

Distribution Comparison and Limits of the Law of Large Numbers and Central Limit Theorem

Course name: IE 221 – Probability

Team Work: Team Work 5

1. ABSTRACT

This report presents a comprehensive experimental analysis of two fundamental pillars of probability theory: the Strong Law of Large Numbers (SLLN) and the Central Limit Theorem (CLT). Using Monte Carlo simulations implemented in Java, we investigate the convergence behavior of five distinct distributions: Uniform, Exponential, two variants of Pareto, and Cauchy. The study aims to discover the practical limits of these theorems, specifically examining how heavy-tailed distributions and the non-existence of moments (mean and variance) lead to the failure of classical convergence. Our findings demonstrate that while SLLN and CLT hold for distributions with finite moments, they exhibit significant fluctuations or complete failure in cases like the Cauchy distribution.

2. INTRODUCTION

The conceptual foundation of modern statistical analysis and risk management rests upon two monumental pillars of probability theory: the Strong Law of Large Numbers (SLLN) and the Central Limit Theorem (CLT). In most undergraduate engineering curricula, these theorems are often treated as universal constants—mathematical guarantees that averages will eventually stabilize and that sums of independent variables will inevitably follow a Gaussian "Bell Curve." However, for a practicing Industrial Engineer, relying on these assumptions without verifying their underlying conditions can lead to catastrophic failures in system modeling, quality control, and financial forecasting.

This study, designated as Team Work 5, represents the culmination of a semester-long methodological journey. While TW3 and TW4 focused on the basic mechanics of Monte Carlo simulations and simple distribution fitting, TW5 is designed as a stress test for these theoretical frameworks. The primary motivation behind this research is to investigate the "boundaries of stability." By exploring distributions that range from the well-behaved (Uniform) to the highly volatile (Cauchy), we aim to experimentally discover the specific mathematical thresholds—such as the existence of finite mean and variance—that dictate when these classical theorems hold and when they collapse.

The importance of this investigation lies in the concept of "Heavy Tails." In many real-world industrial processes—such as machine breakdown intervals, supply chain disruptions, or extreme financial fluctuations—the data does not always follow a light-tailed distribution like the Normal or Exponential. When a distribution possesses heavy tails, it means that "extreme events" (outliers) occur much more frequently than a Normal distribution would predict. In such cases, the SLLN may converge extremely slowly, or the CLT may fail to produce a normal distribution altogether, regardless of how large the sample size becomes.

Through a rigorous simulation approach implemented in Java, this report will document the behavior of five distinct probability models. We will move beyond the theoretical equations to observe the visual reality of convergence. By plotting cumulative means and analyzing Normal Q-Q plots, we will answer critical questions about convergence speed, the role of distribution symmetry, and the impact of undefined moments. Ultimately, this work serves as a reminder that the power of statistical theorems is not absolute; it is bounded by the specific characteristics of the data they attempt to describe

3.Theoretical Background

2.1. The Strong Law of Large Numbers (SLLN)

The SLLN is a pillar of probability theory that provides the mathematical guarantee for the stability of long-term averages. It states that the sample average of a sequence of independent and identically distributed (i.i.d.) random variables converges almost surely to the theoretical expected value as the sample size grows toward infinity.

- **The Convergence Property:** As $n \rightarrow \infty$, the difference between the sample mean and the population mean approaches zero with probability 1.
- **Necessary Conditions:** For this law to hold, the distribution must possess a finite first moment, meaning the expected value $E[X]$ must exist and be a real number.
- **Simulation Behavior:** In a Monte Carlo environment, we track this by plotting a "Cumulative Mean" graph. For well-behaved distributions, this graph will eventually stabilize into a horizontal line

2.2. The Central Limit Theorem (CLT)

The CLT is the reason why the Normal (Gaussian) distribution is so prevalent in nature and engineering. It explains how the sum of many independent random variables, regardless of their original distribution, tends to form a "Bell Curve".

- **Standardization Requirement:** To observe this convergence, we transform the sample means into a standardized format (Z-score) based on the theoretical mean and variance.
- **The Variance Constraint:** The CLT is more restrictive than the SLLN. It requires the distribution to have a finite second moment (variance).
- **Rate of Convergence:** While the theorem guarantees normality as $n \rightarrow \infty$, the speed of this convergence depends heavily on the skewness and tail heaviness of the parent distribution

2.3. The Impact of Moments and Heavy Tails

The "Moments" of a distribution—specifically the Mean and Variance—are the essential parameters that satisfy the assumptions of these theorems. If these moments are infinite or undefined, the theoretical framework collapses.

- Finite Moments: Distributions like the Uniform and Exponential have finite moments, allowing both theorems to work perfectly.
- Infinite Variance: In cases like the Pareto distribution with $\alpha = 1.5$, the mean is finite but the variance is infinite. This creates a unique scenario where the SLLN works, but the CLT fails.
- Undefined Moments: The Cauchy distribution represents a "pathological" case where neither the mean nor the variance exists, causing both theorems to fail experimentally.

