Assignment 8 - Comparing Two Means. Due November 16, 11:59pm 2018

EPIB607 - Inferential Statistics[a]

[a]Fall 2018, McGill University

This version was compiled on November 8, 2018

In this assignment you will practice comparing two means in a regression framework. Be sure to state the regression model in terms of parameters first. Use regression functions to fit these models. Answers should be given in full sentences (DO NOT just provide the number). All figures should have appropriately labeled axes, titles and captions (if necessary). Units for means and CIs should be provided. All graphs and calculations are to be completed in an R Markdown document. Please submit both the compiled HTML report and the source file (.Rmd) to myCourses by November 16, 2018, 11:59pm. Both HTML and .Rmd files should be saved as 'IDnumber_LastName_FirstName_EPIB607_A8'.

Two sample mean | Regression | Indicator variable | Bootstrap

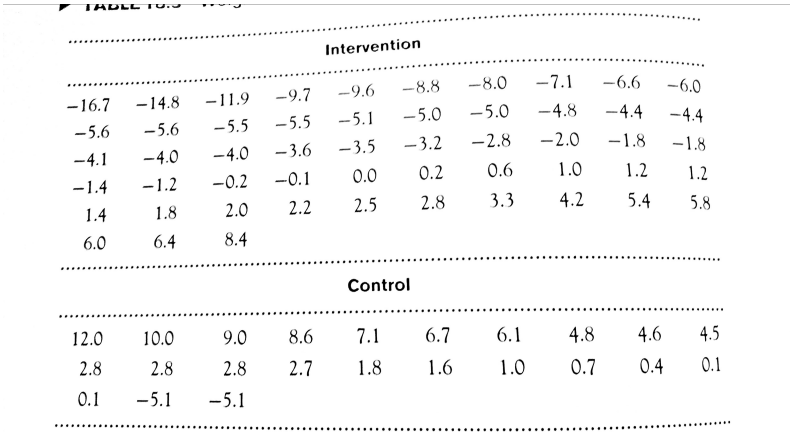
Template

There is no template for this assignment. You may use the same template from previous assignments.

1. Food counseling for obese children

According to the Centers for Disease Control and Prevention (CDC), obesity rates among U.S. children have increased dramatically over the past three decades, from a low of about 5% to 18% as of 2010. A study examined the effect of family-based food-counseling sessions provided by trained professionals. The study randomly assigned obese children aged 9 to 12 years to either the counseling intervention or a control group not receiving any food counseling. The children's weight changes (in pounds, lb) after 15 weeks is displayed in Table 1.

a. Plot the data.
b. We are interested in knowing if obese children receiving food counseling gain less weight over a 15-week period. State the regression model as a function of population parameters. Be sure to define what each parameter represents.
c. Use a regression procedure to answer this question. State the hypotheses, interpret the parameter of interest and a brief statement about your findings.
d. Show how the `Std. Error` was calculated for the parameter of interest.



Intervention									
−16.7	−14.8	−11.9	−9.7	−9.6	−8.8	−8.0	−7.1	−6.6	−6.0
−5.6	−5.6	−5.5	−5.5	−5.1	−5.0	−5.0	−4.8	−4.4	−4.4
−4.1	−4.0	−4.0	−3.6	−3.5	−3.2	−2.8	−2.0	−1.8	−1.8
−1.4	−1.2	−0.2	−0.1	0.0	0.2	0.6	1.0	1.2	1.2
1.4	1.8	2.0	2.2	2.5	2.8	3.3	4.2	5.4	5.8
6.0	6.4	8.4							
Control									
12.0	10.0	9.0	8.6	7.1	6.7	6.1	4.8	4.6	4.5
2.8	2.8	2.8	2.7	1.8	1.6	1.0	0.7	0.4	0.1
0.1	−5.1	−5.1							

Fig. 1. Weight change (lb) of obese children, by group

2. Student drinking

A professor asked her sophomore students, "How many drinks do you typically have per session" (A drink is defined as one 12-ounce beer, one 4-ounce glass of wine, or one 1-ounce shot of liquor.) Some of the students didn't drink. Table 2 gives the responses of the female and male students who did drink. It is likely that some of the students exaggerated a bit. The sample is all students in one large sophomore-level class. The class is popular, so we are tentatively willing to regard its members as an SRS of sophomore students at this college.

a. Do a complete analysis that reports on a comparison of the drinking behavior of women and men.
b. Summarise your findings in a 140 character tweet.
c. (BONUS) Verify the constant variance assumption of your model.

Female students

2.5	9	1	3.5	2.5	3	1	3	3	3	3	2.5	2.5
5	3.5	5	1	2	1	7	3	7	4	4	6.5	4
3	6	5	3	8	6	6	3	6	8	3	4	7
4	5	3.5	4	2	1	5	5	3	3	6	4	2
7	7	7	5.5	3	2.5	10	5	4	9	8	1	6
2	5	2.5	3	4.5	9	5	4	4	3	4	6	7
4	5	1	5	3	4	10	7	3	4	4	4	4
2	1	2.5	2.5									

Male students

7	7.5	8	15	3	4	1	5	11	4.5	6	4	10
16	4	8	5	9	7	7	3	5	6.5	1	12	4
6	8	8	4.5	10.5	8	6	10	1	9	8	7	8
15	3	10	7	4	6	5	2	10	7	9	5	8
7	3	7	6	4	5	2	5	5.5	9	10	10	4
8	4	2	4	12.5	3	15	2	6	3	4	3	10
6	4.5	5										

Fig. 2. Drinks per session claimed by female and male students

3. Ink toxicity

The National Toxicology Program evaluates the toxicity of chemicals found in manufacturing, in consumer products, or in the environment after disposal. Toxicity is assessed through a battery of tests. Here are some results from a study of the toxicity of black newsprint ink in 7-week-old female rats. The rats' fur was locally clipped twice a week for 13 weeks. One group of rats received a dermal application of ink right after each clipping, and a control group of rats was left untreated. Table 3 shows the body weights (in grams) of female rats at the beginning of the study and at the end of the 13 weeks.

a. Verify that the two experimental groups are not significantly different at the beginning of the study
b. Is there good evidence that ink application impairs growth in female rats between 7 and 20 weeks of age? State your answer in terms of the % difference between the mean weight gains of the two populations. Provide a 95% confidence interval for the % difference between the mean weight gains of the two populations.
c. Show how the p-value in the `glm` output was calculated for the parameter of interest.

Control Group		Treatment Group	
Week 0	Week 13	Week 0	Week 13
111.2	191.6	107.3	187.0
105.4	191.2	116.7	189.5
110.8	210.7	112.2	179.2
105.6	185.2	103.4	172.2
106.1	195.0	113.2	178.7
104.4	188.3	110.6	180.9
114.0	188.4	110.6	188.3
115.1	195.6	100.5	188.9
109.2	204.6	106.3	183.1
111.3	195.7	112.5	184.5

Fig. 3. Body weights (g) at beginning and end of ink toxicity study