HW week 8

w203: Statistics for Data Science

The file GPA1.RData contains data from a 1994 survey of MSU students. The survey was conducted by Christopher Lemmon, a former MSU undergraduate, and provided by Wooldridge.

```
load("GPA1.RData")
```

The skipped variable represents the average number of lectures each respondent skips per week. You are interested in testing whether MSU students skip over 1 lecture per week on the average.

a. Examine the skipped variable and argue whether or not a t-test is valid for this scenario.

b. How would your answer to part a change if Mr. Lemmon selected dormitory rooms at random, then interviewed all occupants in the rooms he selected?

c. Provide an argument for why you should choose a 2-tailed test in this instance, even if you are hoping to demonstrate that MSU students skip more than 1 lecture per week.

d. Conduct the t-test using the t.test function and interpret every component of the results.

e. Show how you would compute the t-statistic and p-value manually (without using t.test), using the pt function in R.

f. Construct a 99% confidence interval for the mean number classes skipped by MSU students in a week.

g. Can you say that there is a 99% chance the population mean falls inside your confidence interval?