IE221 PROBABILITY

Experimental Verification of the SLLN and CLT via Monte Carlo Simulation

Fall 2024–2025

Teamwork Group 22

Beyzanur ÇİFTÇİ
Buse YAKUTSOY
Yaren GÖRGÜLÜ
Fatmanur DEMİRCİ

07.01.2026

TABLE OF CONTENTS

LIST OF FIGURES

1.INTRODUCTION

Probability theory provides the mathematical foundation for modeling uncertainty and randomness in many scientific and engineering applications. Among its most fundamental results are limit theorems, which describe the behavior of random variables as the number of observations increases.

In particular, the Strong Law of Large Numbers (SLLN) and the Central Limit Theorem (CLT) play a central role in statistics, simulation, and data-driven decision making. These theorems justify the widespread use of sample averages and normal approximations in practical problems where exact analytical solutions are often unavailable.

While the SLLN guarantees that sample averages converge almost surely to the true expected value, the CLT explains why normalized sums of random variables tend to follow a normal distribution, regardless of the underlying distribution. Although these results are asymptotic in nature, their practical relevance depends on how convergence manifests itself for finite sample sizes. Therefore, numerical experiments are essential for understanding the speed and nature of convergence in real applications.

Monte Carlo simulation provides a powerful experimental framework for investigating such convergence properties. By relying on repeated random sampling, Monte Carlo methods allow complex probabilistic phenomena to be explored through computational experiments. In this project, Monte Carlo simulation is used to experimentally verify the SLLN and CLT and to demonstrate their implications through visualization. In addition, a Monte Carlo approach is employed to estimate the value of π using geometric probability, illustrating how abstract probabilistic concepts can be translated into concrete numerical results.

The main objective of this study is to bridge the gap between theoretical probability and computational practice. By combining theoretical background with simulation-based evidence, the project highlights the differences between pathwise convergence and distributional convergence and emphasizes the importance of careful interpretation of simulation results in applied settings.

2. THEORETICAL BACKGROUND

2.1 Strong Law of Large Numbers (SLLN)

Let $X_1, X_2, \ldots$be i.i.d. random variables with finite expected value $\mathbb{E}[X_i] = \mu$.
The Strong Law of Large Numbers states that the sample mean

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

converges almost surely to μ, i.e.,

$$\mathbb{P}\left(\lim_{n\to\infty} \bar{X}_n = \mu\right) = 1.$$

Almost sure convergence is one of the strongest forms of convergence in probability theory. It implies that, with probability one, the sequence of sample means converges to the true expected value along individual realizations of the experiment. As a result, the effect of random fluctuations diminishes as the number of observations increases.

From a practical perspective, the SLLN provides the theoretical justification for replacing expected values with sample averages in large-scale simulations and empirical studies.

2.2 Central Limit Theorem (CLT)

Under the same assumptions and with finite variance σ^2, the Central Limit Theorem states that the standardized sum

$$Z_n = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}}$$

converges in distribution to a standard normal random variable:

$$Z_n \xrightarrow{d} \mathcal{N}(0,1).$$

Unlike the SLLN, the Central Limit Theorem does not describe convergence of individual sample
paths. Instead, it characterizes the limiting distribution of properly normalized sums of random variables. This distinction is crucial for understanding how CLT-based results should be interpreted in practice.

Consequently, experimental verification of the CLT requires repeated sampling and analysis of empirical distributions rather than single trajectories.

2.3 Monte Carlo Method

Monte Carlo methods rely on repeated random sampling to approximate deterministic quantities that may be difficult or impossible to compute analytically. In this project, Monte Carlo simulation is employed both to investigate the convergence properties of fundamental limit theorems and to estimate the value of π using geometric probability.

The basic idea behind Monte Carlo methods is to exploit randomness to obtain numerical approximations, with accuracy improving as the number of samples increases. According to the law of large numbers, Monte Carlo estimators converge to the true value as the sample size grows, providing a theoretical justification for their use.

One of the main advantages of Monte Carlo methods is their simplicity and flexibility. They can be applied to a wide range of problems without requiring strong assumptions about the underlying structure of the system. However, the convergence rate of Monte Carlo estimators is relatively slow compared to deterministic numerical methods.

Despite this limitation, Monte Carlo techniques are widely used in high-dimensional problems and stochastic systems, where traditional numerical integration becomes impractical or computationally infeasible.

