IE 221 – Probability

Teamwork 4

Simulation Study of the Strong Law of Large Numbers and the Central Limit Theorem

Team Members

Kadir Aslancı – 2111021075

Yiğithan Aldemir – 2311021061

Zehra Sena Gündoğdu – 2211021009

Semester: Fall 2024–2025

Programming Language: Python

Table of Contents

1.Abstract

This study presents a simulation-based investigation of the Strong Law of Large Numbers (SLLN) and the Central Limit Theorem (CLT) using Monte Carlo methods. The main objective is to experimentally demonstrate the different modes of convergence associated with these two fundamental theorems.

Uniform(0,1) random variables are generated to analyze the almost sure convergence of sample means and the convergence in distribution of standardized sums. Multiple replications are performed for various sample sizes in order to observe the rate and nature of convergence. In addition, a Monte Carlo estimation of π is conducted to illustrate a practical application of the Law of Large Numbers.

The results clearly show that while SLLN convergence is observed along individual sample paths, CLT convergence requires distributional analysis through repeated experiments. These findings highlight the methodological and practical differences between the two theorems.

2.Introduction

Probability theory forms the foundation of many disciplines including engineering, data science, economics, and operations research. Real-world systems are often influenced by randomness, uncertainty, and incomplete information. In such environments, it is rarely possible to obtain exact analytical solutions, and therefore simulation and statistical estimation methods play a critical role.

The Strong Law of Large Numbers and the Central Limit Theorem are two of the most important results in probability theory that justify the use of sample-based inference. The SLLN ensures that sample averages converge to their expected values, while the CLT explains why normal distributions appear frequently in practice, even when the underlying data are not normally distributed. Understanding these theorems is essential for interpreting experimental data and designing reliable simulation models.

The purpose of this study is to experimentally verify the Strong Law of Large Numbers and the Central Limit Theorem through computer simulations. By using Monte Carlo techniques, the theoretical concepts of almost sure convergence and convergence in distribution are illustrated in a clear and intuitive way. In addition, a Monte Carlo estimation of π is included to demonstrate a practical application of the Law of Large Numbers.

3. Theoretical Background

3.1 Strong Law of Large Numbers (SLLN)

Let $\{X_1, X_2, \ldots\}$ be a sequence of independent and identically distributed random variables with finite expected value μ. The Strong Law of Large Numbers states that the sample mean converges almost surely to μ as the number of observations tends to infinity. Mathematically, this can be expressed as:

$$P\left(\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}X_i=\mu\right)=1$$

Almost sure convergence means that, with probability one, the sequence of sample means will eventually remain arbitrarily close to the expected value. This strong form of convergence provides a rigorous justification for replacing theoretical expectations with empirical averages in practice.

3.2 Central Limit Theorem (CLT)

The Central Limit Theorem describes the asymptotic behavior of sums of random variables. For a sequence of independent and identically distributed random variables with finite mean μ and variance σ^2, the standardized sum converges in distribution to a standard normal random variable. Formally:

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{d} N(0,1)$$

Unlike the SLLN, which focuses on the convergence of individual sample paths, the CLT is concerned with the convergence of the entire distribution. This distinction is crucial for understanding why different experimental methods are required to observe these two theorems.

3.3 Monte Carlo Methods

Monte Carlo methods are computational techniques that rely on repeated random sampling to approximate deterministic quantities. Their theoretical validity is guaranteed by the Law of Large Numbers. As the number of random samples increases, Monte Carlo estimates converge to the true value of the quantity of interest. These methods are widely used in numerical integration, optimization, finance, risk analysis, and engineering simulations.

3.Modes of Convergence

In probability theory, different theorems rely on different notions of convergence. The Strong Law of Large Numbers is based on almost sure convergence, which refers to the behavior of individual sample paths. This means that, with probability one, each realization of the random experiment eventually stabilizes around the expected value.

In contrast, the Central Limit Theorem is based on convergence in distribution. Rather than focusing on individual sample paths, the CLT describes how the distribution of standardized sums approaches the standard normal distribution as the sample size increases.

In the context of this study, this theoretical difference has important experimental implications. For the SLLN, a single long simulation with increasing sample size is sufficient to observe convergence of the sample mean. However, for the CLT, multiple independent replications are required in order to observe the convergence of the empirical distribution. This explains why line plots are used for SLLN, while histograms and Q–Q plots are used for CLT. Almost sure convergence is a stronger form of convergence than convergence in distribution, as it guarantees convergence with probability one for almost all realizations.