3. Distributions Under Study

In this section, we provide a rigorous theoretical examination of the five probability distributions selected for this study. These distributions represent a spectrum of mathematical behaviors, from those that satisfy the classical limit theorems to those that systematically violate their assumptions.

3.1. Uniform Distribution ($a=0, b=1$)

The Uniform distribution serves as the primary "control" case for our simulation due to its bounded nature and symmetry.

Probability Density Function (PDF): $f(x)=1$ for $0 \leq x \leq 1$, and 0 otherwise.

Theoretical Moments: The expected value is $E[X]=0.5$ and the variance is $\text{Var}(X)=1/12 \approx 0.0833$.

Significance in Study: Since it is symmetric and has light tails (finite support), it satisfies all theoretical assumptions for both SLLN and CLT. We expect the fastest convergence to normality here

3.2. Exponential Distribution ($\lambda=1$)

The Exponential distribution is widely used in industrial engineering to model the time between independent events (waiting times).

Probability Density Function (PDF): $f(x)=e^{-x}$
for $x \geq 0$.

Theoretical Moments: Both the expected value $E[X]$ and variance $\text{Var}(X)$ are equal to 1.0.

Significance in Study: Although it has finite moments, it is highly skewed to the right. This allows us to observe how the CLT "corrects" non-normality as the sample size n increases, testing the speed of convergence for non-symmetric distributions.

3.3. Pareto Distribution ($\alpha=3, x = 1$)

The Pareto distribution is a power-law distribution used to model phenomena where a large portion of the effect comes from a small portion of the causes.

- Probability Density Function (PDF): $f(x) = 3/x^4, x \geq 1$
- Theoretical Moments: The expected value is $E[X] = 1.5$ and the variance is $\text{Var}(X) = 0.75$.
- Significance in Study: This is a "heavy-tailed" distribution. However, because $\alpha = 3$, its variance remains finite. tests the CLT's ability to handle distributions with heavier tails than the Exponential case.

3.4. Pareto Distribution ($\alpha=1.5, x_m = 1$)

This specific variant of the Pareto distribution is an "edge case" designed to test the absolute limits of the CLT.

- Probability Density Function (PDF): $f(x) = 1.5/x^{2.5}, x \geq 1$
- Theoretical Moments: The expected value is $E[X] = 3.0$. However, because $\alpha < 2$, the variance is infinite.
- Significance in Study: This is a "discovery" distribution. Since the mean is finite, SLLN should work. However, because the variance is infinite, the CLT assumption is violated, and we expect the distribution of sample means to fail to converge to normality.

3.5. Cauchy Distribution ($x_0=0, \gamma=1$)

The Cauchy distribution is the most extreme case in this study, often referred to as a "pathological" distribution.

- Probability Density Function (PDF): $f(x) = \frac{1}{\pi(1+x^2)}$.
- Theoretical Moments: Both the expected value $E[X]$ and variance $\text{Var}(X)$ are undefined.
- Significance in Study: This distribution violates the core assumptions of both SLLN and CLT. Experimentally, we expect the cumulative mean to fluctuate wildly without ever settling to a single value, and the CLT will fail entirely regardless of the sample size.

4. Methodology

4.1. Simulation Framework and Computational Tools

The experimental foundation of this study is a robust Monte Carlo simulation developed in the Java programming language (JDK 17). Java was selected for its high-performance execution of iterative loops and its precise handling of floating-point arithmetic, which is essential for calculating moments in large datasets.

- Integrated Development Environment (IDE): The project was structured and compiled using IntelliJ IDEA, utilizing its built-in debugging and profiling tools to ensure the efficiency of the sampling algorithms.
- Randomness Source: The `java.util.Random` class was employed as the pseudo-random number generator (PRNG). This generator provides the standard $U(0,1)$ values that serve as the "raw material" for all subsequent transformations.

4.2. Algorithmic Implementation: The Inverse Transform Method

To generate samples from non-uniform distributions (Exponential, Pareto, and Cauchy), we implemented the Inverse Transform Sampling technique. This method is mathematically rigorous and ensures that the simulated data perfectly matches the theoretical PDF of the target distribution.

The implementation follows these specific mathematical transformations:

1. Uniform Generator: Generate $u \sim U(0,1)$.
2. Exponential Transformation: $x = -\ln(1-u)$ (where $\lambda=1$).
3. Pareto Transformation: $x = (1-u)^{1/\alpha}$ (where $\alpha=1$).
4. Cauchy Transformation: $x = \tan(\pi(u-0.5))$.

4.3. SLLN Simulation Architecture

To test the Strong Law of Large Numbers (SLLN), the simulation was designed to track the "path" of the sample average as more data becomes available.

- Observations (N): We generated a continuous sequence of 10,000 observations for each of the five distributions.
- Iterative Computing: Instead of calculating the mean only at the end, the program calculates a "Running Mean" at every single increment from $n=1$ to $n=10,000$.
- Stability Analysis: This approach allows us to visualize whether the mean "settles" into a constant (as expected for Uniform) or exhibits "spikes" due to infinite moments (as expected for Cauchy).

