Multiple Testing

Your name here

mm/dd/yyyy

Create a Word document from this R Markdown file for the following exercises. Submit the R markdown file and the knitted Word document.

## Exercise 1

For this exercise we’re going to follow along with the first part of the presentation and compute the error rates for three types of multiple testing correction. Like the presentation, in this example there are 1000 tests and we know when the null and alternative hypotheses are true. For each test we have: - value is from a normal distribution with - value is from a normal distribution with

Here is the data and p-values:

# Do not change this chunk  
set.seed(124)  
T0 = 900;  
T1 = 100;  
x = c( rnorm(T0), rnorm(T1, mean = 3))  
P = pnorm(x, lower.tail = FALSE )

For the first 900 tests is true while for the last 100 tests is true. Throughout this problem you should test with .

### Part 1a

Using the p-values from above how many discoveries are made? If the testing were working perfectly how many discoveries should have been made?

### -|-|-|-|-|-|-|-|-|-|-|- Answer 1a -|-|-|-|-|-|-|-|-|-|-|-

Replace text with your answer.

### Part 1b

If no correction (PCER) is made for multiple testing, then compute the Type I error rate, Type II error rate, and False Discovery Rate. Is the Type I error rate similar to what you would expect? Explain.

### -|-|-|-|-|-|-|-|-|-|-|- Answer 1b -|-|-|-|-|-|-|-|-|-|-|-

Replace text with your answer.

### Part 1c

Now attempt to control the family-wise error rate (FWER) using Bonferroni correction. Compute the Type I error rate, Type II error rate, and False Discovery Rate. How do these results compare to using no correction as in Part 1b?

### -|-|-|-|-|-|-|-|-|-|-|- Answer 1c -|-|-|-|-|-|-|-|-|-|-|-

Replace text with your answer.

### Part 1d

Repeat Part 1c using the Bonferroni-Holm Step-Down procedure to control FWER. Are the results any different than when using the ordinary Bonferroni correction?

### -|-|-|-|-|-|-|-|-|-|-|- Answer 1d -|-|-|-|-|-|-|-|-|-|-|-

Replace text with your answer.

### Part 1e

Would either of the Bonferroni correction methods be recommended if you were trying to discover possibly significant results for conducting further research into those results? Explain.

### -|-|-|-|-|-|-|-|-|-|-|- Answer 1e -|-|-|-|-|-|-|-|-|-|-|-

Replace text with your answer.

### Part 1f

Now apply the Benjamin-Hochberg procedure to achieve a target average False Discovery Rate (FDR) of 5%. Compute the Type I error rate, Type II error rate, and False Discovery Rate. Write a brief summary comparing these results to those above.

### -|-|-|-|-|-|-|-|-|-|-|- Answer 1f -|-|-|-|-|-|-|-|-|-|-|-

Replace text with your answer.

## Exercise 2

A pharmaceutical company is doing preliminary hypothesis testing of hundreds of compounds to see which ones might be useful in treating a rare form of cancer. What form of multiple testing correction should the company use (none, Bonferroni, or FDR)? Explain your reasoning.

### -|-|-|-|-|-|-|-|-|-|-|- Answer 2 -|-|-|-|-|-|-|-|-|-|-|-

Replace text with your answer.

## Exercise 3

Cars were selected at random from among 1993 passenger car models that were listed in both the Consumer Reports issue and the PACE Buying Guide. Pickup trucks and Sport/Utility vehicles were eliminated due to incomplete information in the Consumer Reports source. Duplicate models (e.g., Dodge Shadow and Plymouth Sundance) were listed at most once. Use the data set Cars93 to do the following. (Type ?Cars93 to learn more about the data.)

For the next two exercises we are going to use the Cars93 data set from the MASS package. We’ll delete the data having to do with vans so that we are only dealing with cars. The code to load and prepare the data is here:

Throughout this exercise we’ll compare population mean engine revolutions per minute at maximum horsepower (RPM) of the different types of cars (Type).

### Part 3a

Make a boxplot of RPM~Type. Is it reasonable to assume the RPM distributions are normal and have equal variances for the different types of cars?

### -|-|-|-|-|-|-|-|-|-|-|- Answer 3a -|-|-|-|-|-|-|-|-|-|-|-

Replace text with your answer.

### Part 3b

Conduct pairwise t-tests with no correction for multiple testing to see which mean RPM’s are different from each other. Summarize your findings including a brief description of which types of cars have significanlt different mean RPM ().

### -|-|-|-|-|-|-|-|-|-|-|- Answer 3b -|-|-|-|-|-|-|-|-|-|-|-

Replace text with your answer.

### Part 3c

Repeat 3b, but this time use Bonferroni correction.

