STAT 110: Homework 9 Solutions (20 pts)
Fall 2017

Note: You must attach a copy of the JMP output for both problems to obtain full credit.

1. Eleven people diagnosed as being dependent on caffeine were subjects in a study. During the study, they were barred from coffee, colas, and other substances containing caffeine. Instead, during the first time period, they were randomly assigned to take either (1) capsules containing their normal caffeine intake, or (2) placebo capsules with no caffeine. During the second time period, they were assigned to the other treatment group. The subjects did not know when they got each pill. Subjects were assessed on the Beck Depression Inventory, which is a psychological test that measures depression. Higher scores on the test mean the subject shows more symptoms of depression. Data from this study can be found in the file **Caffeine Depression.jmp**.

   The research question is as follows: *Does depriving caffeine-dependent people of caffeine cause them to become more depressed?*

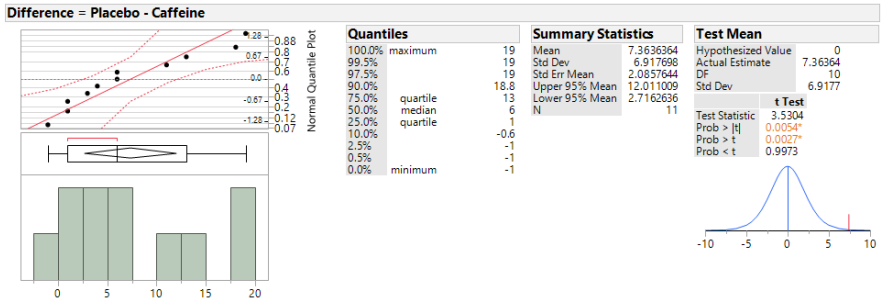
   a. Explain why a paired-test (for dependent samples) is appropriate for investigating this research question. (1 pt)

      Each subject was measured twice in this study. So, the observations in the two groups being compared (placebo vs. caffeine) are paired according to subject so that we can account for the variation in depression scores that exists from one subject to another in order to isolate the differences between the placebo and caffeine groups.

   b. Set up the null and alternative hypotheses to test the research question. (2 pts)

      Let $\mu_{\text{difference}} = \mu_{\text{placebo - caffeine}}$.
      H_0: **$\mu_{\text{difference}} = 0$**
      H_a: **$\mu_{\text{difference}} > 0$**

   c. Use a paired t-test to address the research question. Include in your solutions both the p-value and a conclusion in context of the problem. (3 pts)



   p-value: **0.0027**

   Conclusion: **The study provides evidence that depriving caffeine-dependent people of caffeine cause them to become more depressed.**

d. Are there any assumptions that must be met in order to use the paired t-test? If so, state these conditions and discuss whether these conditions have been met. (1 pt)

Yes. Either the sample size must be sufficiently large or the distribution of differences must be approximately normal. In this case, the sample size was only 11. The normal quantile plot, however, suggests that it is reasonable to assume the data come from a normal distribution; so, the t-test is OK to use.

e. Use JMP to find the 95% confidence interval for the mean difference between depression scores in the caffeine and placebo groups. Interpret this interval. (3 pts)

Lower endpoint: **2.7**

Upper endpoint: **12.0**

Interpretation: **We are 95% certain that, on average, depression scores are anywhere from 2.7 up to 12 points higher when a caffeine-dependent person is deprived of caffeine.**

2. Researchers count the number of breeding horseshoe crabs on beaches on Delaware Bay every year; the data from 2011 and 2012 are given in the file **HorseshoeCrabs.jmp**. Note that the number of horseshoe crabs found on each beach is given for both years.

   The research question is as follows: *Has the number of horseshoe crabs changed from 2011 to 2012?*

   a. Explain why a paired-test (for dependent samples) is appropriate for investigating this research question. (1 pt)

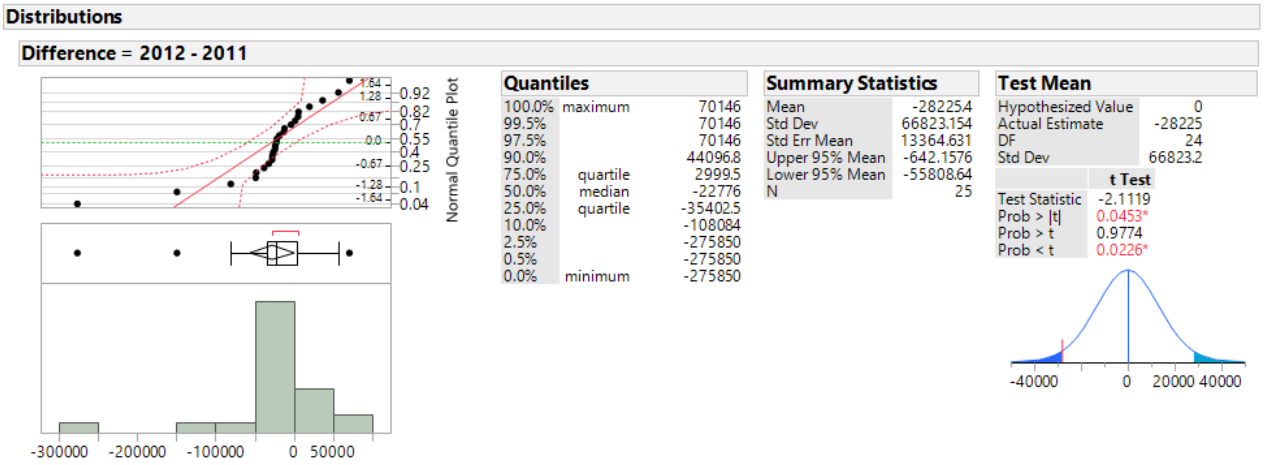
      Each beach was measured twice in this study. So, the observations in the two groups being compared (2011 vs. 2012) are paired according to beach so that we can account for the variation in the number of horseshoe crabs that exists from one beach to another in order to isolate the effect of 2011 vs. 2012.

   b. Set up the null and alternative hypotheses to test the research question. (2 pts)

      Let $\mu_{\text{difference}} = \mu_{\text{2012-2011}}$.

      H_0: $\mu_{\text{difference}} = 0$
      H_a: $\mu_{\text{difference}} \neq 0$

   c. Use an appropriate hypothesis testing procedure to address the research question. Include in your solutions both the p-value and a conclusion in context of the problem. (3 pts)



   p-value: **0.0453**

   Conclusion: **The data provide evidence that the number of horseshoe crabs has changed from 2011 to 2012.**

d. Are there any assumptions that must be met in order to use the hypothesis testing procedure you used in part c? If so, state these conditions and discuss whether these conditions have been met. (1 pt)

Yes. Either the sample size must be sufficiently large or the distribution of differences must be approximately normal. In this case, the sample size was only 25. The normal quantile plot also suggests that the original population might not be normal distributed; so, the results of the t-test might be suspect. It would be worth consulting a statistician to discuss alternative approaches.

e. Use JMP to find the 95% confidence interval for the mean difference between the number of horseshoe crabs found from 2011 to 2012. Interpret this interval. (3 pts)

Lower endpoint: **-55,808.6**

Upper endpoint: **-642.2**

Interpretation: **We can be 95% certain that there were anywhere from 642.2 to 55,808.6 fewer horseshoe crabs on Delaware Bay beaches, on average, in 2012 than in 2011.** ***Note that this confidence interval might also be suspect because the assumptions for using the t-distribution weren't met.***