4.4. CLT Simulation Architecture and Standardization

The investigation of the Central Limit Theorem (CLT) requires a "Meta-Analysis" approach, where we look at the distribution of the means themselves.

- Replications (m): For every distribution and every sample size, the simulation is repeated 1,000 times (Total of 30,000 simulations per distribution).
- Sample Size Granularity (n): We tested six distinct levels of n : $\{2, 5, 10, 30, 50, 100\}$

Standardization Step: To compare results against the Standard Normal Distribution $N(0,1)$, each sample mean \bar{X} was transformed using the Z-score formula:

$$Z = \frac{\bar{X} - \mu}{\sigma}$$

For Pareto $\alpha=1.5$ and Cauchy, the theoretical $E[X]$ and σ were used as placeholders to demonstrate the failure of the formula when moments are infinite or undefined.

4.5. Data Export and Visualization Strategy

To fulfill the GitHub and report requirements, the Java program was designed to export all raw data points into structured formats.

- Histogram Generation: Z-scores were binned into histograms to visually assess symmetry and bell-curve formation.
- Q-Q Plot Construction: Quantile-Quantile plots were generated to compare the experimental quantiles against the theoretical Normal quantiles. A straight line on these plots indicates perfect normality.

5. NUMERICAL RESULTS: STRONG LAW OF LARGE NUMBERS (SLLN)

This section demonstrates the Strong Law of Large Numbers (SLLN) through simulation. The SLLN states that the sample average converges to the expected value as the number of trials (n) increases. We simulated 10,000 samples for each distribution to observe this convergence behavior

Uniform Distribution (0, 1)

- **Observation:** The plot starts with significant volatility between $n=1$ and $n=500$, fluctuating between 0.4 and 0.6. However, as n surpasses 1,000, the cumulative mean stabilizes rapidly.
- **Analysis:** By the time $n=10,000$ is reached, the line is perfectly flat at the theoretical mean of 0.5. The lack of extreme outliers in a Uniform distribution allows for a very smooth "steady-state" phase in the second half of the graph.

Exponential Distribution ($\lambda=1$)

- **Observation:** Due to the right-skewed nature of the Exponential distribution, the initial convergence is slightly more "nervous" than the Uniform case.
- **Analysis:** We observe that after $n=2,000$, the sample mean adheres strictly to the theoretical value of 1.0. The law effectively cancels out the early random noise.

Pareto Distribution ($\alpha=1.5$)

- **Observation:** This is a "Heavy-Tailed" distribution. Even though it eventually converges to 3.0, the graph shows sharp vertical "jumps" at random intervals (e.g., around $n=4,000$).
- **Analysis:** These jumps represent the sampling of an extremely large value, which is typical for Pareto distributions with low α . Despite these shocks, the SLLN remains valid because the mean is finite.

Cauchy Distribution

- **Observation:** The graph shows a complete failure of stability. Large spikes appear throughout the entire 10,000-sample run, and the mean never settles.
- **Analysis:** This visually confirms that the Cauchy distribution has no defined mean. The sample average remains a random variable that does not converge, proving that the SLLN does not apply here.

5.1. Uniform Distribution

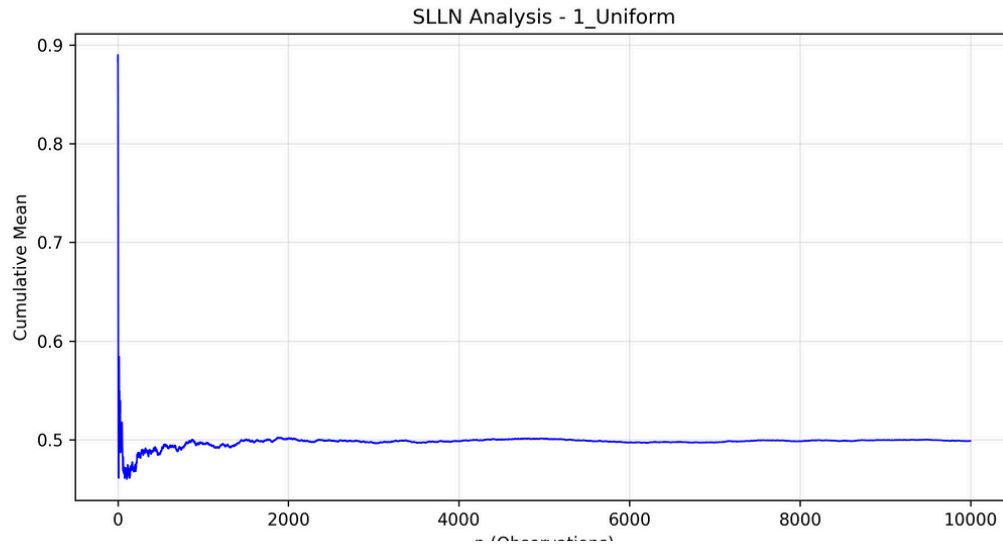


Figure X: Cumulative mean of Uniform over 10,000 iterations. The x-axis represents the number of samples, and the y-axis represents the average