4.Methodology

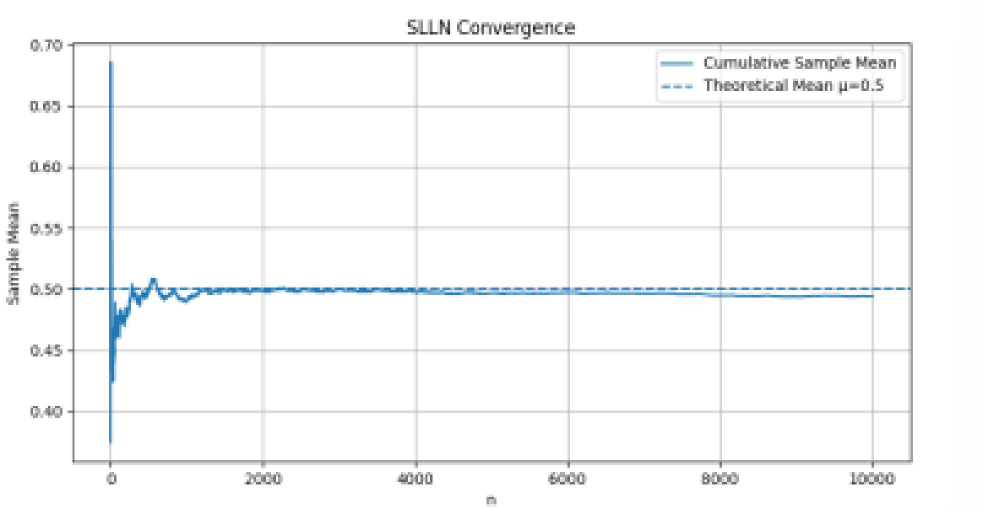
All simulations in this study were implemented in Python using the NumPy, Matplotlib, and SciPy libraries. Random numbers were generated from the Uniform(0,1) distribution.

For the SLLN simulation, a single sequence of random variables was generated and the cumulative sample mean was computed up to $n = 10{,}000$ in order to clearly observe the stabilization behavior. For the CLT simulation, sample sizes $n = 2, 5, 10, 30$, and 50 were selected to analyze the effect of increasing sample size on the distribution of standardized sums. For each value of n, $m = 1000$ independent replications were performed.

In addition, Monte Carlo estimation of π was carried out by generating random points in the unit square and counting the proportion that fall inside the unit circle. All experiments were designed to ensure reproducibility and numerical stability.

