Convergence to Expected Value in Repeated Trials

Ayman Azad

Memorial University of Newfoundland

MATH-3030 Project-1

Abstract

Monte Carlo simulation estimates theoretical quantities by averaging the outcomes of repeated random experiments. This project studies this idea in a basic yet fundamental setting using repeated Bernoulli trials, where each trial yields a success or failure with fixed probability. The main focus is the convergence of the sample mean to the expected value as the number of trials increases.

The analysis extends beyond visual convergence by examining the decay rate of the estimation error, its asymptotic distribution, and principled uncertainty quantification through confidence intervals. Theoretical predictions are compared with numerical simulations, including finite-sample bounds derived from concentration inequalities. Additional experiments explore rare-event behavior and demonstrate the failure of convergence when the conditions of the Law of Large Numbers are not met.

1   Introduction

Repeated random experiments form the basis of many methods in probability theory, statistics, and computational mathematics. When analytical evaluation of expected values or probabilities is infeasible, simulation provides an effective alternative: random experiments are repeated, and numerical summaries of the outcomes are used to approximate theoretical quantities. This approach, known as Monte Carlo simulation, plays a central role in modern numerical analysis [1].

A key estimator arising from repeated sampling is the sample mean, whose convergence to the expected value is guaranteed by the Law of Large Numbers. While this result establishes the existence of convergence, practical applications require a deeper understanding of how rapidly convergence occurs, how much uncertainty remains for finite samples, and when convergence may fail.

Using repeated Bernoulli trials as a fundamental model, this project investigates these questions by combining probabilistic theory with numerical simulation. The analysis focuses on quantitative aspects of convergence, including rates of convergence, confidence interval construction, and analytical error bounds, as well as examples that illustrate violations of the assumptions underlying the Law of Large Numbers.

2 Methods

2.1 Probability model

We consider a sequence of repeated random experiments modeled by independent and identically distributed Bernoulli random variables[3]. Let $X_1, X_2, \ldots$ be i.i.d. Bernoulli(p) variables defined by

$$X_i = \begin{cases} 1, & \text{with probability } p, \\ 0, & \text{with probability } 1 - p. \end{cases}$$

Each X_i represents the outcome of the ith trial, where a value of 1 denotes success and 0 denotes failure. The expected value and variance of a single trial are

$$\mathbb{E}[X_i] = p, \qquad \text{Var}(X_i) = p(1 - p).$$

2.2 Estimator: sample mean

To estimate the expected value from N repeated trials, we use the sample mean

$$\bar{X}_N = \frac{1}{N} \sum_{i=1}^{N} X_i.$$

For Bernoulli trials, the parameter of interest is $p = \mathbb{E}[X_1]$. Thus, the convergence problem studied in this project is to understand how $\bar{X}_N$ approaches p as the number of trials N increases.

2.3 Law of Large Numbers LLN

The Law of Large Numbers explains why averaging many independent observations yields more accurate estimates of an underlying expected value[3]. Specifically, let $X_1, X_2, \ldots$ be independent and identically distributed random variables with finite mean $\mu = \mathbb{E}[X_1]$. The sample mean

$$\bar{X}_N = \frac{1}{N} \sum_{i=1}^{N} X_i.$$

Then the Law of Large Numbers states that

$$\bar{X}_N \xrightarrow{P} \mu \quad \text{as } N \to \infty.$$

Equivalently, for any tolerance $\varepsilon > 0$,

$$\lim_{N \to \infty} \mathbb{P}\big(|\bar{X}_N - \mu| > \varepsilon\big) = 0.$$

In the case of Bernoulli(p) trials, the mean is $\mu = p$. Therefore, as the number of trials increases, the sample proportion of successes $\bar{X}_N$ converges in probability to the true success probability p.