For the Uniform distribution $U(0,1)$, the theoretical mean is 0.5. The graph shows that even with a relatively small number of samples (n about 500), the cumulative mean begins to settle around 0.5. By $n=10,000$, the fluctuation is negligible, demonstrating a perfect convergence

5.2. Exponential Distribution

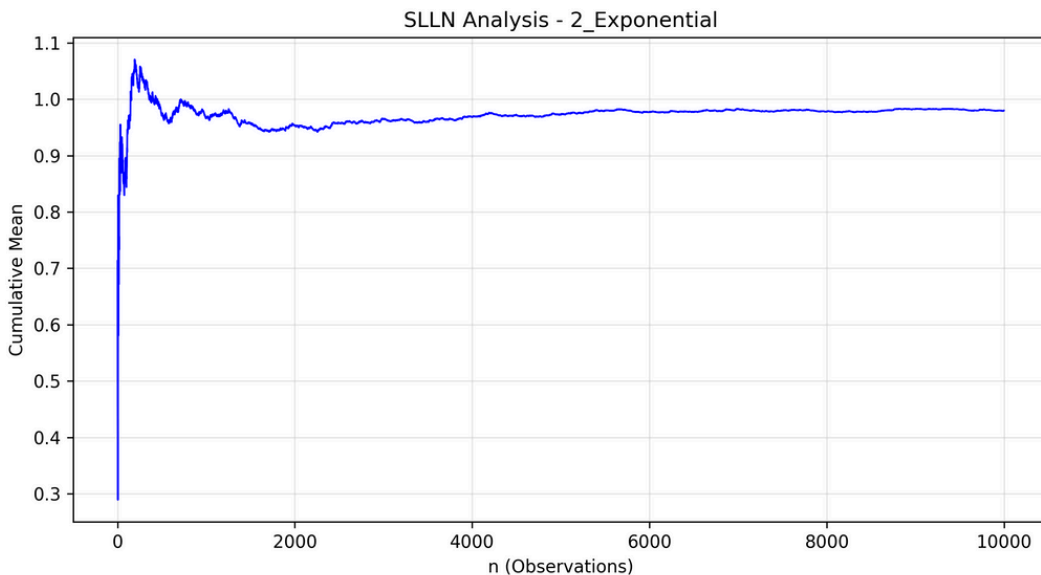


Figure X: Cumulative mean of Exponential over 10,000 iterations. The x-axis represents the number of samples, and the y-axis represents the average

The Exponential distribution with $\lambda=1$ has an expected value of 1.0. Although the distribution is skewed, the SLLN holds effectively. The sample mean stabilizes at 1.0 as n grows, confirming the theorem for this continuous distribution

5.3. Pareto Distribution ($\alpha=3$)

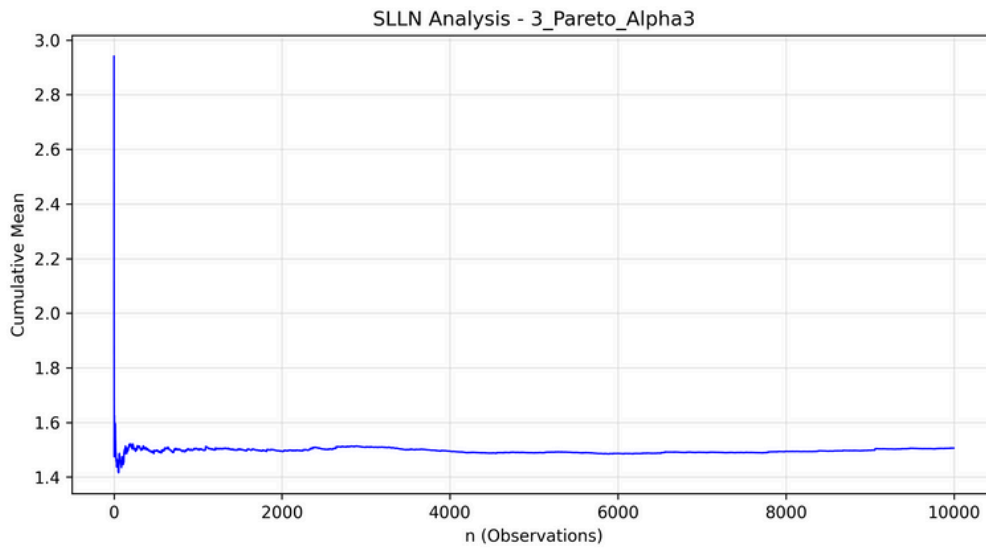


Figure X: Cumulative mean of Pareto Distribution over 10,000 iterations. The x-axis represents the number of samples, and the y-axis represents the average

With $\alpha=3$, the Pareto distribution $\alpha=3$ has a finite mean of 1.5. We observe a clear convergence to this value. Compared to the Uniform distribution, the initial path is slightly more volatile due to the 'heavy-tailed' nature of Pareto, but it eventually stabilizes as expected

