Name: Mateusz Romaniuk
Instructor: Mahsa Ghasemi
GitHub Username: RomanczuG
Purdue Username: mromaniu

Problem 1

1. Suppose the instructor of the course is convinced that the mean engagement of students who become knowledgeable in the material (i.e., the engagement_1 population) is 0.75.

    a. *[5 points]* Formulate null and alternative hypotheses for a statistical test that seeks to challenge this belief. What are the null and alternative hypotheses?

    ANSWER:
    Null Hypothesis H_0: The mean engagement of students is equal to professors predictions**,** $\mu = 0.75$

    Alternative Hypothesis H_1: The mean engagement of students is not equal to professors predictions, $\mu != 0.75$

    b. *[5 points]* What type of test should be used and why?

    ANSWER: We should perform hypothesis testing because it lets us reject the null hypothesis

2. Carry out the statistical test defined in (1b) using the *`engagement_1`* sample.

    a. *[1 point]* What is the sample size?

    ANSWER: 937

    b. *[1 point]* What is the sample mean?

    ANSWER: 0.7430304110448239

    c. *[2 points]* What is the standard error?

    ANSWER: 0.004153027288269652

    d. *[2 points]* What is the standard score?

    ANSWER: -1.6781948375012814

    e. *[2 points]* What is the p-value?

    ANSWER: 0.09330906925243751

f. *[2 points]* Are the results statistically significant at a level of 0.05? How about 0.10? What (if anything) can we conclude (i.e., what is the interpretation of the result)?

ANSWER: 90% confidence interval, reject the null hypothesis
95% confidence interval, accept the null hypothesis
99% confidence interval, accept the null hypothesis

3. *[10 points]* What is the largest standard error for which the test will be significant at a level of 0.05? What is the corresponding minimum sample size? (You may assume that the population variance and mean does not change.)

ANSWER:
Standard error: 0.004237209220916103
Minimum sample size: 900.1385114802096

4. Suppose the instructor is also convinced that the mean engagement is different between students who become knowledgeable (the engagement_1 population) and those who do not (the engagement_0 population).

a. *[5 points]* Formulate null and alternative hypotheses that seek to validate this belief. What are the null and alternative hypotheses?

ANSWER:

Null Hypothesis H_0: The mean engagement of students is equal to the second sample.

Alternative Hypothesis H_1: The mean engagement of students is not equal to the second sample.

b. *[5 points]* What type of test should be used and why?

ANSWER: We should perform hypothesis testing because it lets us reject the null hypothesis

5. Carry out the statistical test defined in (4b) using the *`engagement_1`* and *`engagement_2`* samples.

a. *[1 point]* What are the sample sizes?

ANSWER: 1977

b. *[1 point]* What are the sample means?

ANSWER: 0.6399545077035914

c. *[2 points]* What is the standard error?

ANSWER: 0.005715989588773277

d. *[2 points]* What is the standard score?

ANSWER: -19.252220562568542

e. *[2 points]* What is the p-value?

ANSWER: 1.352401028607762e-82

f. *[2 points]* Are the results statistically significant at a level of 0.05? How about 0.10? What (if anything) can we conclude (i.e., what is the interpretation of the result)?

ANSWER:
Reject the null hypothesis for 90% confidence interval
Reject the null hypothesis for 95% confidence interval
Reject the null hypothesis for 99% confidence interval

Problem 2

1. Use the sample to construct a 90% confidence interval for the number of points by which the team wins on average.

    a. *[3 points]* Will you use a t-test or z-test (Hint: Think which distribution should you use here if very few data points are available)? Justify your answer.

    ANSWER: I will perform the t-test.

    b. *[3 points]* What is the sample mean?

    ANSWER: 7.363636363636363

    c. *[3 points]* What is the standard error?

    ANSWER: 5.0762776757504415

    d. *[3 points]* What is the standard statistic (t or z value)?

    ANSWER: 1.4407478926091581

    e. *[3 points]* What is the 90% confidence interval?

    ANSWER: from 0.049999999999998934 to 14.677272727272728

2. Repeat Q1 for a 95% confidence interval.

    a. *[2 points]* What is the standard statistic (t or z value)?

    ANSWER: 1.4456727611047144

    b. *[2 points]* What is the 95% confidence interval?

    ANSWER: from 0.025000000000000355 to 14.702272727272726

    c. *[1 point]* Is your interval wider or narrower compared to using the 90% confidence interval in Q1?

    ANSWER: The interval is wider

3. Repeat Q2 if you are told that the population standard deviation is 15.836.

    a. *[5 points]* Will you use a t-test or z-test (Hint: Think which distribution should you use here now that you have the true population standard deviation)? Justify your answer.

ANSWER: Perform z-test

b. *[3 points]* What is the standard error?

ANSWER: 4.774733652733465

c. *[3 points]* What is the standard statistic (t or z value)?

ANSWER: 1.385131997855033

[3 points] What is the 95% confidence interval?

ANSWER: 0.16601210933385924

d. *[6 points]* Is your interval wider or narrower than the interval computed in Q2?

ANSWER: The range is narrower.

4. *[10 points]* Assume you no longer know the population standard deviation. With what level of confidence can we say that the team is expected to win on average? (Hint: What level of confidence would you get a confidence interval with the lower endpoint being 0?)

ANSWER: Level of confidence is 0.9229747134327154.