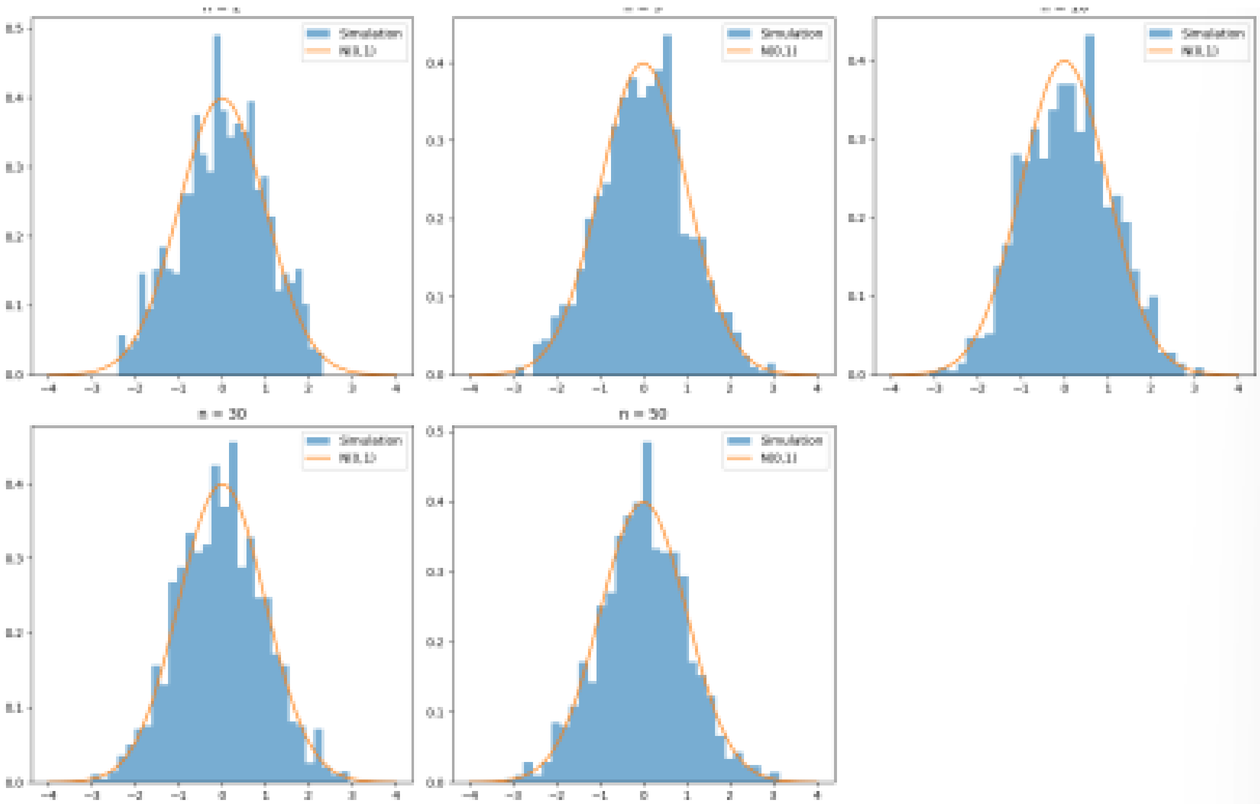
5. Results

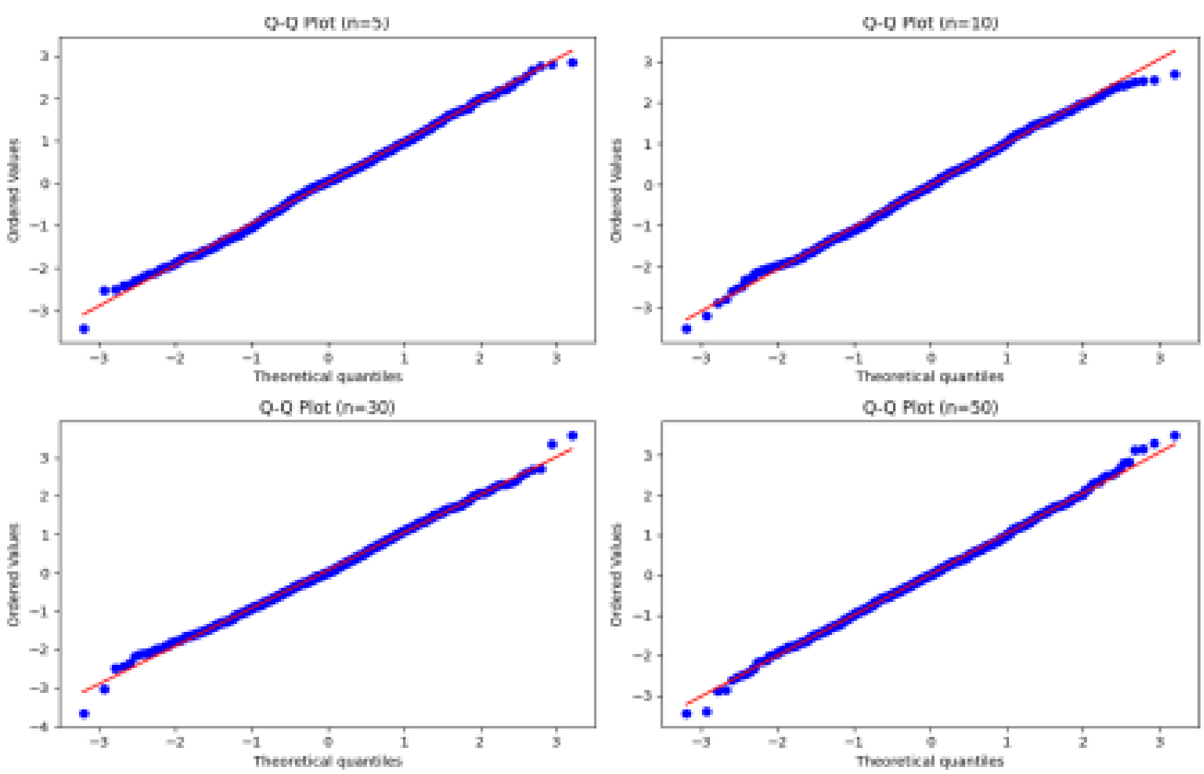
5.1 SLLN Results



The SLLN simulation shows that for small values of n, the sample mean exhibits significant fluctuations due to randomness. As the number of observations increases, these fluctuations decrease and the sample mean stabilizes around the theoretical mean of 0.5. After approximately n = 2000 observations, the sample mean remains within a narrow band around 0.5, indicating practical convergence. This behavior is consistent with the almost sure convergence predicted by the SLLN.