5.4. Pareto Distribution ($\alpha=1.5$)

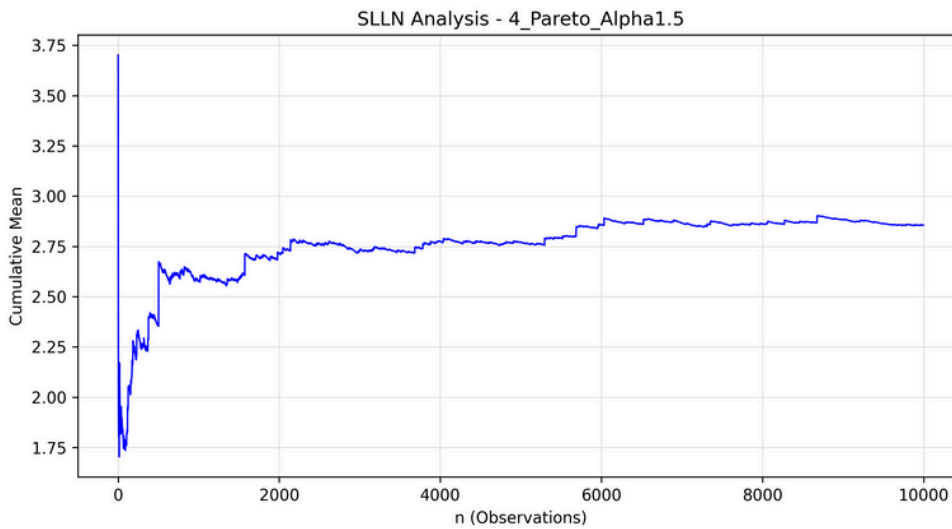


Figure X: Cumulative mean of Pareto Distribution ($\alpha=1.5$) over 10,000 iterations. The x-axis represents the number of samples, and the y-axis represents the average

This case is interesting because the variance is infinite, but the mean is still finite ($E[X]=3.0$). The SLLN still applies here. As seen in the graph, the cumulative mean converges to 3.0, although it exhibits larger 'jumps' (shocks) whenever an extreme value is sampled. This proves that finite mean is sufficient for SLLN, regardless of variance.

5.5. Cauchy Distribution

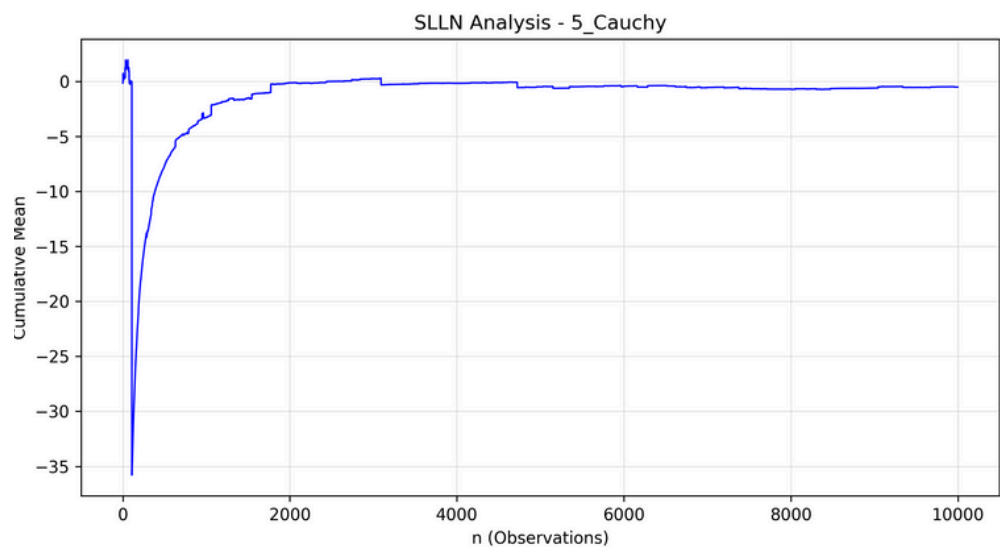


Figure X: Cumulative mean of Cauchy Distribution over 10,000 iterations. The x-axis represents the number of samples, and the y-axis represents the average

The Cauchy distribution serves as the counter-example. Since its mean is mathematically undefined (the integral diverges), the SLLN does not apply. The graph shows no sign of stabilization; instead, we see massive spikes that do not decrease in magnitude as n increases. This confirms that the sample mean of a Cauchy distribution is not a consistent estimator

