

Instructions for this homework

Please ensure that you include necessary information for any hypothesis tests. The appropriate formulation and statements about statistical inference are at least as important as correct calculations. You may use Python or hand-calculate values if you choose, but in either case please show your work. At a minimum, your solutions should include the following information:

- Executive summary-style synopsis of results (plainly worded short sentence or paragraph)
- Null and Alternative Hypotheses
- Type of test being performed and why (t , z , Wilcoxon Signed Rank Sum, Mann-Whitney U, etc.). Justify any assumptions.
- Level of significance, justified if not given
- Calculation of test statistic, critical value, and/or p-value as appropriate
- Test statistic comparison or p-value comparison leading to your conclusion.
- Correct and full conclusion statement to include appropriate language (i.e., reject or fail to reject etc.)

Problem 1: 40 points*You batter bell-lieve it*

Use the file **BattingAverages.csv**, containing batting averages for all players with at least 100 at bats for the 2009 season, for the following questions. Assume this is a random sample rather than a census.

(a)

Are the batting averages data (**BattingAvg**) approximately normally distributed? Use both graphical and analytical methods to make your argument.

(b)

Is the mean value of batting averages greater than .265? Perform a test to find out. Use $\alpha = 0.05$ for your test.

(c)

Is there a difference between batting averages in the National League and American League (column **League**)? Use $\alpha = 0.05$ for your test.

Problem 2: 60 points*Cowbell usage must have increased...*

A researcher studying true body temperature in adult humans collected the data in **BodyTemp.csv** in degrees Fahrenheit.

(a)

Is body temperature approximately normally distributed? Use graphical and analytical methods to make your argument.

(b)

Is the mean body temperature equal to 98.6°F?

(c)

For the α you selected, what is the power to detect a difference of 0.2°F? Assume the population variance is equal to the sample variance.

Remember that effect size is difference to detect divided by standard deviation. See the progress check for examples.

(d)

Create a plot showing how α (x-axis) affects the power to detect a difference of 0.2°F (y-axis).

(e)

Is there a difference between body temperature in males and females?