AGU IE221 PROBABILITY

TEAMWORK 4 TECHNICAL REPORT

Prepared by

Nil Belinay Akoyun 2211021005

Mina Soyçeri 2311021014

Gönül Hülya Atik 2211021057

Date of the Submission

10.01.2026

Kayseri/Turkey

Table of Contents

List of Tables

List of Figures

1. INTRODUCTION

Probability is an important area of mathematics that is widely used in engineering, statistics, and data analysis. Many theoretical results in probability are difficult to understand only by formulas. Therefore, computer simulations are very useful to observe and understand these concepts in a practical way.

The main purpose of this project is to study two fundamental theorems in probability theory: the Strong Law of Large Numbers (SLLN) and the Central Limit Theorem (CLT). These theorems explain how random variables behave when the number of observations increases. Although both theorems describe convergence, they represent different types of convergence and have different interpretations.

In this project, simulations are used to experimentally verify the SLLN and the CLT. The behavior of sample means and distributions is analysed for increasing sample sizes. In addition, the Monte Carlo method is applied to estimate the value of the mathematical constant π. This method shows how random sampling can be used to solve numerical problems.

All simulations are implemented using the Python programming language. The results are presented with graphical visualizations, such as convergence plots and histograms, to clearly demonstrate the theoretical concepts. The project aims to provide a clear and practical understanding of probabilistic convergence and simulation-based methods.

2. THEORETICAL BACKGROUND

2.1 Strong Law of Large Numbers (SLLN)

The Strong Law of Large Numbers states that the sample mean of independent and identically distributed random variables converges almost surely to the expected value as the sample size goes to infinity. Mathematically, it can be expressed as:

$$\frac{1}{n}\sum_{i=1}^{n} Xi \overset{as}{\rightarrow} \mu$$

where Xi are independent and identically distributed random variables with expected value $\mu = E[Xi]$

Assumptions of SLLN:

- The random variables are independent.
- The random variables are identically distributed.
- The expected value E[Xi] exists and is finite.

This theorem explains why long-run averages become stable and approach the true mean with probability one.

2.2 Central Limit Theorem (CLT)

The Central Limit Theorem describes the distributional behavior of the sample mean. According to the CLT, for a sufficiently large sample size, the distribution of the standardized sample mean converges to the standard normal distribution. The mathematical statement is:

$$\frac{\sqrt{n}(\overline{X}n - \mu)}{\sigma} \overset{d}{\rightarrow} \mathbb{N}(0,1)$$

Assumptions of CLT:

- The random variables are independent and identically distributed.
- The mean μ and variance σ2 are finite.
- The sample size n is sufficiently large.

The CLT is important because it allows the use of the normal distribution in many practical applications.

2.3 Theoretical Basis of the Monte Carlo Method

The Monte Carlo method is a numerical simulation technique based on randomsampling. Its mathematical foundation mainly relies on the Law of Large Numbers, which ensures convergence of sample averages to expected values.

Let X1,X2,…,Xn be independent and identically distributed random variables with finite expected value E[X]. According to the Law of Large Numbers, the Monte Carlo estimator can be written as:

$$\frac{1}{n}\sum_{i=1}^{n} Xi \xrightarrow{as} E[\mathrm{X}]$$

This expression shows that as the number of simulations nnn increases,the Monte Carlo estimate converges to the true expected value.

In the estimation of π, random points are generated uniformly inside the unit square [0,1]×[0,1]. Let be an indicator variable defined as:

$$Ii = \begin{cases} 1, & if\ x_i^2 + y_i^2 \leq 1 \\ 0, & otherwise \end{cases}$$

The Monte Carlo estimator of π is then given by:

$$\widehat{\pi_n} = \widehat{4\,x\,\frac{1}{n}}\sum_{\iota=1}^{n} I\iota$$

As the number of points n increases, the estimated value $\widehat{\pi_n}$ converges to the true value of π. This demonstrates how random sampling can be used to obtain numerical approximations of mathematical constants.

3. MODES OF CONVERGENCE

Although both the Strong Law of Large Numbers (SLLN) and the Central Limit Theorem (CLT) describe limiting behaviors of sequences of random variables, they are based on different modes of convergence. Understanding these differences is essential for correctly interpreting both the theoretical results and the simulation outcomes presented in this study.

Theorem	Mode of Convergence	Description
SLLN	Almost sure convergence	The sample mean converges to the true mean μ with probability 1
CLT	Convergence in distribution	The distribution of standardized sums converges to N(0,1)

Table 1:Convergence Types

Almost sure convergence is a stronger form of convergence than convergence in distribution. While almost sure convergence focuses on the behavior of individual sample paths, convergence in distribution concerns the limiting shape of probability distributions.