2.4 Rate of Convergence of the Sample Mean

To quantify how fast the sample mean converges to the expected value, we examine its variance. Since $X_1, X_2, \ldots, X_N$ are independent and identically distributed, the variance of the sample mean is

$$\text{Var}(\bar{X}_N) = \text{Var}\left(\frac{1}{N}\sum_{i=1}^{N} X_i\right) = \frac{1}{N^2}\sum_{i=1}^{N}\text{Var}(X_i) = \frac{\text{Var}(X_1)}{N}.$$

For Bernoulli(p) trials, $\text{Var}(X_1) = p(1-p)$, and therefore

$$\text{Var}(\bar{X}_N) = \frac{p(1-p)}{N}, \qquad \text{SD}(\bar{X}_N) = \sqrt{\frac{p(1-p)}{N}}.$$

This expression shows that the typical magnitude of the estimation error decreases at the rate $1/\sqrt{N}$. Consequently, to reduce the standard deviation(SD) of the estimator by a factor of two, the number of trials must be increased by a factor of four. We will prove this theory numerically in the result sections[4].

2.5 Central Limit Theorem and error distribution

While the Law of Large Numbers guarantees convergence of the sample mean, it does not describe the distribution of the estimation error. This is provided by the Central Limit Theorem. If $X_1, X_2, \ldots$ are i.i.d. random variables with mean μ and variance $\sigma^2 > 0$, then

$$\frac{\sqrt{N}\,(\bar{X}_N - \mu)}{\sigma} \xrightarrow{D} \mathcal{N}(0,1) \quad \text{as } N \to \infty.$$

For Bernoulli(p) trials, $\mu = p$ and $\sigma^2 = p(1-p)$, so the normalized error

$$Z_N = \frac{\bar{X}_N - p}{\sqrt{p(1-p)/N}}$$

converges in distribution to a standard normal random variable. This result explains how the sample mean typically fluctuates around its limiting value and motivates the use of normal approximations when quantifying uncertainty.

2.6 Confidence intervals

The Central Limit Theorem provides a practical way to quantify uncertainty in the sample mean. For large N, the distribution of $\bar{X}_N$ can be approximated by a normal distribution with mean p and variance $p(1-p)/N$. As a result, an approximate 95% confidence interval for p is given by

$$\bar{X}_N \pm 1.96 \sqrt{\frac{p(1-p)}{N}}.$$

In practice, the parameter p is unknown and must be estimated from the data. A common approach is to use the plug-in estimator $p \approx \bar{X}_N$, which yields the empirical confidence interval

$$\bar{X}_N \pm 1.96 \sqrt{\frac{\bar{X}_N(1-\bar{X}_N)}{N}}.$$

These confidence intervals provide a clear numerical measure of the uncertainty in the estimator and will be used in the Results section to describe convergence in a precise, quantitative way. [5].

2.7 Concentration inequalities

Concentration inequalities bound the probability that the sample mean deviates from its expected value for a finite number of trials. Chebyshev's inequality gives the general bound

$$\mathbb{P}\big(\big|\bar{X}_N - p\big| \geq \varepsilon\big) \leq \frac{p(1-p)}{N\varepsilon^2}.$$

For Bernoulli trials, Hoeffding's inequality provides a sharper bound:

$$\mathbb{P}\big(\big|\bar{X}_N - p\big| \geq \varepsilon\big) \leq 2\exp\big(-2N\varepsilon^2\big).$$

These results give explicit finite-sample error bounds for the sample mean.

2.8 Simulation design and implementation

Monte Carlo simulations are used to study the convergence behavior predicted by the theoretical results. For increasing values of the sample size N, independent Bernoulli(p) samples are generated and the corresponding sample means are computed.

To assess convergence rates and variability, the simulation is repeated multiple times for each value of N. This enables simulation-based estimation of the standard deviation of $\bar{X}_N$ and direct comparison with the theoretical rate $\sqrt{p(1-p)/N}$.

