Stat414; Fall 2024; Worksheet 10; 20 Points;

Module10 deals with a wellknown topic, but more in depth than you may have studied in your previous stat course (like Stta 350 or 351 or 355). The worksheet will require you to use the `tTestPower` function of EnvStats. Read the help file i n t he E nvStats c arefully–will h elp y ou in completing the workseet.

1. we use the two-sample t-test to compare sulfate concentrations (`EPA.09.Ex.16.1.sulfate.df`) at a background and downgradient well. The resulting t-statistic is 5.66 with 12 degrees of freedom.

   (a) Plot the pdf of a t-distribution with 12 degrees of freedom and add a vertical line at x = 5.66.

   (b) Explain what part of this plot represents the p-value for the test of the null hypothesis that the average sulfate concentrations at the two wells are the same against the alternative hypothesis that the average concentration of sulfate at the downgradient well is larger than the average concentration at the back-ground well.

2. Consider the copper concentrations stored in the data frame `EPA.09.Ex.16.4.copper.df` in EnvStats package. Use the t-test and Wilcoxon rank sum test to compare the data from the two background wells.

3. The following data shows age at diagnosis of Type II diabetes among young adults. Is the age at diagnosis different for males and females.

Males	19, 22,16,29,24
Females	20,11,17,12

   (a) What test procedures are applicable to this data structure? Write down the assumptions of for each of the candidate procedure.

   (b) Which procedure would you recommend and why?

   (c) Regardless of your recommendation above, apply both the parametric and nonparametric methods to this example and compare the results.