Assignment #2 (I2)
Basic Statistics
Individual Assignment

1. You work at a popular pizzeria in suburban Rochester that is considering implementing a delivery time guarantee (e.g., "Guaranteed delivery within 50 minutes, or your pizza is free!"). However, you are concerned about the potential number of pizzas that could be given away for free under this policy. Your shift manager *asserts* that delivery times are uniformly distributed between 30 minutes and an hour, but you have reservations about this claim. Fortunately, you have recorded the actual delivery times for the past several years and compiled a dataset of 1,400 observed actual delivery times (DeliveryTimes.xlsx).

    a. Using the included dataset, compute the sample average, the sample standard deviation and the sample variance of delivery times.

    b. Based on your findings from part (a), are these results consistent with the delivery times being uniformly distributed between 30 and 60 minutes? Why or why not?
    Note: A formal test is not required here; however, just an informal discussion is sufficient (for which the formulas for the mean and variance of a uniform distribution that you can look up on Wikipedia will be useful).

    c. Suppose you assume instead that delivery times are normally distributed (rather than uniform) with the mean and standard deviation you found in part (a). Using the estimates from part (a), construct an interval (cut-off times) into which you expect 95% of delivery times to fall. Repeat this exercise but replace 95% with 80%.
    How do these intervals compare with what you would have concluded had you just *assumed* that delivery times were uniformly distributed between 30 and 60 minutes? A formal comparison is required here (i.e., construct intervals under the normal and uniform distribution assumptions and discuss).
    Note: This question is not asking students to construct a *confidence interval* for the *population mean*. This question is asking students to construct an interval in *the normal (and uniform) distribution.* Recall the questions Q1(e)-(g) in Assignment 1?
    Note: When constructing the interval (cut-off times), assign the probability (95% or 80%) at the center of the distribution.

    d. Continuing to assume a normal distribution, what is the probability that a given pizza is delivered in 45 minutes or less? How about 40 minutes or less? 50 minutes or more? Compare these answers to what you would conclude under the uniform assumption (i.e., repeat this exercise with a uniform distribution).

    e. If you were to implement a policy of only charging for pizzas that are delivered in under 50 minutes, would it matter which distribution was the correct one? Why or why not? (Note: you do not need to compute anything here, just answer the question from an intuitive standpoint).

f.   Again using the information from part (a), construct a 95% *confidence interval* for the population mean. How does this compare to the 95% interval you constructed in part (c) for the Normal case? If it is different, what is the reason for this?

g.   Would the *confidence interval* you constructed in part (f) change if you assumed the population distribution was Uniform instead of Normal? Why or why not?

h.   Using the information from part (a), test the null hypothesis that the population average delivery time is 50 minutes against the two-sided alternative hypothesis that it is not 50 minutes at both the 1% and 5% levels. What do you conclude?

Submission Guidelines for Assignments:

Please submit your completed assignment as a PDF document, naming it "YOURLASTNAME_I2.pdf." Use the assignment submission link on Blackboard to upload your file.

(1) **Word Document**: You may create your assignment in a Word document, including all answers and formulas. Please save this document as a PDF file and submit it.
(2) **Python Notebook**: Alternatively, you can complete your assignment in a Jupyter Notebook. After executing all cell blocks, your notebook will display your code, formulas, comments, as well as outputs. To convert your notebook to a PDF file, navigate to *File → Print Preview → Right Click → Print → Save as PDF* (This is just one way of creating the PDF file - there can be other ways, too). TAs will not be running your code, so please ensure that the PDF file *clearly displays your final answers*, not just the code and formulas.
(3) **Excel**: You also have the option to use Excel. Type your formulas as "text" within the cells. When saving your Excel file as a PDF, ensure all content is visible by setting the print area under *Page Layout → Page Setup → Print Area → Set Print Area* or adjusting the scale at *Page Layout → Scale to Fit → Scale*.
(4) **Note:** Handwritten submissions will not be accepted as they can be difficult for TAs to read.

Grading:

If you provide only your final answer and it is incorrect, we cannot award any points. However, if you detail your calculation process, we can follow your reasoning and may award partial credit even if your final answer is incorrect. Therefore, please show all steps in your calculations.

Others:

For this assignment, you may use Python, Excel, or a Normal CDF table. The values obtained from a Normal CDF table might slightly differ from those in Excel or Python due to rounding errors in the table. Our grading TAs are aware of this, and you don't have to worry.