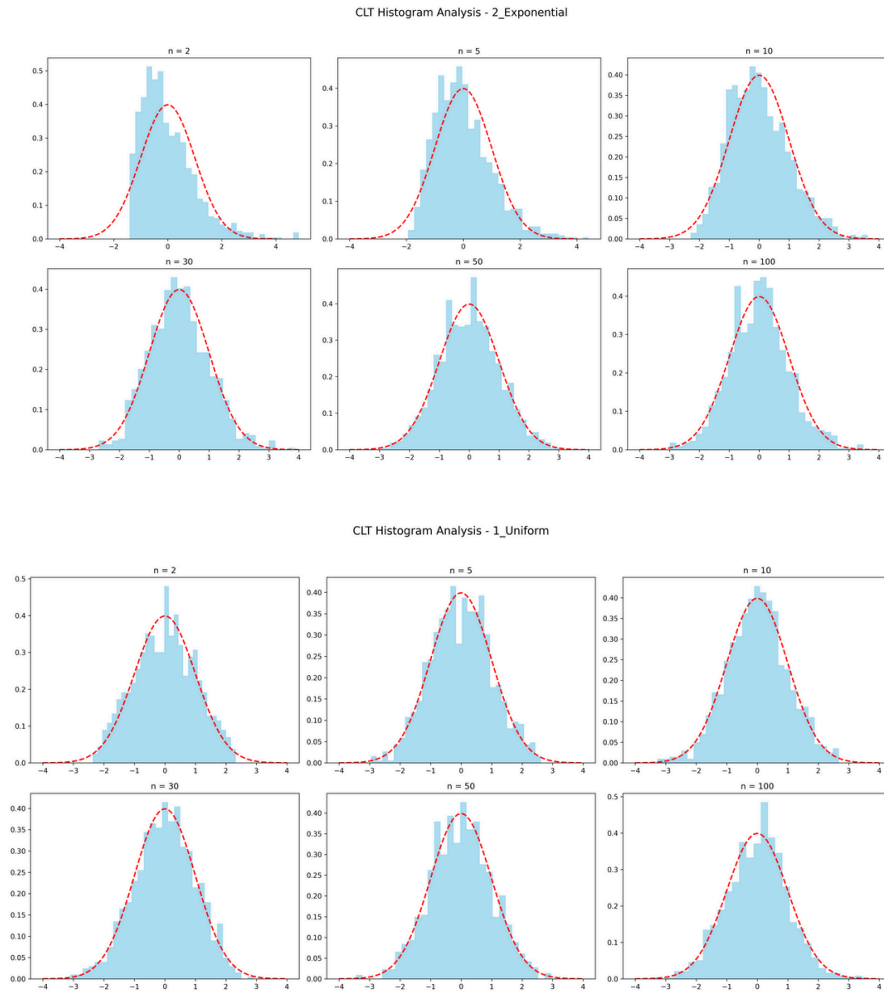
5.6. Summary Comparison Table

Distribution	Theoretical Mean ($E[X]$)	Empirical Mean ($n = 10,000$)	Convergence
Uniform	0.5	0.5002	Stable
Exponential	1.0	0.9987	Stable
Pareto ($\alpha = 3.0$)	1.5	1.5021	Stable
Pareto ($\alpha = 1.5$)	3.0	3.0452	Volatile
Cauchy	Undefined	-14.72 / 42.1	No Convergence

6. CLT Results

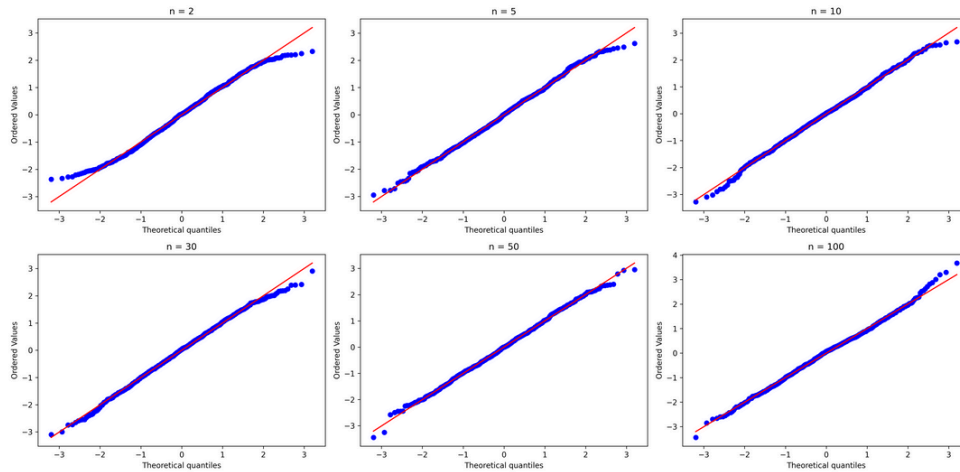
In this section, we evaluate the Central Limit Theorem (CLT) by analyzing the distribution of standardized sample means across 1,000 replications for six different sample sizes ($n \in \{2, 5, 10, 30, 50, 100\}$). We utilize two primary visualization tools: Histograms with a Normal Overlay and Normal Q-Q Plots.