3.1 SLLN and Almost Sure Convergence

In the context of the Strong Law of Large Numbers, almost sure convergence means that, for almost every realization of the random sequence, the sample mean converges to the expected value μ as the number of observations tends to infinity. This convergence occurs along a single sample path and guarantees that deviations from the true mean become negligible with probability one.

The experimental results presented in Figure 1 support this theoretical property. The convergence of the cumulative sample mean toward $\mu = 0.5$ demonstrates that, as the sample size increases, the sample mean stabilizes and remains close to the true mean. This behavior is consistent with the definition of almost sure convergence and confirms the validity of the SLLN in practice.

3.2 CLT and Convergence in Distribution

In the Central Limit Theorem, convergence occurs in distribution rather than almost surely. Convergence in distribution means that the distribution of the standardized sum of random variables approaches the standard normal distribution as the sample size increases. Unlike almost sure convergence in the Strong Law of Large Numbers, convergence in distribution does not guarantee convergence along individual sample paths.

Instead, it describes the limiting behavior of the distribution as a whole. Therefore, observing a single realization is not sufficient to verify the CLT experimentally. For this reason, repeated simulations are required in order to examine the distribution of standardized sums. In this study, histograms and Q-Q plots are used to compare the empirical distributions with the theoretical standard normal distribution. These visualization techniques provide clear evidence of convergence in distribution as predicted by the Central Limit Theorem.

3.3 Experimental Implications

The difference between almost sure convergence and convergence in distribution directly affects the experimental methodology. For SLLN, a single long simulation is sufficient to observe convergence, since the theorem concerns pathwise behavior. In contrast, CLT requires repeated simulations to capture distributional properties.

This distinction explains why different visualization techniques are used in this study: a convergence plot for SLLN and histograms with Q-Q plots for CLT. Together, these results highlight the fundamental conceptual difference between the two theorems and demonstrate how the choice of convergence mode determines the appropriate experimental approach.

4. METHODOLOGY

This section describes the simulation methodology used to experimentally verify the Strong Law of Large Numbers (SLLN), the Central Limit Theorem (CLT), and the Monte Carlo estimation of π. All simulations were implemented using the Python programming language and were designed to clearly demonstrate the theoretical properties discussed in the previous sections.

4.1 Programming Environment and Tools

All simulations were implemented in Python, which provides efficient tools for numerical computation and visualization. The simulations were developed and executed either on a local machine or using GitHub Codespaces. The following libraries were used throughout the study:

- NumPy for random number generation and numerical operations,
- Matplotlib for visualizing simulation results,

- SciPy (where applicable) for statistical analysis such as Q-Q plots.

4.2 SLLN Simulation Setup

To experimentally verify the Strong Law of Large Numbers, independent and identically distributed random variables were generated from the Uniform(0,1) distribution. The theoretical mean of this distribution is $\mu = 0.5$.

A sequence of random variables $X_1, X_2, \dots, X_n$ was generated, and the cumulative sample mean was computed as

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

The sample size was increased up to $n = 20{,}000$, and the cumulative sample mean was plotted against the number of observations to observe convergence behavior. A fixed random seed was used to ensure reproducibility of the results.

4.2 CLT Simulation Setup

To experimentally verify the Central Limit Theorem, independent and identically distributed random variables were generated from the Uniform(0,1) distribution. The theoretical mean and standard deviation of this distribution are $\mu = 0.5$ and $\sigma = \sqrt{(1/12)}$, respectively. For the simulation, $m = 1000$ independent experiments were conducted for each sample size $n \in \{2, 5, 10, 30, 50\}$. In each experiment, n independent Uniform(0,1) random variables were generated and summed. To analyze convergence in distribution, the sums were standardized according to the Central Limit Theorem as $Z = (S_n - n\mu) / (\sigma\sqrt{n})$, where S_n denotes the sum of n random variables. For each value of n, histograms of the standardized sums were plotted and compared with the probability density function of the standard normal distribution N(0,1). Additionally, Q-Q plots were constructed to visually assess the agreement between the empirical distribution and the theoretical normal distribution.

4.3 Monte Carlo Estimation of π

The Monte Carlo method was applied to estimate the value of π by randomly generating points uniformly in the unit square $[0,1] \times [0,1]$. For each generated point, it was checked whether the point lies inside the quarter unit circle defined by $x^2 + y^2 \leq 1$. The estimate of π was computed using the ratio of points inside the quarter circle to the total number of generated points. The number of points was increased gradually to observe the convergence behavior of the estimator. Additionally, the absolute estimation error was calculated to analyze how the accuracy improves as the sample size increases.

