

## BIOS 6611 Homework 7

Due Tuesday, October 29, 2019 by 11:59 pm to Canvas Assignment Basket

Provide all relevant code and output with your answers to each question summarizing your conclusion based on your results.

1. Bayesitis, a disease that causes a fear of incorporating prior knowledge to everyday decision making, is treatable through an in-patient stay at the hospital. Upon discharge, patients have returned to their baseline behavior and are able to resume their normal daily activities. You learn that one hospital has developed a new treatment for Bayesitis, and you are able to obtain information for the length of stay (LOS) on a random sample of patients from different hospitals that either use the new method or the standard of care (SOC):

Hospital (Treatment)	LOS (days)
Cauchy General (New)	3, 3, 4, 5, 5, 5, 6, 7, 7, 8, 9, 15
Skellam Memorial (SOC)	6, 7, 7, 7, 8, 8, 8, 9, 9, 10, 10, 11, 13, 13, 15

- a. A study based on 1,500 observations using the SOC identified a median LOS of 9 days. For both hospitals, test if **their median LOS is significantly different from the historic median of 9 days using a function in R** (e.g., `wilcox.exact`, `binom.test`, `SIGN.test`, etc.) and interpret your results.
  - b. Create a 2x1 panel figure showing histograms of the LOS for Cauchy General and Skellam Memorial that have a range from 0 to 16 for their x-axis, have bins of width 2 (e.g., “breaks” at 0, 2, 4, 6, ..., 14, and 16 in the base R `hist` function), and informative titles and labels. [Hint: One option to creating a panel figure is to use `par( mfrow=c( number rows, number columns ) )`.]
  - c. Carry out a nonparametric procedure for testing the hypothesis that lengths of stay are comparable in the two hospitals by using a function in R (e.g., `wilcox.exact`, `binom.test`, `SIGN.test`, etc.). For your interpretation, if appropriate, reference your histogram from part (b) for discussion of the shape of the two groups.
2. Suppose we have separately analyzed the effects of 10 single nucleotide polymorphisms (SNPs; [https://en.wikipedia.org/wiki/Single-nucleotide\\_polymorphism](https://en.wikipedia.org/wiki/Single-nucleotide_polymorphism)) comparing people with type I diabetes vs. controls. The p-values from these analyses are given below.

Effects of 10 SNPs on Type I Diabetes			
SNP	p-value	SNP	p-value
1	0.040	6	0.620
2	0.100	7	0.001
3	0.400	8	0.010
4	0.550	9	0.800
5	0.340	10	0.005

Use the **FDR method in R**, SAS or by hand to correct for multiple testing using an FDR = 0.05. After correction, which SNPs show statistically significant effects?

3. Twenty-two young asthmatic volunteers were studied to assess the short-term effects of sulfur dioxide ( $\text{SO}_2$ ) exposure under various conditions. The baseline data in the table (Table 12.30 from Rosner) were presented regarding the relationship of bronchial reactivity to  $\text{SO}_2$  (cm  $\text{H}_2\text{O/s}$ ) stratified by lung function (as defined by forced expiratory volume / forced vital capacity [ $\text{FEV}_1/\text{FVC}$ ]) at screening.

Lung-Function Group		
Group A $\text{FEV}_1/\text{FVC} < 74\%$	Group B $\text{FEV}_1/\text{FVC} 75\text{-}84\%$	Group C $\text{FEV}_1/\text{FVC} \geq 85\%$
20.8	7.5	9.2
4.1	7.5	2.0
30.0	11.9	2.5
24.7	4.5	6.1
13.8	3.1	7.5
	8.0	
	4.7	
	28.1	
	10.3	
	10.0	
	5.1	
	2.2	

Using SAS or R:

- Assume that the variances across the groups are equal and test the hypothesis that there is an overall mean difference in bronchial reactivity among the three lung-function groups.
- If justified, compare the means of each pair of groups using the Tukey HSD method and summarize the results. Otherwise note why it isn't justified.
- EXTRA CREDIT:** Carry out part (a) assuming that the variances across the groups are not equal. If justified, describe a way to compare the means of each pair of groups, but do not carry out any further analysis.