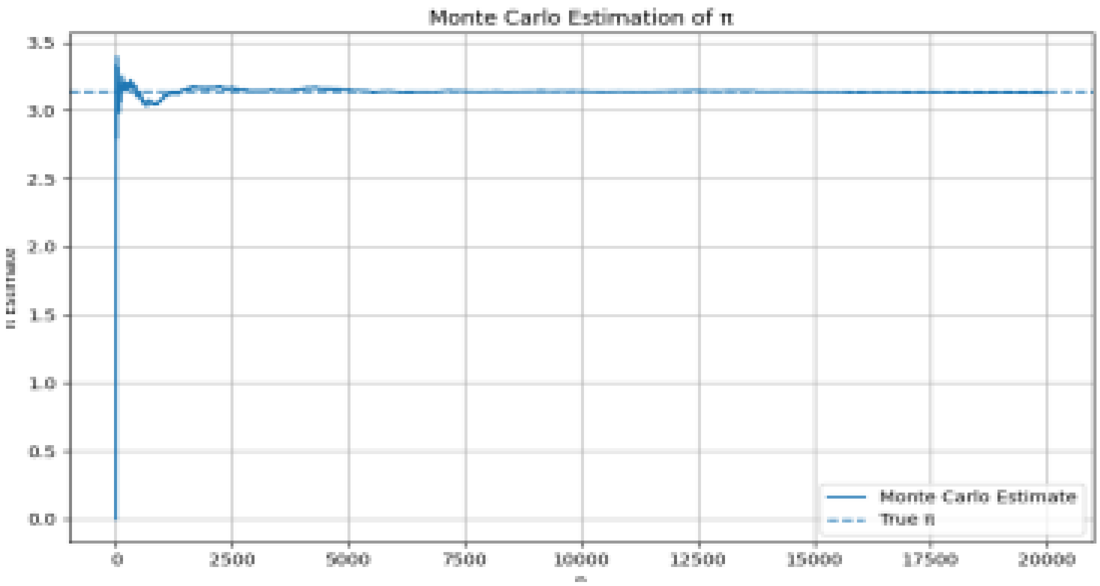
5.2 CLT Results





For small sample sizes such as n = 2 and n = 5, the histograms of standardized sums show noticeable skewness and deviation from the normal shape. As n increases to 30 and 50, the distributions become increasingly symmetric and bell-shaped. The Q–Q plots show improved linearity, indicating closer agreement with the standard normal distribution. These observations confirm the gradual nature of convergence in distribution predicted by the CLT.

5.3 Monte Carlo Estimation of π



The Monte Carlo estimation of π demonstrates a practical application of the Law of Large Numbers. As the number of simulated points increases, the estimated value of π approaches the true value. The estimation error decreases approximately at a rate proportional to $1/\sqrt{n}$, which is consistent with theoretical expectations. This result illustrates the effectiveness of Monte Carlo methods for numerical approximation.

6. Discussion and Conclusion

The simulation results clearly illustrate the fundamental differences between the Strong Law of Large Numbers and the Central Limit Theorem. While the SLLN demonstrates convergence along a single sample path, the CLT requires repeated experiments to observe distributional convergence. In practice, this means that relatively large sample sizes are needed before normal approximation becomes reliable.

Based on the experiments, values of n greater than 30 appear to be sufficient for the standardized sums to exhibit approximately normal behavior. However, much larger values are required for the sample mean to stabilize in an almost sure sense. These findings highlight the importance of understanding the mode of convergence when designing simulations and interpreting statistical results. In engineering applications, misinterpreting convergence type may lead to incorrect conclusions about reliability and variability.

8. References

1. **Ross, S. M. (2014). *Introduction to Probability Models* (11th ed.). Academic Press.**
2. **Grimmett, G., & Stirzaker, D. (2001). *Probability and Random Processes* (3rd ed.). Oxford University Press.**
3. **Durrett, R. (2019). *Probability: Theory and Examples* (5th ed.). Cambridge University Press.**
4. **Casella, G., & Berger, R. L. (2002). *Statistical Inference* (2nd ed.). Duxbury Press.**
5. **Fishman, G. S. (1996). *Monte Carlo: Concepts, Algorithms, and Applications*. Springer.**
6. **Kroese, D. P., Taimre, T., & Botev, Z. I. (2011). *Handbook of Monte Carlo Methods*. Wiley.**
7. **Devroye, L. (1986). *Non-Uniform Random Variate Generation*. Springer.**
8. **NumPy Developers (2023). *NumPy Documentation*. https://numpy.org/doc/**
9. **Matplotlib Development Team (2023). *Matplotlib Documentation*. https://matplotlib.org/stable/**
10. **SciPy Developers (2023). *SciPy Documentation*. https://docs.scipy.org/doc/**