STAT 515 hw 7

CIs for mean with σ unknown, sample size calculations

Attach a sheet with the R plots and R code printed on it. You may write out your other answers by hand if you want. Just try to make it easy for me grade!!

1. Open R and enter `data(Loblolly)` into the console. This imports the Loblolly data set into the workspace. Type `?Loblolly` into the console to read a description of the data set.
   (a) On how many trees was data collected?
   (b) How many times was the height of each tree recorded?
   (c) At what ages was the height of each tree recorded?
   (d) Compute the mean $\bar{X}_n$ and the sample standard deviation s for the heights of Loblolly pines which are 3 years old. Hint: Enter the command

   ```
   x <- Loblolly$height[Loblolly$age==3]
   ```

   Then to compute the mean $\bar{X}$, you can simply type `mean(x)` and for the standard deviation, you can type `sd(x)`.
   (e) Generate a Normal QQ plot of the heights of the Loblolly pines at age 3. Turn in this plot. Use `qqnorm(x)`.
   (f) Based on the QQ plot, do you think that the heights follow a Normal distribution?
   (g) Compute a 95% confidence interval for the mean height of three-year-old Loblolly pines.
   (h) Interpret this interval.
   (i) Give a 95% percent confidence interval for the mean height of twenty-year-old Loblolly pines.
   (j) If you had constructed 99% confidence intervals for the Loblolly heights, would they have been wider or narrow than the 95% confidence intervals?
   (k) You plan to estimate the mean height of 3-year-old Loblolly pines in a different region of North America, and you need to know how many trees to measure. Give a recommended sample size if you want
      i. a 95% confidence interval no wider than 0.25 feet.
      ii. a 99% confidence interval with margin of error no greater than 0.10.

2. Make a 95% confidence interval for the variance σ^2 of the heights of Loblolly trees which are three years old in the following steps:
   (a) Compute S_n^2.
   (b) Find the degrees of freedom of the relevant Chi-square distribution.
   (c) Find $\chi^2_{\nu,1-\alpha/2}$ and $\chi^2_{\nu,\alpha/2}$, where ν is your answer to part (b).
   (d) Compute the confidence interval.

3. You wish to estimate the proportion of bees in a beehive that are drones within 0.02 with confidence level 95%. A sample of 307 bees from a previous hive contained 44 drones.

(a) How many bees should you sample?

(b) If you ignore the data from the previous hive, how many bees would you recommend sampling?

Optional (do not turn in) problems for additional study from McClave, J.T. and Sincich T. (2017) *Statistics*, 13th Edition: 7.38, 7.40, 7.50