ST 228: Data Analysis, ML and AI

Assignment # 1 (10 marks)

(Central Limit Theorem and Hypothesis Testing)

10th of Jan 2025

Due on: 15th **of August 2025 before 5 PM**

Instructions :

- Clearly state all the assumptions made.
- Clearly quote the source of data used.
- Clearly show your work

1. Generate the following distributions with specified parameters:
    i. A uniform distribution with a known population mean and variance.
    ii. A normal distribution with a known population mean and variance.
    iii. An exponential distribution with a known population mean and variance.

   For each distribution:

   a. Demonstrate the Central Limit Theorem by showing how the sampling distribution of the mean approaches normality as sample size increases.
   b. Investigate how varying the population size and sample size influences the sample mean and sample variance. Focus on a single sample scenario to illustrate the effects.
   c. Examine how the skewness of the original population distribution affects the conclusions drawn from the Central Limit Theorem and the behavior of the sample mean and variance. (**5 marks)**

2. The shelf life of a beverage is of interest. Ten bottles are randomly selected and tested, and the following results are obtained:

Days	
108	138
124	163
124	159
106	134
115	139

   a. We would like to demonstrate that the mean shelf life exceeds 120 days. Set up appropriate hypotheses for investigating this claim.
   b. Test these hypotheses using *alpha* = 0.01. What are your conclusions?
   c. Find the *P*-value for the test in part (b).
   d. Construct a 99 percent confidence interval on the mean shelf life.
   e. Explain Type I and Type II error **(5 marks)**