3. MODES OF CONVERGENCE

3.1 Theoretical Comparison

Theorem	**Convergence Type**	**Meaning**
SLLN	Almost sure convergence	Sample paths converge to μ with probability 1
CLT	Convergence in distribution	Distributions converge to $\mathcal{N}(0,1)$

In probability theory, different modes of convergence are used to describe how sequences of random variables behave as the number of observations increases. Among these, almost sure convergence and convergence in distribution play a central role in the formulation of the Strong Law of Large Numbers (SLLN) and the Central Limit Theorem (CLT), respectively.

Almost sure convergence is one of the strongest forms of convergence. It implies that, with probability one, the sequence of random variables converges to a fixed value along individual sample paths. The SLLN is a classical example of this concept, as it guarantees that the sample mean converges almost surely to the true expected value. This provides a strong assurance that long-run averages stabilize as the sample size increases.

In contrast, convergence in distribution concerns the limiting behavior of probability distributions rather than individual realizations. The CLT states that properly standardized sums of independent and identically distributed random variables converge in distribution to a normal distribution. Importantly, convergence in distribution does not imply convergence of individual sample paths, but instead describes how the overall shape of the distribution evolves as the number of observations grows.

These theoretical differences highlight that almost sure convergence is a stronger notion thanconvergence in distribution. While the former guarantees pathwise stability, the latter explains the widespread appearance of the normal distribution in statistical applications.

3.2 Experimental Comparison

The theoretical distinctions between convergence modes are directly reflected in the design and interpretation of simulation experiments. In the case of almost sure convergence, a single long simulation is sufficient to observe stabilization of the sample mean over time. Therefore, SLLN is typically illustrated using running mean or trajectory plots.

On the other hand, convergence in distribution requires repeated sampling in order to construct empirical distributions. For this reason, CLT experiments rely on multiple independent replications for each sample size. Histograms and quantile–quantile (Q–Q) plots are commonly used to compare empirical distributions with the theoretical normal distribution.

The simulation results clearly demonstrate these differences. For the SLLN, the running mean exhibits noticeable fluctuations in the early stages but gradually stabilizes as the number of observations increases, confirming almost sure convergence. For the CLT, empirical distributions of standardized sums increasingly resemble the bell shape of the normal distribution, and Q–Q plots show improved alignment with the reference line as the sample size grows.

These observations emphasize that different convergence results require different visualization and interpretation strategies. Recognizing these differences is essential for correctly analyzing simulation outputs and for applying probabilistic results in practice.

4. METHODOLOGY

All simulations in this study were implemented using the Python programming language. Numerical computations and random number generation were carried out using the NumPy library, while SciPy was used for statistical tools and reference probability distributions. Visualization of simulation results was performed using Matplotlib.

Random samples were generated from a Uniform(0,1) distribution for all experiments. This distribution was chosen due to its simplicity and well-defined theoretical properties, allowing the focus to remain on convergence behavior rather than distributional complexity. A sufficiently large number of samples was used in each experiment to ensure that asymptotic convergence trends could be clearly observed.

For the Strong Law of Large Numbers (SLLN) experiment, a single long sequence of random variables was generated. The running mean of this sequence was computed incrementally to visualize almost sure convergence of the sample average to the theoretical expected value. This approach reflects the pathwise nature of almost sure convergence and allows stabilization behavior to be observed directly.

In contrast, the Central Limit Theorem (CLT) experiment required multiple independent replicationsfor each selected sample size. For each replication, sums of random variables were standardized using the theoretical mean and variance of the underlying distribution. The resulting standardized values were then analyzed using histograms and quantile–quantile (Q–Q) plots to assess convergence in distribution toward the standard normal distribution.

The Monte Carlo estimation of π was performed using a geometric probability approach. Random points were generated uniformly within the unit square, and the proportion of points falling inside the unit quarter circle was recorded. This proportion was multiplied by four to obtain an estimate of π, and the convergence of this estimator was examined as the number of samples increased.

All experiments were conducted in a controlled computational environment, and intermediate results were stored to ensure consistency between numerical outputs and visual representations. This methodological framework enables a systematic experimental verification of theoretical probabilistic results through simulation.

