MA256 Lesson 15 - Multiple Means (9.1-9.2)

Review: Comparing two means.

Hypotheses: When comparing two means our hypotheses (in symbols) were:

H_0 :

H_a :

Multiple-means: Similar to our method used when considering multiple-proportions:

Hypothesis test:

H_0 :

H_0 :

Like in chapter 8 when we compared multiple proportions, we need a statistic other than the *Mean Group Difference* statistic to make the transition to theory-based a smooth one.

This new statistic is called an F-statistic and the theory-based distribution that estimates our null distribution is called an F-distribution. Unlike the Mean Group Diff statistic, the F-statistic takes into account the variability within each group.

The analysis of variance F-statistic is (see p. 658):

Validity conditions:

As sample size increases, strength of evidence…

As the means move farther apart, strength of evidence…

As the standard deviations increase, strength of evidence…

Warm up:

If the balloons popped, the sound wouldn't be able to carry since everything would be too far away from the correct floor. A closed window would also prevent the sound from carrying, since most buildings tend to be well insulated. Since the whole operation depends on a steady flow of electricity, a break in the middle of the wire would also cause problems. Of course, the fellow could shout, but the human voice is not loud enough to carry that far. An additional problem is that a string could break on the instrument. Then there could be no accompaniment to the message. It is clear that the best situation would involve less distance. Then there would be fewer potential problems. With face to face contact, the least number of things could go wrong.

The statement on the previous page is from a study that examined college students' comprehension of the ambiguous prose passage. Before the college students were tested to see whether they understood the passage, they were randomly assigned to one of three groups, and then each group was read the passage under one of the following cue conditions: 1) students were shown a picture before they read the passage; 2) students were shown a picture after they heard the passage; 3) students were not shown a picture before or after hearing the passage.

Did you understand what the passage was describing? Would it help to have a picture?

Q1. Fifty-seven students from a small Midwestern college were randomly assigned to be in one of the three groups with 19 in each group. After hearing the passage under the assigned cue condition, they were given a test and their comprehension of the passage was graded on a scale of 1 to 7 with 7 being the highest level of comprehension. What was the null hypothesis in symbols?

Q2. Download the data. Plot the data with a dotplot, overlaid with a boxplot (use `alpha = 0.25`), and `facet_wrap()` to separate the figures by treatment level. Does there appear to be a difference between the three treatments?

```
# library(tidyverse)
# comp <- read.table("http://www.isi-stats.com/isi/data/chap9/Comprehension.txt",
#                     header = TRUE,
#                     stringsAsFactors = TRUE)
#
# comp %>% ggplot(...
```

Q3. Calculate sample sizes, sample means, and sample standard deviations. Also calcuate the overall mean of the comprehension scores. Using these values, calculate the F-statistic.

```
# comp %>% xxx
# xbar <- xxx
#
# VBG <- xxx
# VWG <- xxx
# fstat <- VBG / VWG
# c(VBG, VWG, fstat)
```

Q4. What does the value of the F-statistic mean in the context of the problem? Do you think you could calculate the p-value to determine the strength of evidence? Try it. What can we conclude about the study?

Q5. ANOVA. Take a look at the output from the `aov()` command. What are these numbers?

```
# aov.comp <- comp %>% aov(Comprehension ~ Condition, data = .)
# summary(aov.comp)
```

Diet Comparisons

Because about two-thirds of Americans are considered overweight, weight loss is big business. There are many different types of diets, but do some work better than others? Is low fat better than low carb or is some combination best? Researchers (Garnder et al. 2007) conducted a study involving four popular diets: Atkins (very low carb), Zone (40:30:30 ratio of carbs, protein, fat), LEARN (high carbohydrate, low fat), and Ornish (low fat). They randomly assigned women aged 25–50 with a body mass index (BMI) of 27-40 (overweight and obese) to one of the four diets. The 311 women who volunteered for the program were educated on their assigned diet and were observed periodically as they stayed on the diet for a year. At the end of the year, the researchers calculated the change in BMI (e.g., negative means reduction in BMI) for each woman and compared the results across the four diets.

```
diet <- read.csv("https://raw.githubusercontent.com/jkstarling/MA256/main/data/Diet_Data.csv", stringsAsFa
```

D1. What is the overarching research question the researchers hoped to answer?

