DSL201: Statistical Programming

Dr. Anil Kumar Sao

Assignment 3

Deadline: 20/03/2024 11:59pm

Instructions for Submission: You can submit your solution as a Jupyter Notebook/ Matlab file with comments and discussions on the results obtained in each step

1. Follow Standard Report Format: Include sections like Introduction, Data, Methodology, Results, Discussion, and Conclusion.
2. **File Naming Convention:** Adhere to the specified naming convention for each file you submit (e.g. RollNumber_FirstName_Asg2).
3. Refrain from using zip files. If necessary, submit multiple files.
4. Include comments in the code explaining the logic and any assumptions made.
5. Include References: Cite any external sources or references used in your assignment.
6. Code Quality: Ensure your code follows best practices and is well-organized and **avoid plagiarism** as a plagiarism check will be conducted.
7. Be aware that late submissions are not permitted; ensure timely submission.

---

1. Title: Unveiling Gaussian Secrets

   Task 1: Write a Python function to generate 10,000 random numbers following a Normal distribution, embodying the symmetrical beauty of Gaussian fluctuations.

   Task 2: Decode the essence by unveiling the parameters μ (mean) and σ (standard deviation) for the Normal distribution.

   Task 3: Compute the expected point, μ, embracing the predictability of Gaussian occurrences.

   Task 4: Dive into statistical depths and compute a 95% confidence interval for the Normal distribution, shedding light on the range within which the true mean, μ, lies. Visualize your findings through a histogram, capturing the harmonious rhythm of Gaussian fluctuations.

   Please write the functions from scratch.

2. Title: The Poisson Generator

   Task 1: Generate $10,000$ random numbers using the Poisson distribution. Set the λ parameter to represent the average rate of events per interval. Document your chosen value for λ and explain how it influences the distribution.

Task 2: Calculate the expected point, the mean of your generated Poisson numbers. What insights does this provide into the central tendency of your dataset?

Task 3: Dive into the statistical depths and compute a 95% confidence interval for your Poisson distribution. Unveil the range within which you are confident the true mean lies. How does the interval reflect the reliability of your data?

Task 4: Visualize the Poisson distribution using a histogram. What patterns emerge, and how do they align with the theoretical expectations? Share your insights and observations.

Please write the functions from scratch.

3. A market researcher is conducting a study to estimate the proportion (p) of customers satisfied with a new product launch. The researcher assumes that the quality of her sample is such that it can be safely assumed to be a random sample from the Bernoulli distribution with the unknown parameter p.

   The researcher wants to determine the smallest sample size (n) needed to estimate the satisfaction proportion with $\pm 3\%$ accuracy, at a 99% confidence level. The estimator for the proportion is given by $\hat{p} = \frac{Y}{n}$, where Y is the number of satisfied customers in the sample.

   Assuming that the sampling distribution of $\hat{p}$ is well approximated by a normal distribution, your task is to write a Python function to find the smallest sample size (n) under the given conditions.

   Please write the functions from scratch.