5. RESULTS

5.1 SLLN Results

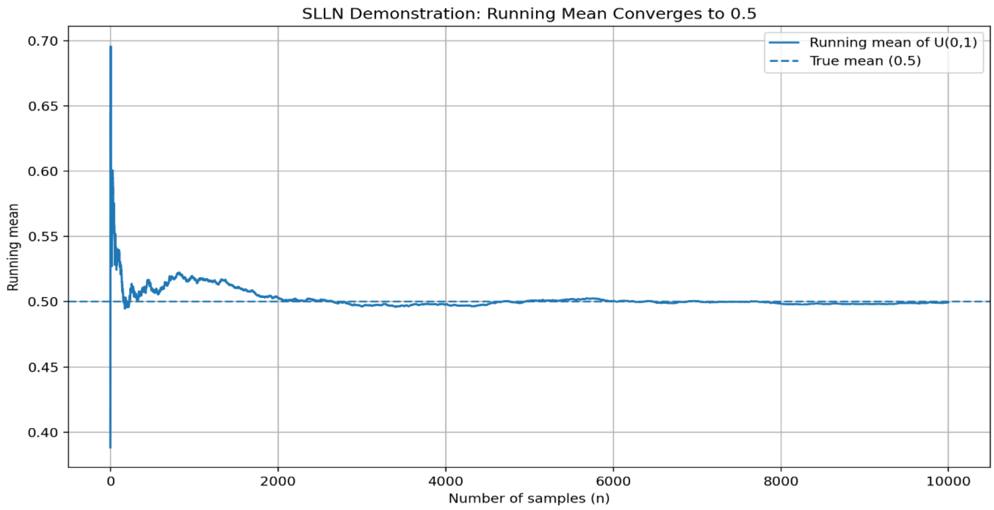

Figure 1.*Running mean illustrating SLLN convergence*

Figure 1 illustrates the running mean of a sequence of independent and identically distributed random variables generated from a Uniform(0,1) distribution. At small sample sizes, the running mean exhibits noticeable fluctuations due to randomness. However, as the number of observations increases, the running mean gradually stabilizes around the theoretical expected value.

This behavior provides empirical evidence of almost sure convergence, as predicted by the Strong Law of Large Numbers. The diminishing impact of individual observations demonstrates that long-run averages become increasingly reliable as the sample size grows.