3 Results

3.1 Running mean and convergence

We first examine convergence of the sample mean for Bernoulli trials with $p = 0.5$. Figure 1 shows the running sample mean $\bar{X}_N$ as a function of N. As predicted by the Law of Large Numbers, the estimator fluctuates for small N but stabilizes around the expected value $p = 0.5$ as the number of trials increases.
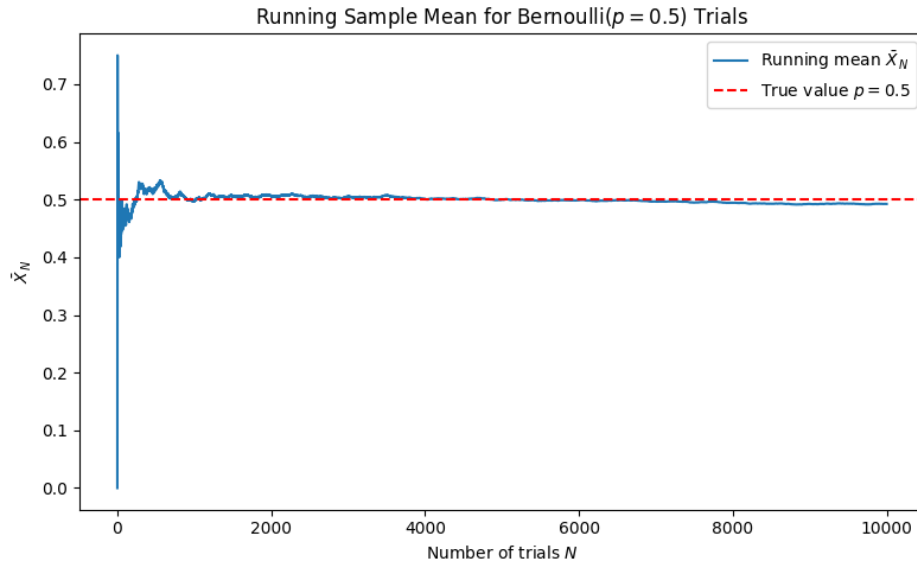

Figure 1: Running sample mean $\bar{X}_N$ for Bernoulli($p = 0.5$) trials.

3.2 Verifying the $1/\sqrt{N}$ convergence rate

A quantitative prediction from the variance calculation is

$$\mathrm{SD}(\bar{X}_N) = \sqrt{\frac{p(1-p)}{N}},$$

which implies that the standard deviation of the sample mean decays at the rate $1/\sqrt{N}$. To verify this numerically, we fix $p = 0.5$ and, for each sample size N, repeat the simulation many times and compute the empirical standard deviation of $\bar{X}_N$ across repetitions.

Figure 2 compares the empirical standard deviation with the theoretical curve $\sqrt{p(1-p)/N}$. The close agreement confirms the predicted $1/\sqrt{N}$ scaling. Table 1 reports the numerical values.

N	Empirical SD	Theory SD	Ratio
50	0.07130	0.07071	1.008
100	0.04987	0.05000	0.997
200	0.03593	0.03536	1.016
500	0.02282	0.02236	1.021
1000	0.01547	0.01581	0.979
2000	0.01113	0.01118	0.996
5000	0.00701	0.00707	0.991

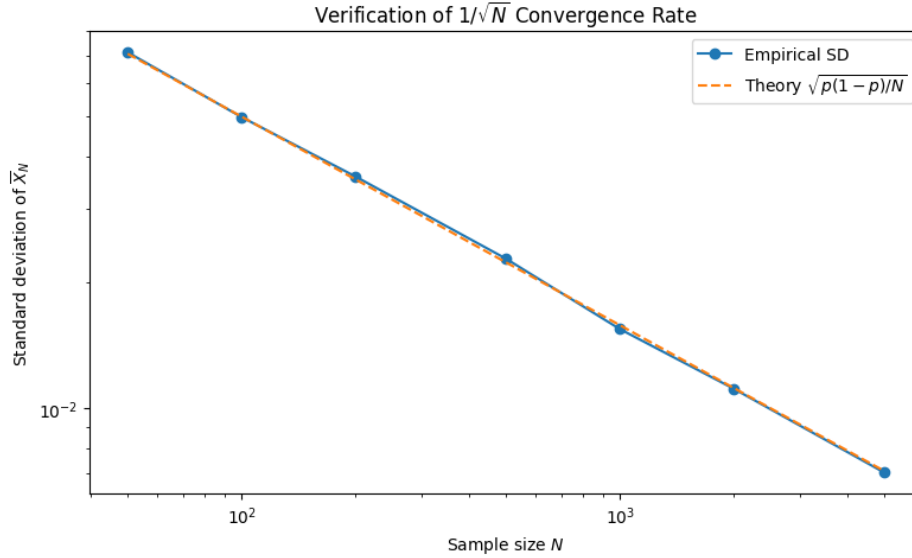Table 1: Empirical and theoretical standard deviations of the sample mean for Bernoulli($p = 0.5$) trials.