4.4 Reproducibility and Visualization

To ensure reproducibility, fixed random seeds were used in all simulations. Simulation results were visualized using convergence plots, histograms, and Q-Q plots, which provided clear graphical representations of the theoretical concepts and facilitated interpretation of the results.

5. RESULTS

5.1 SLLN Results

In this part of the study, the Strong Law of Large Numbers was experimentally investigated using independent and identically distributed Uniform(0,1) random variables. The cumulative sample mean was computed and plotted as the number of observations increased in order to observe the convergence behavior predicted by the theorem.

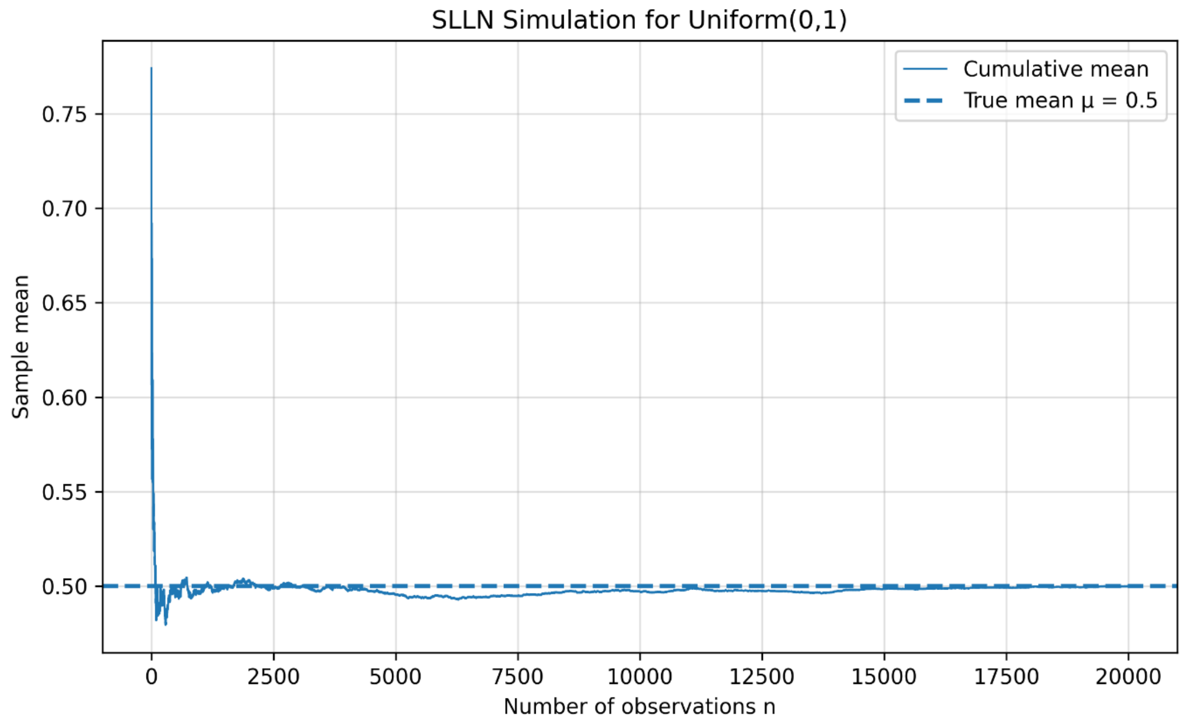

Figure 1:Convergence of the sample mean under the Strong Law of Large Numbers

Figure 1 illustrates the convergence of the sample mean of i.i.d. Uniform (0,1) random variables toward the theoretical mean $\mu = 0.5$. At small sample sizes, the sample mean exhibits noticeable fluctuations due to randomness. However, as the number of observations increases, these fluctuations gradually decrease and the sample mean stabilizes around the true mean.

From the simulation results, it can be observed that for sufficiently large sample sizes (approximately $n \geq 10{,}000$), the sample mean remains very close to $\mu = 0.5$. This behavior provides clear experimental evidence of the Strong Law of Large Numbers, which states that the sample mean converges almost surely to the expected value.

5.2 CLT Results

In this section, the Central Limit Theorem is examined through simulation results based on standardized sums of Uniform (0,1) random variables.