D2. What are the observational units in this study? Identify the explanatory and response variables. Classify them as categorical or quantitative. For categorical variables, indicate how many categories are used.

D3. Does this study make use of random sampling, random assignment, both, or neither? What are the implications of your answer with regard to scope of inference? Did the researchers collect the data as paired data or as independent samples? In other words, according to the study design, are the responses from one treatment group paired with or independent of the responses from other treatment groups?

D4. State the null and alternative hypotheses, both in words and symbols, for testing the research conjecture. (Recall that the response variable is change in BMI; positive values indicate an increase in BMI and negative values indicate a decrease in BMI from the beginning to the end of the study.)

D5. Using the dataframe, produce a histogram separated by diet. Calculate the five number summaries and statistics to compute the ANOVA.

```
# diet %>% ggplot(xxx
#
# diet %>% xxx
# gmean <- xxx
```

D6. Calculate the F-statistic and p-value.

```
# VBG <-  XXX
# VWG <-  XXX
# fstat <- VBG / VWG
# c(VBG, VWG, fstat)
#
# 1-pf(fstat, XXX, XXX)
#
# aov.bmi <- diet %>% aov(xxx)
# summary(aov.bmi)
```

D7. Calculate the 95% CI for the following differences:

Atkins diet vs. Zone diet; Atkins diet vs. Ornish diet; Atkins diet vs. LEARN diet; Zone diet vs. Ornish diet;

```
# n=xxx
#
# xbar_Atkins =
# xbar_LEARN =
# xbar_Ornish =
# xbar_Zone =
# s_Atkins =
# s_LEARN =
# s_Ornish =
# s_Zone =
# n_Atkins =
# n_LEARN =
# n_Ornish =
# n_Zone =
#
# null = 0
# multiplier = qt(xxx, xxx)
#
# sd_A_Z = sqrt(s_Atkins^2/n_Atkins+s_Zone^2/n_Zone)
# sd_A_O = xxx
# sd_A_L = xxx
# sd_Z_O = xxx
#
# statistic_AZ = xbar_Atkins-xbar_Zone
# statistic_AO = xbar_Atkins-xbar_Ornish
# statistic_AL = xbar_Atkins-xbar_LEARN
# statistic_ZO = xbar_Zone-xbar_Ornish
#
# t_AZ = (statistic_AZ-null)/sd_A_Z
# t_AO = (statistic_AO-null)/sd_A_O
# t_AL = (statistic_AL-null)/sd_A_L
# t_ZO = (statistic_ZO-null)/sd_Z_O
#
# se_AZ = sd_A_Z
# se_AO = sd_A_O
# se_AL = sd_A_L
# se_ZO = sd_Z_O
#
# CI_AZ = c(statistic_AZ - multiplier*se_AZ, statistic_AZ + multiplier*se_AZ)
# CI_AO = c(statistic_AO - multiplier*se_AO, statistic_AO + multiplier*se_AO)
# CI_AL = c(statistic_AL - multiplier*se_AL, statistic_AL + multiplier*se_AL)
# CI_ZO = c(statistic_ZO - multiplier*se_ZO, statistic_ZO + multiplier*se_ZO)
#
# CI_AZ
# CI_AO
# CI_AL
# CI_ZO
```

D8. Which of the confidence intervals calculated in #17 indicate a significant difference between groups? How are you deciding? Based on your analysis, would you conclude that there is one diet that works significantly better than the other three? If so, which one?

D9. Do the sample data for the diet study appear to satisfy the validity conditions for conducting an ANOVA F-test? How are you deciding?