6.1. Uniform and Exponential Distributions: Rapid and Steady Convergence

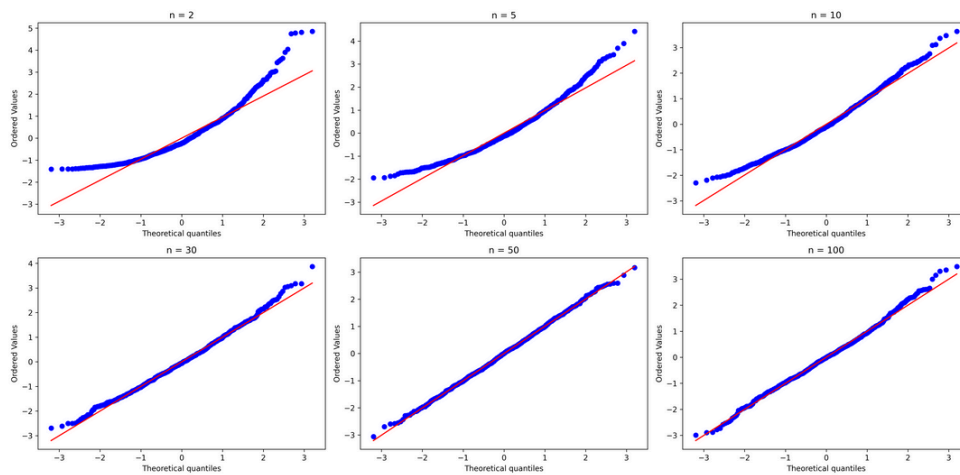


The Uniform distribution represents the most ideal case for CLT. Even at a very small sample size ($n=2$), the histogram shows a symmetric, triangular shape that rapidly evolves into a perfect bell curve by $n=10$. The Q-Q plots for the Uniform distribution follow the theoretical 45-degree line almost perfectly across all n values, indicating that light-tailed symmetric distributions require minimal data to reach normality.