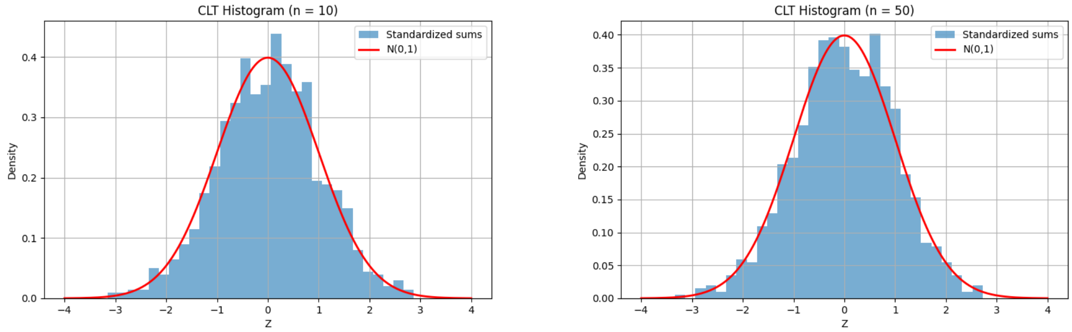

Figure 2:Histograms of standardized sums for different sample sizes with the standart normal density function

Figure 2 presents histograms of the standardized sums for different sample sizes (n = 2, 5, 10, 30, and 50), together with the probability density function of the standard normal distribution. For very small sample sizes, such as n = 2 and n = 5, the empirical distribution deviates significantly from normality and exhibits visible skewness. However, as the sample size increases, the histograms become more symmetric and increasingly resemble the bell-shaped normal distribution. For n = 30 and n = 50, the agreement between the empirical distributions and the standard normal density is remarkably close, providing strong visual evidence for convergence in distribution as predicted by the Central Limit Theorem.

Further evidence of convergence in distribution is provided by the Q-Q plots shown in Figure 3. For small sample sizes, noticeable deviations from the reference line can be observed, indicating departures from normality. As the sample size increases, the plotted quantiles align more closely with the theoretical normal quantiles.

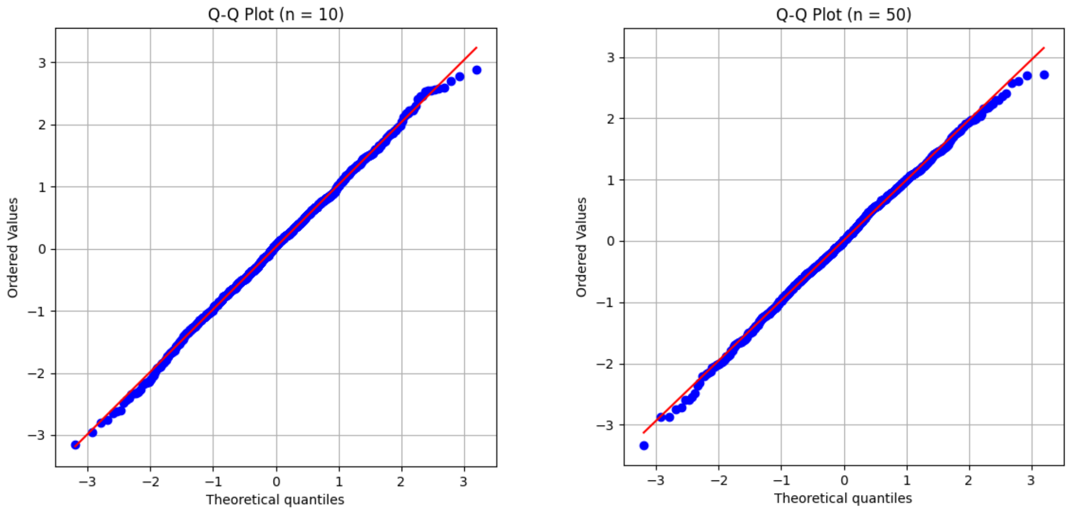

Figure 3:Histograms of standardized sums for n = 10 and n = 50 with the standard normal density

In particular, for n = 30 and n = 50, the points lie almost entirely along the reference line, confirming that the standardized sums follow a distribution that is well-approximated by the standard normal distribution.

5.3 Monte Carlo π Estimation

In this part of the study, the Monte Carlo method was used to estimate the value of π. Random points were generated uniformly in the unit square [0,1] × [0,1], and the ratio of points falling inside the quarter unit circle was used to approximate π.

