Problem set 9

S520

Upload a PDF/HTML file/Word document with your answers and a .R/.Rmd file with your code through the Assignments tab on Canvas by 11:59 pm, Thursday 4th November.

Trosset question numbers refer to the hardcover textbook. You may work with others, but you must write up your homework independently — you should not have whole sentences in common with other students or other sources.

1. (5 points.) "Glycemic index" is a measure of how quickly blood sugar level rises after eating a particular food. (Glucose has a glycemic index of 100, while water has a glycemic index of 0.) A group of researchers wished to study glycemic index when dates and coffee were consumed together by individuals with type 2 diabetes. They performed a study on 10 subjects with diabetes. Firstly, they measured glycemic index for each patient after consuming dates without coffee. The mean was 53 with standard deviation 19. Then (several days later) they measured glycemic index for each patient after consuming dates with coffee. The mean was 41.5 with standard deviation 17. The differences between the measurements ("without coffee" minus "with coffee") had mean 11.5 with standard deviation 21. There were no outliers in the data.

   (a) Explain why this data should be analyzed using a one-sample t-test rather than a two-sample t-test.

   (b) Give null and alternative hypotheses for an appropriate two-tailed t-test, and calculate the t-statistic and P-value for the test.

   (c) Find a 95% confidence interval for the population mean difference between the date glycemic index measurements ("without coffee" minus "with coffee.")

   (d) The P-value you found was greater than 0.05, so the authors did not reject the null hypothesis. Based on this and the rest of analysis, is it correct to conclude that we are sure that on average, dates have the same glycemic index with or without coffee? Explain.

2. (5 points.) The heights of men and women are approximately normally distributed. Suppose we wish to estimate the average difference in heights between men and women attending a certain university. A random sample of seven men has average height 68.5 inches with standard deviation 3.0 inches, while a random sample of seven women has average height 65.5 inches with standard deviation 2.5 inches.

   (a) Explain why we would use the t-distribution and not the standard normal distribution to calculate P-values and confidence intervals in this case.

(b) According to Welch's approximation, the number of degrees of freedom is 11.6. Using the fact that the R code `qt(.975,df=11.62)` gives the value 2.187, find an approximate 95% confidence interval for the average difference in heights between men and women at the university.

(c) The confidence interval you calculated in part (b) contains the value 0. Should you conclude that there is no difference between the average height of men and the average height of women at the university? Explain why or why not.

3. (10 points.) Trosset chapter 11.4 Problem Set C, parts 1–3.

4. (10 points.) Trosset chapter 11.4 Problem Set D, parts 1–4.