5.2 CLT Results

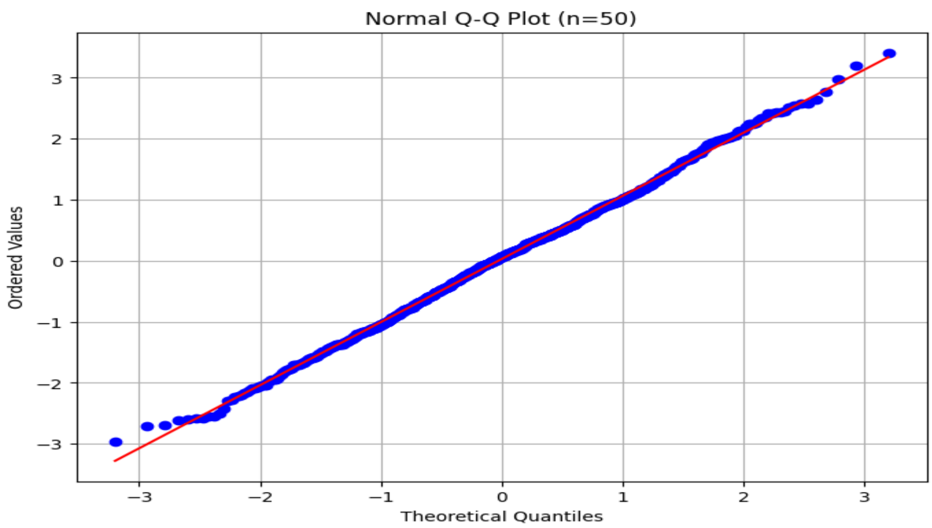

Figure 2.*Histograms of standardized sums for different sample sizes.*

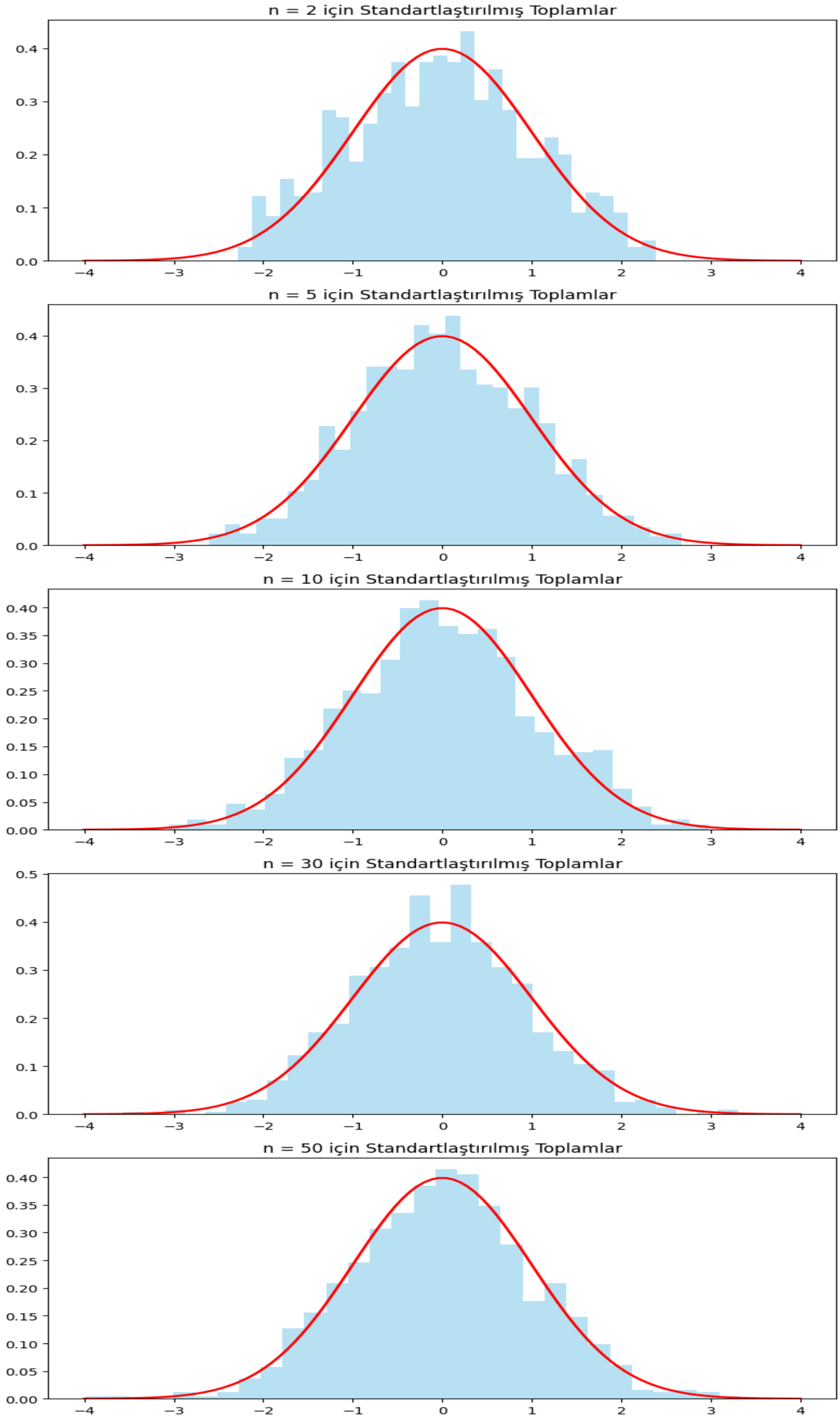

Figure 3. *Q–Q plot of standardized sums compared with the standard normal distribution.*

Figures 2 and 3 present the results of the Central Limit Theorem experiments. Figure 2 shows histograms of standardized sums for different sample sizes. As the sample size increases, the empirical distributions increasingly resemble the bell-shaped curve of the standard normal distribution.

Figure 3 displays the corresponding Q–Q plot, where the alignment of points with the reference line improves as the sample size grows. These results confirm convergence in distribution toward normality and illustrate that larger sample sizes are required to clearly observe CLT behavior.