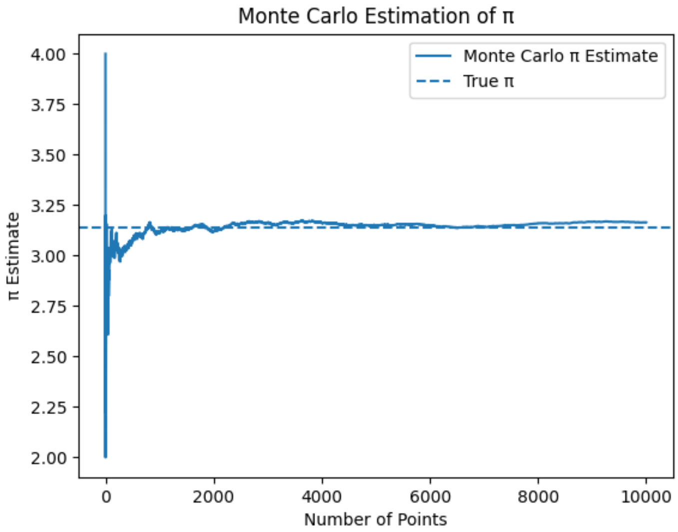

Figure 4:Monte Carlo estimation of π convergence

Figure 4 shows the convergence of the Monte Carlo estimation of π as the number of random points increases. At small sample sizes, the estimation fluctuates significantly due to randomness. As the number of points increases, the estimated value stabilizes and converges to the true value of π.

The accuracy of the Monte Carlo estimation was evaluated using the absolute error, defined as the absolute difference between the estimated value of π and the true value of π.

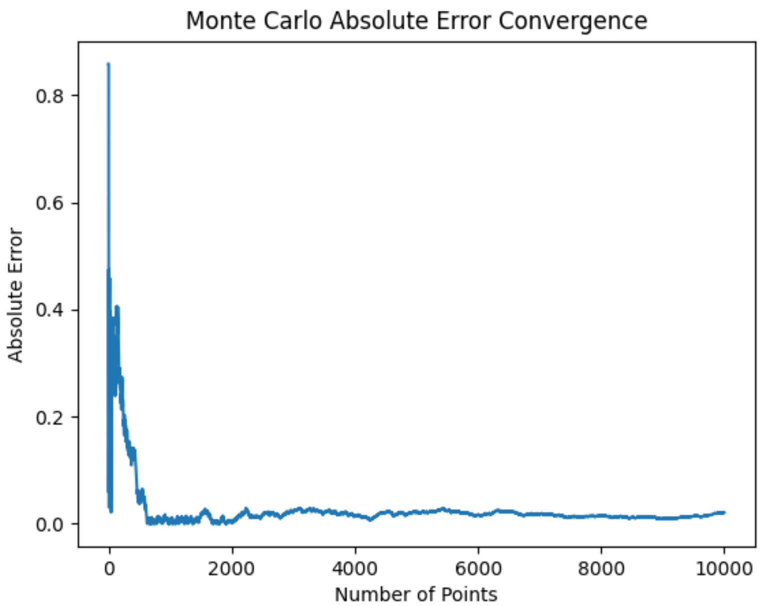

Figure 5:Monte Carlo Absolute Error Convergence

Figure 5 shows that the absolute error of the Monte Carlo π estimation decreases as the number of random points increases. At low iteration counts, the error is relatively large due to the randomness of the sampling process. As the sample size grows, the error rapidly declines and stabilizes around small values. This behavior indicates that the Monte Carlo method becomes more accurate with larger numbers of samples and confirms the convergence of the estimation toward the true value of π.

6. CONCLUSION

In this study, fundamental concepts of probability theory were investigated through computer simulations. The Strong Law of Large Numbers, the Central Limit Theorem, and the Monte Carlo method were examined both theoretically and experimentally using Python-based simulations.

The results clearly showed that the Strong Law of Large Numbers guarantees almost sure convergence of sample means to the expected value, which was observed through the stabilization of cumulative averages. In contrast, the Central Limit Theorem was shown to describe convergence in distribution, requiring repeated simulations and distributional analysis to observe the emergence of the normal distribution. This distinction highlights the importance of understanding different modes of convergence and their implications for experimental design.

Additionally, the Monte Carlo estimation of π demonstrated how probabilistic methods can be applied to solve numerical problems. The convergence of the estimator and the reduction of estimation error with increasing sample size confirmed the practical usefulness of simulation-based approaches.

Overall, this project illustrates how computational simulations can enhance the understanding of abstract probabilistic theorems by providing clear visual and numerical evidence. Such simulation-based methods are especially valuable in engineering and applied sciences, where analytical solutions may be difficult or impossible to obtain.