ST511 HW #4

Reading: Read Chapters 5 and 6 of Ott & Longnecker.
See Canvas Calendar for due date.
40 points total, 2 points per problem part unless otherwise noted.

1. A national agency sets recommended daily allowances for many supplements. In particular, the allowance for zinc for adult men is 15 mg/day. The agency would like to determine if the average intake of zinc for adult men is greater than 15 mg/day. Suppose from a previous study they estimate the standard deviation to be 3 mg/day and they conjecture that the true population mean is 17 mg/day. The investigators plan to use a one-sample t-test with α=0.05.
    A. Find the power with n=12 for the scenario above.
    B. If the standard deviation was larger (more than 3) would the power be higher or lower than that calculated in part A?
    C. If the sample size was larger (more than 12) would the power be higher or lower than that calculated in part A?
    D. If we used α=0.10 (instead of 0.05), would the power be higher or lower than that calculated in part A?
    E. Using a conjectured mean of 16 mg/day (instead of 17), would the power be higher or lower than that calculated in part A?
    F. Return to the original scenario and find the sample size required to achieve 90% power. Remember to "round" up to an integer value.

2. Use the data from Problem 5.29, which deals with lead concentrations in estuarine creeks.
    A. Construct a histogram, qqplot and run SW test of normality. What do you conclude about the normality of the data? Do the various plots and tests agree? (**4 pts**)
    B. Give the sample mean and median for this data.
    C. Use the sign test to test the null hypothesis that the median is equal to 30. Give the p-value and make a conclusion.
    D. Give a 95% confidence interval for the median. Note: For consistency, please report the "Upper Achieved CI".
    E. Using the standard one-sample t-test, test the null hypothesis that the mean is equal to 30. Give the p-value and conclusion.
    F. Give a (standard) 95% confidence interval for the mean.
    G. It should be clear from the diagnostics in part A that the assumption of normality is not met. Hence the test and CI from parts E and F are questionable. Give a 95% bootstrap studentized confidence interval for the mean. Hint: See "boot example2", but use a different value for set.seed.
    H. Assuming that cumulative lead exposure is of interest, would the mean or the median be of more interest.

3. Read problem 6.60 which concerns flare of two mixtures of rocket propellant.
   A. Construct side-by-side boxplots.
   B. Give the sample means and standard deviations for each mixture.
   C. Assuming equal variances, give the 95% confidence interval for the difference between the means. Based on this interval, can we conclude that there is a difference between the population means? Explain.
   D. Considering the summary statistics in part B, is the pooled variance t-test or Welch-Satterthwaite t-test appropriate here? Justify your response.
   E. Regardless of your answer from part D, run the pooled t-test to test H0: H_0:$\mu_1 - \mu_2 = 0$ versus a two-sided alternative. Give the p-value and conclusion.