Name: Yolanda Chen

Instructor: Mahsa Ghasemi

GitHub Username: codery12321

Purdue Username: chen3633

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.
   1. *[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: mean engagement is 0.75

Alternative: mean engagement is not 0.75

* 1. *[5 points]* What type of test should be used and why?

**ANSWER**: Z test should be used because the sample size is greater than 50**.**

1. Carry out the statistical test defined in (1b) using the *`engagement\_1`* sample.
   1. *[1 point]* What is the sample size?

**ANSWER:** 937

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

**ANSWER:** ﻿0.7430304110448239

* 1. *[2 points]* What is the standard error?

**ANSWER:** ﻿ 0.004153027288269652

* 1. *[2 points]* What is the standard score?

**ANSWER:** ﻿-179.8481258566727

* 1. *[2 points]* What is the p-value?

**ANSWER:** ﻿0.09364288349711912

* 1. *[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:** Not statistically significant at 0.05 but is statistically significant at 0.1. We can reject the null hypothesis because p-value is > 0.05.

1. *[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:** ﻿0.003555978074164272

1. 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).
   1. *[5 points]* Formulate null and alternative hypotheses that seek to validate this belief. What are the null and alternative hypotheses?

**ANSWER:** Null hypothesis: Mean engagement is the same.

Alternative hypothesis: mean engagement is different.

* 1. *[5 points]* What type of test should be used and why?

**ANSWER**: z test should be used because sample size is larger than 50

1. Carry out the statistical test defined in (4b) using the *`engagement\_1`* and *`engagement\_2`* samples.
   1. *[1 point]* What are the sample sizes?

**ANSWER:** ﻿937, 1977

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

**ANSWER:** ﻿0.7430304110448239, ﻿0.6399545077035914

* 1. *[2 points]* What is the standard error?

**ANSWER:** ﻿0.004153027288269652, ﻿0.007065420910043284

* 1. *[2 points]* What is the standard score?

**ANSWER:** ﻿-179.8481258566727, ﻿-14.588784540028351

* 1. *[2 points]* What is the p-value?

**ANSWER:** ﻿3.3104307168195455e-48

* 1. *[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:** No the results are not statistically significant at a level of 0.05 nor at 0.1. We can reject the null hypothesis and conclude that the mean engagement is different.

Problem 2

1. Use the sample to construct a 90% confidence interval for the number of points by which the team wins on average.
   1. *[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:** T test should be used

* 1. *[3 points]* What is the sample mean?

**ANSWER:** ﻿7.363636363636363

* 1. *[3 points]* What is the standard error?

**ANSWER:** ﻿5.0762776757504415

* 1. *[3 points]* What is the standard statistic (t or z value)?

**ANSWER:** ﻿1.8124611228107335

* 1. *[3 points]* What is the 90% confidence interval?

**ANSWER:** [﻿ -1.8369195722533416, ﻿16.56419229952607]

1. Repeat Q1 for a 95% confidence interval.
   1. *[2 points]* What is the standard statistic (t or z value)?

**ANSWER:** ﻿2.2281388519649385

* 1. *[2 points]* What is the 95% confidence interval?

**ANSWER:** [ ﻿-3.9470151490654715, ﻿18.674287876338198]

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

**ANSWER:** wider

1. Repeat Q2 if you are told that the population standard deviation is 15.836.
   1. *[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**: z test should be used because the sample size is smaller than 50 and we know the true population standard deviation.

* 1. *[3 points]* What is the standard error?

**ANSWER:** ﻿1.4396363636363636

* 1. *[3 points]* What is the standard statistic (t or z value)?

**ANSWER:** ﻿1.96

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

**ANSWER:** [ ﻿ -1.994841595721228, ﻿16.722114322993953]

* 1. *[6 points]* Is your interval wider or narrower than the interval computed in Q2?

**ANSWER:** narrower

1. *[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:** ﻿91.12376063813866