### -|-|-|-|-|-|-|-|-|-|-|- Answer 3c -|-|-|-|-|-|-|-|-|-|-|-

Replace text with your answer.

### Part 3d

Now suppose we actually need to estimate the differences in the population mean RPM types while controlling for Type I errors using the Bonferroni correction. Use the onewayComp() function from the DS705data package with adjust = ‘bonferroni’ to compute the CI’s with 95% overall confidence. How much larger is the mean RPM for small cars than for large cars according to your estimates?

### -|-|-|-|-|-|-|-|-|-|-|- Answer 3d -|-|-|-|-|-|-|-|-|-|-|-

Replace text with your answer.

### Part 3e

Repeat 3d, but this time use the Tukey-Kramer procedure (use onewayComp() with adust = ‘one.step’ and var.equal=TRUE ). Again, how much larger is the mean RPM for small cars than for large cars according to your estimates?

### -|-|-|-|-|-|-|-|-|-|-|- Answer 3e -|-|-|-|-|-|-|-|-|-|-|-

Replace text with your answer.

### Part 3f

Simultaneous confidence intervals increase the width of the individual intervals to limit the probability that one or more of the intervals are wrong. Both Bonferroni and Tukey-Kramer can provide the family of simultaneous confidence intervals and maintain an overall confidence level of 95%. Compare your results from 3d and 3e. Which set of intervals do you think is better? Why?

### -|-|-|-|-|-|-|-|-|-|-|- Answer 3f -|-|-|-|-|-|-|-|-|-|-|-

Replace text with your answer.

### Part 3g

Even when you’re using a parametric procedure (one that assumes normality for instance), it can be useful to bootstrap the results to validate the choice of parametric procedure. Use onewayComp() to get the Tukey-Kramer confidence intervals with 95% confidence, but add nboot=10000 to have the code appoximate the critical values used to construct the intervals using bootstrapping. How do the results compare the theoretically derived Tukey-Kramer results from 3e? Does this increase your belief in the validity of the theoretical Tukey-Kramer results?

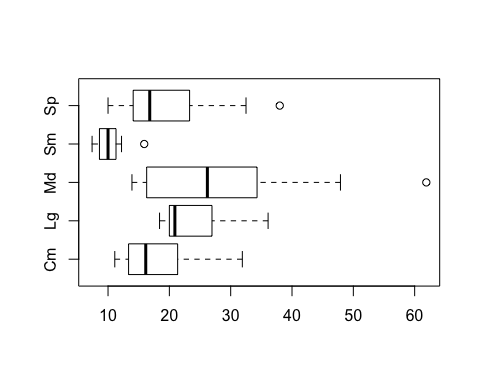
### -|-|-|-|-|-|-|-|-|-|-|- Answer 3g -|-|-|-|-|-|-|-|-|-|-|-

Replace text with your answer.

## Exercise 4

Now we are going to analyze differences in prices for different types of cars in the Cars93 data set. The boxplot below shows that the prices are skewed and variances are different.

boxplot(Price~Type,horizontal=TRUE,data=Cars93)



### Part 4a

It should be fairly clear that the price data is not from normal distributions, at least for several of the car types, but ignore that for now and use the Games-Howell procedure with confidence level 90% to do simultaneous comparisons (if interpreting the -values use ). (You can use onewayComp() with var.equal=FALSE and ajust = ‘one.step’). Use the CI’s to identify and interpret the largest significant difference in population mean prices.

### -|-|-|-|-|-|-|-|-|-|-|- Answer 4a -|-|-|-|-|-|-|-|-|-|-|-

# replace this with your code

Replace text with answer.

### Part 4b

Now repeat 4a, but since the price distributions are skewed use bootstrapping by specifying nboot=10000 in onewayComp(). Summarize how these results are different than in 4a.

### -|-|-|-|-|-|-|-|-|-|-|- Answer 4b -|-|-|-|-|-|-|-|-|-|-|-

# replace this with your code

Replace text with answer.

### Part 4c

Are the results in 4a and 4b very different? Which results seem more trustworthy, those in 4a or in 4b? Explain.

### -|-|-|-|-|-|-|-|-|-|-|- Answer 4c -|-|-|-|-|-|-|-|-|-|-|-

Replace text with answer.

### Part 4d

Since the prices are skewed it might be better to report differences in medians than in means. Use the boot package and Bonferroni correction to bootstrap 4 simultaneous confidence intervals with overall confidence level 90% for the difference in population median price between Sporty Cars and each of the other four car types. You don’t need to interpret the results.

### -|-|-|-|-|-|-|-|-|-|-|- Answer 4d -|-|-|-|-|-|-|-|-|-|-|-

# replace this with your code