Figure 2: Empirical SD of $\bar{X}_N$ versus the theoretical $\sqrt{p(1-p)/N}$ on a log–log scale (Bernoulli $p = 0.5$).

3.3 Central Limit Theorem: empirical verification

To assess the accuracy of the normal approximation, we examine the distribution of the normalized estimation error

$$Z_N = \frac{\bar{X}_N - p}{\sqrt{p(1-p)/N}}$$

for a fixed large sample size N. The experiment is repeated many times to obtain an empirical distribution of Z_N.

Figure 3 shows a histogram of the normalized errors together with the standard normal density. This verifies that, for large sample sizes, the fluctuations of the sample mean around the true parameter are well described by a normal distribution, as predicted by the Central Limit Theorem.
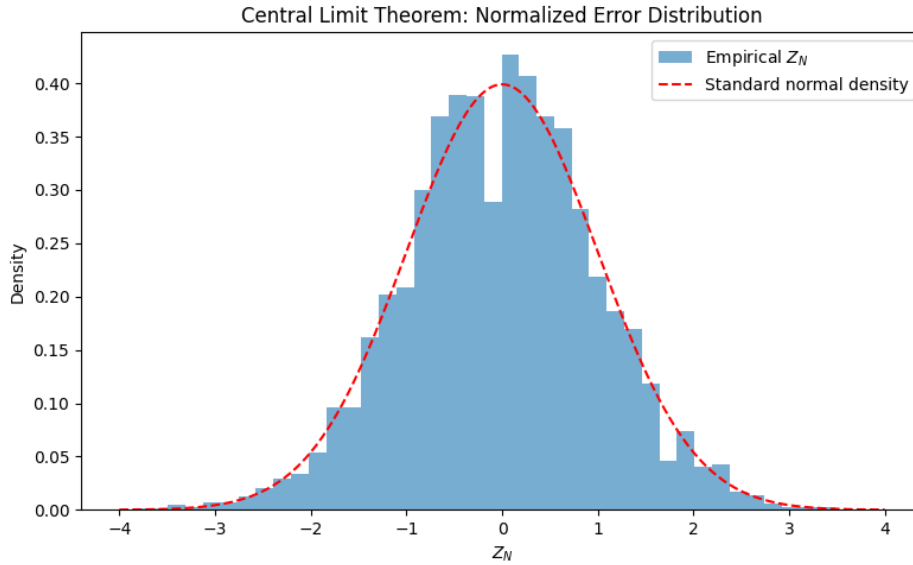

Figure 3: Histogram of normalized errors Z_N for Bernoulli($p = 0.5$) trials, compared with the standard normal density.

3.4 Confidence interval coverage

We evaluate the empirical coverage of nominal 95% confidence intervals for p. For each sample size N, the experiment is repeated many times and the proportion of intervals containing the true value is recorded. The empirical coverage approaches 95% as N increases, consistent with the Central Limit Theorem.
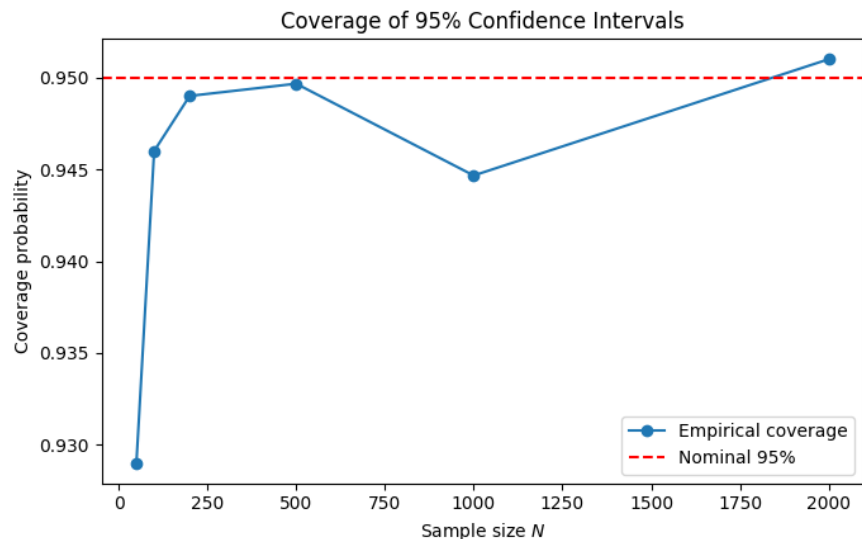
Figure 4: Empirical coverage of 95% confidence intervals versus sample size.

