ENGRD 2700, Basic Engineering Probability and Statistics, Fall 2019
Homework 8

Due **Friday December 6** at 11:59 pm. Submit to Gradescope by clicking the name of the assignment. See `https://people.orie.cornell.edu/yudong.chen/engrd2700_2019fa.html#homework` for detailed submission instructions.

The same stipulations from Homework 1 (e.g., independent work, computer code, etc.) still apply.

1. For some reason, Harry has kept meticulous records of Harry's muffins, and found that they usually contain a number of blueberries that is normally distributed with mean $\mu_0 = 25$ and standard deviation $\sigma = 4$. However, recently Jimmy suspects that Harry has been adding more blueberries than usual to his muffins. The last 12 muffins Jimmy bought contained the following numbers of blueberries:

   33 21 28 25 24 31 17 31 29 30 29 31

   Jimmy is considerate, and decides that he doesn't want to alarm Harry unless his findings are significant at the $\alpha = 0.05$ level.

   (a) Write down the (one-sided) hypothesis test being conducted above.

   (b) Treating σ^2 as the true variance of the distribution, compute the test statistic associated with the above data.

   (c) Does Jimmy reject H_0, the null hypothesis?

   (d) If your answer to part (c) is "yes", how small would the sample mean $\bar{x}_{12}$ need to be for Jimmy not to reject H_0? Alternatively, if your answer to part (c) is "no", how large would $\bar{x}_{12}$ need to be for Jimmy to reject H_0?

   (e) Repeat parts (b) and (c), but assume this time that Jimmy has no idea what σ^2 is.

2. Harriet suspects that the quarter in her pocket may not be a fair coin. She flips it 50 times, and to conduct a two-sided hypothesis test at the $\alpha = 0.025$ significance level. Heads appears $h = 28$ times.

   (a) Does Harriet reject H_0, the null hypothesis that the coin is fair, for a 2-sided test?

   (b) For what values of h would Harriet reject H_0?

3. According to the CDC, 17% of school-age children in the United States are obese, while 33.8% of adults in the U.S. are obese (having a Body Mass Index, or BMI, of at least 30).

   (a) In 2005, the Health Department in Marion County, Indiana measured the heights and weights of 90,147 school-age children, allowing exact determination of their BMIs. Among the children participating in the study, 22% were considered obese. Does this indicate that the true obesity rate for children in Marion County is different from the national average? Conduct a two-sided hypothesis test.

   (b) The Marion County Health Department simultaneously conducted a telephone survey of 4784 adults. 25% of participants reported as being obese. Does this indicate that the true adult obesity rate in Marion County is <u>lower than</u> the national average? Conduct a <u>one-sided</u> hypothesis test.

   (c) What are the potential issues with the study above?

4. Consider once again the temperature data `ithaca.csv` and `syracuse.csv` from Homework 7. We want to conduct the hypothesis test

   $$H_0 : \mu_i = \mu_s \qquad\qquad H_1 : \mu_i \neq \mu_s$$

   at the $\alpha = 0.05$ significance level, where μ_i and μ_s denote the mean temperatures in both cities during the month of March. Attach your code for the following questions.

   (a) If we make the (unrealistic) assumption that the two samples are independent, do we reject H_0?

(b) Repeat part (a), but relax the assumption that the two cities are independent. That is, we build a hypothesis test for paired, dependent data. (Now that this is only possible when the two samples contain the same number of observations for the same dates.)

5. A certain NBA player had a field goal percentage (i.e., probability of making a shot) of $p_0 = 60\%$ before needing to take a season off to recover from an injury.

   (a) Since returning to the game from injury, the player has made 13 out of $n = 20$ shots. Is the player's new, post-injury field goal percentage higher than his old percentage p_0? Perform a suitable one-sided hypothesis test and state your conclusion, taking $\alpha = 0.05$.

   (b) Suppose that the true new field goal percentage is p, where $p \in (0.6, 1)$. If we perform a one-sided test as above and want to achieve type-I error rate of 0.05 and type-II error rate of 0.025, what is the number of shots n needed since returning from injury? Provide an approximate formula as a function of p, and compute the values of n for each of $p = 0.8, 0.7, 0.61$ (Notice that if p is very close to 0.6 then you may need a very large number of shots.)