_covid is significantly greater than the mean of RR_healthy"
 - 'Greater than', not 'different'
3. A statement of how many permutations they conducted
4. A statement of their p-value, which should be **0.566** ($\pm 5\%$)
5. A concluding statement, which should be like **"Given that the p-value is above our alpha-threshold of 0.05, we cannot reject the null hypothesis"**.

Optionally (for 'bonus marks'), they could illustrate the distribution of differences of means under H_0 , which will look *something* like:



Common Errors:

- A description is given, but it is missing 1-2 elements listed above
 - If more elements are missing, this becomes a 'major' error
- The wrong p-value is calculated (off by more than 5%)
 - This is likely due to using a 2-tailed test, when they should have used a 1-tailed test
- Their conclusion doesn't agree with their p-value
 - Note that if they arrived at the wrong p-value, then their conclusion should align with that incorrect p-value (e.g., if they computed a p-value of 0.02, and therefore concluded that H_0 can be rejected and H_1 accepted, then this represents a single error (p-value only)).