

Biostatistics Methods, STAT 5500/6500

Spring 2021 Homework 1

Assignment Due (by 11:59 P.M.): Wed 10 Feb

Directions: In completing these Exercises, use the SAS tools presented in class. You may find the SAS Help useful. Where appropriate, you may perform calculations by hand. Note that neatness and format (including SAS code in an appendix) will contribute 5 points to the total score. This assignment will be graded out of 50 points (70 points for 6500 students).

1. Spirometry involves measuring lung function, and values can be reported in terms of peak flow rate (liters per minute). Researchers in the Williams lab wanted to see if aminophylline treatment affects spirometric [peak flow rate] values in severe acute asthmatic patients any differently than did salbutamol, another potential asthma treatment. Nine patients were randomly assigned to receive aminophylline, and eleven randomly assigned to receive salbutamol. Each patient's spirometric value was recorded, then they received their designated treatment, and their spirometric value was re-measured one hour later. The aminophylline group saw a mean spirometric increase of 44, with a standard deviation of 20. The salbutamol group saw a mean increase of 86, with a standard deviation of 15.
 - (a) (5 points) Conduct an appropriate test of significance for the Williams lab, using significance threshold $\alpha = .05$. Be sure to specify (i) the name of your chosen test, (ii) null and alternative hypotheses, (iii) numeric value of test statistic, (iv) p-value, and (v) conclusion (in context). You may assume the spirometric increase values for both groups of patients are approximately normally distributed.
 - (b) (5 points) Clearly explain what your p-value from 1a means. Do not just repeat what your conclusion is – what does the number itself mean?
 - (c) (5 points) Assuming a common standard deviation as estimated from the data, what would be the power of this test if the true effect of aminophylline were an increase of 20 liters per minute more than salbutamol (i.e., the mean difference is 20) ?
 - (d) (5 points) Clearly explain what power means in this context.
 - (e) (5 points) Construct a clear visual representation of how the power of this test varies with the underlying true mean difference.
2. Twelve other labs independently assessed the same effect of aminophylline treatment on spirometric values in severe acute asthmatic patients, as compared to another potential treatment. The p-values from these twelve other tests of significance are in the following table:

0.350697	0.88816	0.005715	0.000684	0.010598	0.850057
0.651455	0.000005	0.075073	0.006792	0.9034	0.949922

Add your p-value from Exercise 1a to this list of p-values.

- (a) (5 points) In considering this family of 13 tests, explain clearly why it might be worthwhile to adjust for multiple hypothesis tests.
 - (b) (5 points) Report a table of raw and adjusted p-values, using the following adjustments: Bonferroni, Sidak, Holm (step-down Bonferroni), Hochberg (step-up Bonferroni), and False Discovery Rate (Benjamini-Hochberg).
 - (c) (5 points) Which of these adjustments would you recommend, and why?
 - (d) (5 points) After making this adjustment, what change, if any, would you make to your conclusion regarding the aminophylline effect in the Williams lab?
3. (6500 students only; 15 points) First, read the paper by Raimund Alt (link in Canvas), paying particular attention to his “closed LSD-procedure”. Then consider a multiple regression model with many predictors:

$$Y_i = \beta_0 + \beta_1 X_{i,1} + \beta_2 X_{i,2} + \dots + \beta_k X_{i,k} + \epsilon_i.$$

When we fit such a model and look for significant predictors, we test a family of null hypotheses $H_{0,j} : \beta_j = 0$ for $j \in \{1 \dots k\}$. However, the typical approach is to not adjust the resulting set of k p-values for multiple hypothesis testing. Is this lack of adjustment justified by Alt’s “closed LSD-procedure” approach (i.e., is the closed LSD-procedure unnecessary because these two approaches are equivalent)? Write a convincing paragraph explaining why or why not. As part of your discussion, report the conclusions of the two approaches when applied to the two sample data sets on the course website (both use $k = 3$). Do the two approaches yield equivalent conclusions in both data sets? You may find the following SAS code to be a helpful template.

```
proc reg data=a1;
  model Y = X1 X2 X3;
  X1X2: test X1=X2=0;
run;
```

4. (6500 students only; 5 points) After reading the Hoenig and Heisey paper (link in Canvas), write a paragraph explaining in your own words why post-experiment power calculations are inappropriate.

(5 points) SAS code in appendix