3.5 Rare-event convergence

Figure 5 shows the running sample mean for Bernoulli trials with a rare success probability ($p = 0.01$). The estimator exhibits large fluctuations for small and moderate sample sizes due to long stretches of failures before successes occur. As the number of trials increases, the sample mean gradually stabilizes near the true value, consistent with the Law of Large Numbers. However, compared to the symmetric case $p = 0.5$, substantially larger sample sizes are required to observe stable convergence.
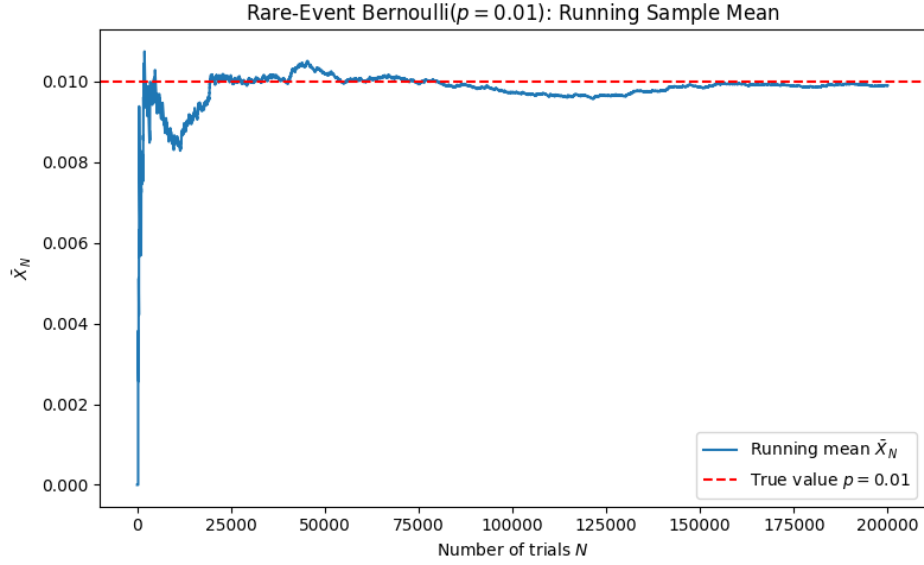
Figure 5: Running sample mean for Bernoulli($p = 0.01$) trials, illustrating slow convergence in a rare-event setting.

3.6 Failure of convergence for heavy-tailed distributions

Figure 6 shows the running mean of samples drawn from a Cauchy distribution. Unlike the Bernoulli case, the sample mean does not stabilize as the number of observations increases. Large fluctuations persist even for very large sample sizes, illustrating the failure of the Law of Large Numbers when the expected value is undefined.
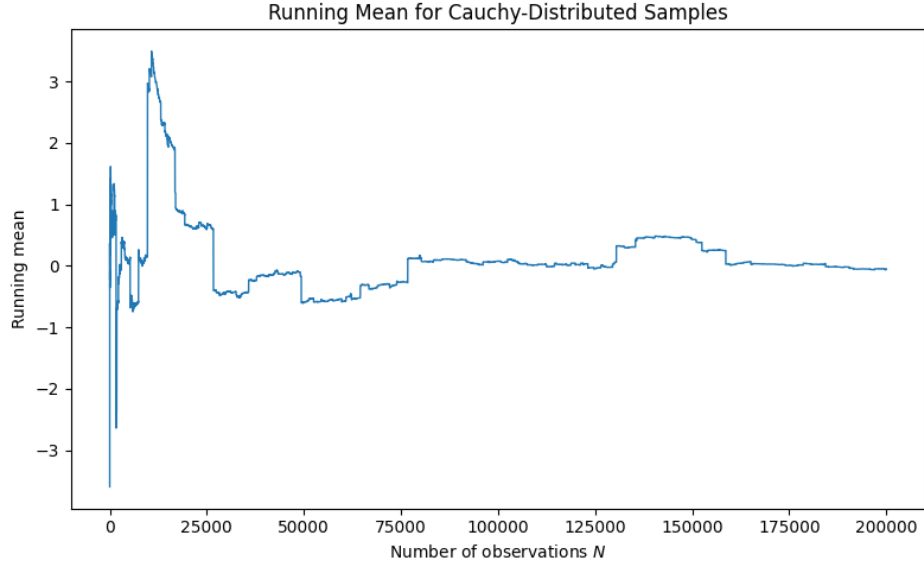
Figure 6: Running mean of Cauchy-distributed samples, showing persistent fluctuations and lack of convergence.