5.3 Monte Carlo π Estimation

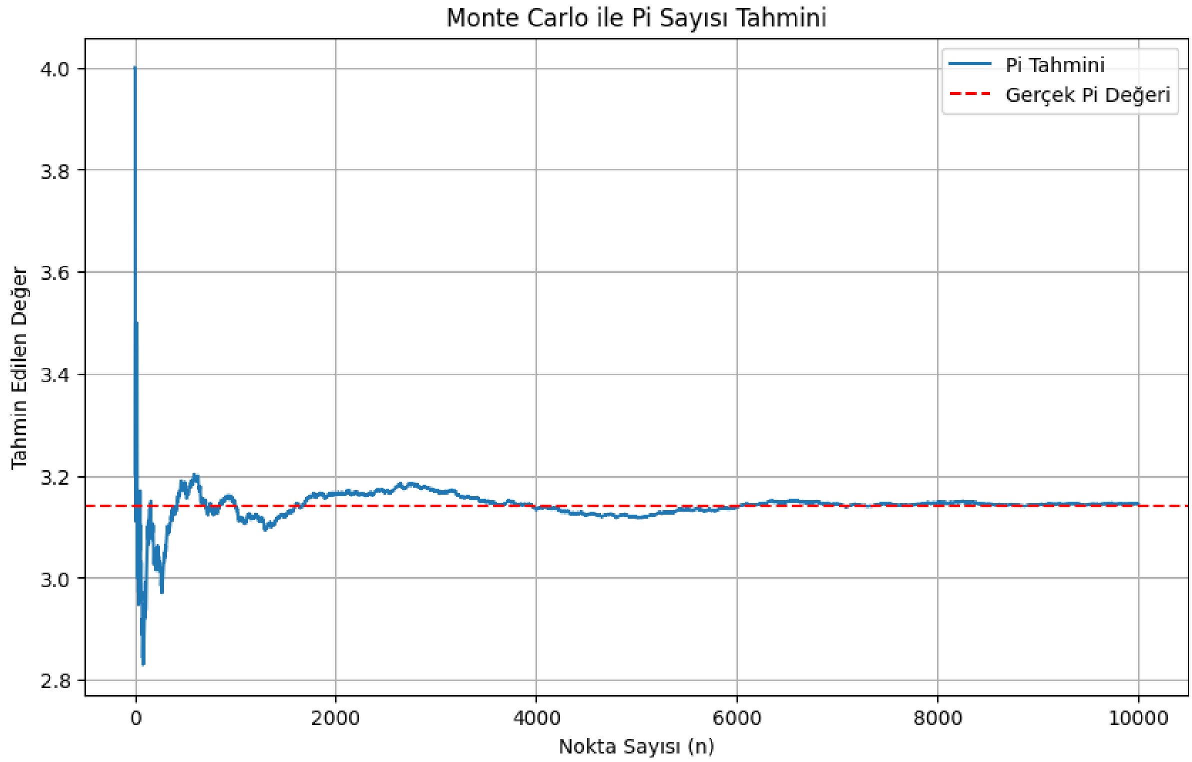

Figure 4.*Monte Carlo estimation of π using geometric probability.*

Figure 4 shows the results of the Monte Carlo estimation of π using geometric probability. For small numbers of samples, the estimator exhibits significant variability. As the number of random points increases, the estimated value converges toward the true value of π.

Although the convergence is relatively slow, the overall trend demonstrates the effectiveness of Monte Carlo simulation for approximating deterministic quantities. The results highlight the trade-off between simplicity and computational efficiency inherent in Monte Carlo methods.

6. DISCUSSION AND CONCLUSION

The simulation experiments clearly demonstrate that although both the Strong Law of Large Numbers (SLLN) and the Central Limit Theorem (CLT) describe limiting behavior as the number of observations increases, their convergence characteristics differ fundamentally. While the SLLN guarantees almost sure convergence of the sample mean along individual sample paths, the CLT describes convergence in distribution of standardized sums. As observed in the results, this distinction has important practical implications for how convergence is interpreted in numerical experiments.

The SLLN results show that the running mean stabilizes around the theoretical expected value after an initial period of fluctuation. These early fluctuations are a consequence of randomness and finite-sample effects and should not be interpreted as violations of theoretical results. Instead, they highlight the importance of sufficiently large sample sizes when relying on sample averages in practice. This observation is particularly relevant in engineering and data-driven applications, where decisions are often based on limited data.

In contrast, the CLT results emphasize that distributional convergence requires repeated sampling and larger sample sizes. Histograms and Q–Q plots illustrate that convergence to the normal distribution becomes more apparent as the sample size increases. This demonstrates that CLT-based approximations may be unreliable for small sample sizes and must be applied with caution in practical settings.

The Monte Carlo estimation of π further illustrates both the strengths and limitations of simulation-based methods. Although the estimator converges slowly, its simplicity and robustness make Monte Carlo simulation a valuable tool when analytical solutions are unavailable or computationally infeasible. The observed convergence behavior is consistent with theoretical expectations and highlights the trade-off between accuracy and computational effort.

In conclusion, this project demonstrates how fundamental probabilistic theorems can be experimentally verified through simulation. By combining theoretical background with numerical experiments, the study bridges the gap between abstract probability theory and computational practice. The results emphasize the importance of understanding different modes of convergence and carefully interpreting simulation outputs, particularly when applying probabilistic methods in real-world applications.