Long Problem Set 3

For each problem below, complete any requested calculations and answer any accompanying questions. Your responses are evaluated on the appropriateness of your approach and the insightfulness of your analysis. Be sure to consider significant figures when interpreting the output from R (which, as with a calculator, often ignores such niceties).

You will share your answers to these problems in two ways: a text document that contains your written responses to the questions, and a `.R` script file that contains your code (use comments to separate your code by problem). Save your script file using `yourlastname_LPS03.R` as a file name and share it with by email. You may use any program you wish for your text document.

Your answers to the following questions are due by 4:00 pm on Tuesday, September 18th.

1. The data file "BloodPressure.RData" contains results from a clinical study of the efficacy of calcium supplements as a treatment for blood pressure in males. In this study 21 subjects received either a calcium supplement ("calcium.yes") or a placebo that contained no calcium ("calcium.no") for 12 weeks. Each subject's blood pressure was measured before and after the treatment period and the difference recorded (a positive value represents a decrease in blood pressure). Determine at $\alpha = 0.05$ whether there is any evidence that calcium lowers blood pressure.

2. The file "SpeedLight.RData" contains two sets of results from Michelson's 1879 determination of the speed of light ("sol.one" and "sol.two"). To make it easier to enter the data into a file, all values are offset by subtracting 299,000 km/sec from the measured value. Determine at $\alpha = 0.05$ if there is a significant difference between the results of these two experiments.

3. The data at this link (http://www.rsc.org/images/CO2_methods_tcm18-57755.txt) reports results for the determination of CO_2 by six different methods. The data itself uses Na_2CO_3 as a reference standard; presumably, a portion of the standard was treated to release the CO_2, which subsequently was determined and reported as % w/w CO_2 in the sample. Using the data for the gravimetric method, determine at $\alpha = 0.05$ if there is any evidence for a determinate error in the analysis. Note that the authors report a known value of 41.518% w/w CO_2 for the reference standard.

4. The file "Clouds.RData" contains results for the amount of rainfall recorded from 26 clouds, half of which were randomly seeded with AgI (the units are in acre-feet, or a volume equivalent to the feet of rain covering one acre of ground). For each data set, "seeded" and "unseeded", make a convincing argument that the data are not normally distributed and then, using an appropriate statistical test, determine at $\alpha = 0.05$ if seeding clouds has any effect. By the way, the writer Kurt Vonnegut's brother, Bernard, was an atmospheric scientest at General Electric who discovered that AgI could be used to seed clouds.

5. Sometimes it is possible to transform a strongly right-skewed distribution by using a log function. In R you can take the log of an object using the command log10(object). Transform the data from Problem 4 and show that the data are now normally distributed. Repeat your significance test using a test appropriate for normally distributed data and compare this result to your results from Problem 4.