4 Conclusion

This project studied how the average of repeated random experiments behaves when the same experiment is carried out many times. Using ideas from probability theory together with Monte Carlo simulation, we focused on how the sample mean approaches the expected value as the number of trials increases. Starting with Bernoulli trials as a simple reference case, we showed that the sample mean does move toward the expected value, and that the size of its fluctuations becomes smaller as more trials are performed. In particular, we observed that the variability of the sample mean decreases at a rate proportional to $1/\sqrt{N}$, and we confirmed the Central Limit Theorem by examining the shape of the distribution of the scaled errors.

To describe uncertainty in a clear and precise way, we used confidence intervals. These intervals provide a numerical range that is likely to contain the true expected value and allow uncertainty to be stated in probabilistic terms rather than through informal descriptions of fluctuation.

Additional simulations demonstrated important limits of this convergence. In situations involving very rare outcomes, convergence still occurs in theory, but an extremely large number of trials is needed before it becomes visible in practice. On the other hand, when the underlying distribution does not have a well-defined average, as in the case of the Cauchy distribution, the sample mean does not settle toward any value at all. Together, these results emphasize that averaging works well only when the assumptions of the model are

satisfied. They also show both the usefulness and the limitations of Monte Carlo methods when estimating quantities through repeated random sampling.

Declaration of Use of Generative AI

Generative AI were used in the preparation of this report for the purposes of improving clarity, organization, presentation and LaTeX formatting. All mathematical reasoning, interpretation of results, and final decisions regarding content were reviewed and verified by the author.

References

[1] C. P. Robert and G. Casella. *Monte Carlo Statistical Methods*. Springer, 2nd edition, 2004.

[2] W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963.

[3] S. M. Ross. *Introduction to Probability Models*. Academic Press, 11th edition, 2014.

[4] S. M. Ross. *A First Course in Probability*. Pearson, 10th edition, 2019.

[5] L. Wasserman. *All of Statistics*. Springer, 2004.

A   Appendix: Python Code Summary

A.1   Core Monte Carlo simulation components

This appendix summarizes the key Python code components used to generate the results in this project. Only the essential parts of the implementation are shown.

Bernoulli trials and running mean

```python
import numpy as np


p = 0.5
N = 10000
X = np.random.binomial(1, p, size=N)
running_mean = np.cumsum(X) / np.arange(1, N + 1)
```

Empirical standard deviation of the sample mean

```python
M = 2000
samples = np.random.binomial(1, p, size=(M, N))
means = samples.mean(axis=1)


empirical_sd = np.std(means, ddof=1)
theory_sd = np.sqrt(p * (1 - p) / N)
```

Normalized error for CLT verification

```python
Z = (means - p) / np.sqrt(p * (1 - p) / N)
```

Confidence interval construction

```python
se = np.sqrt(means * (1 - means) / N)
lower = means - 1.96 * se
upper = means + 1.96 * se
```

Rare-event simulation

```python
p = 0.01
X = np.random.binomial(1, p, size=N)
running_mean = np.cumsum(X) / np.arange(1, N + 1)
```

Cauchy distribution counterexample

```
X = np.random.standard_cauchy(size=N)
running_mean = np.cumsum(X) / np.arange(1, N + 1)
```