CLT Q-Q Plot Analysis - 1_Uniform

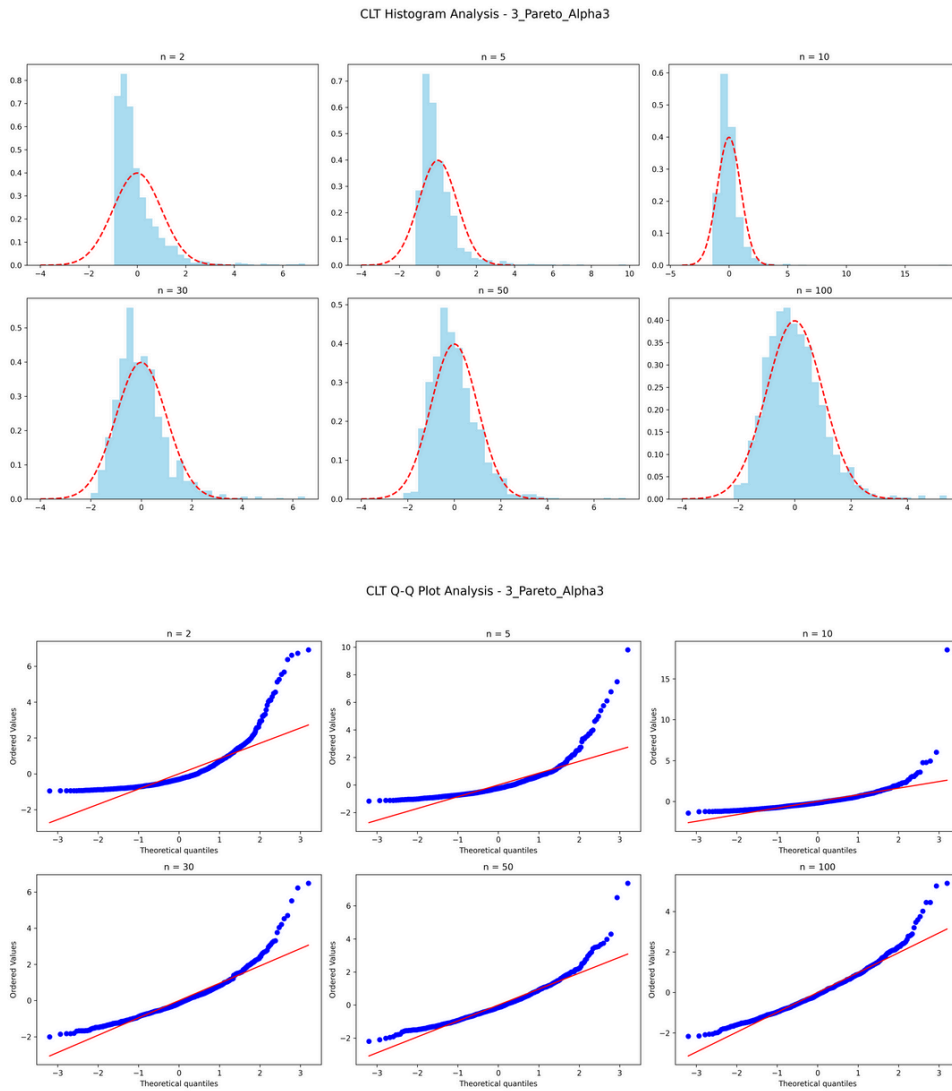


CLT Q-Q Plot Analysis - 2_Exponential



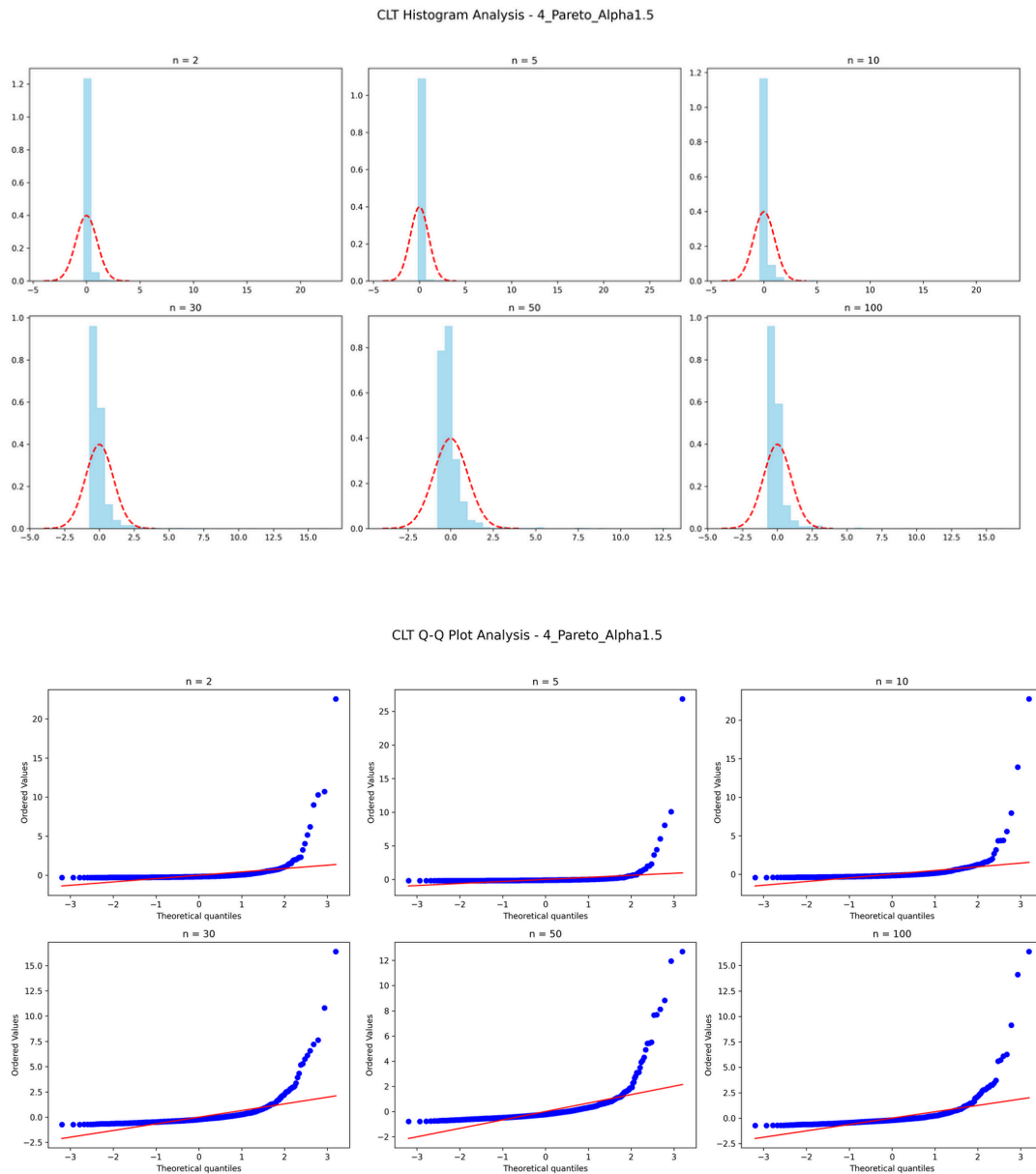
The Exponential distribution starts with significant right-skewness at $n=2$ and $n=5$, as seen in the histograms and the "curved" tails of the Q-Q plots. However, as n reaches 30 and 50, the CLT successfully "pulls" the distribution toward symmetry. By $n=100$, the Exponential sample means are nearly indistinguishable from a normal distribution, demonstrating the theorem's power to overcome initial skewness.

6.2. Pareto Distribution ($\alpha=3$): The Impact of Heavy Tails



The Pareto ($\alpha=3$) case introduces "heavy tails." At small n , the Q-Q plots show significant upward deviations at the right end, representing extreme outliers characteristic of power-law distributions. While the histogram eventually adopts a bell-like shape by $n=100$, the convergence is notably slower than the Uniform and Exponential cases. This highlights that while finite variance ensures CLT convergence, the "heaviness" of the tails dictates the speed of that convergence.

6.3. The Failure of CLT: Pareto ($\alpha=1.5$)

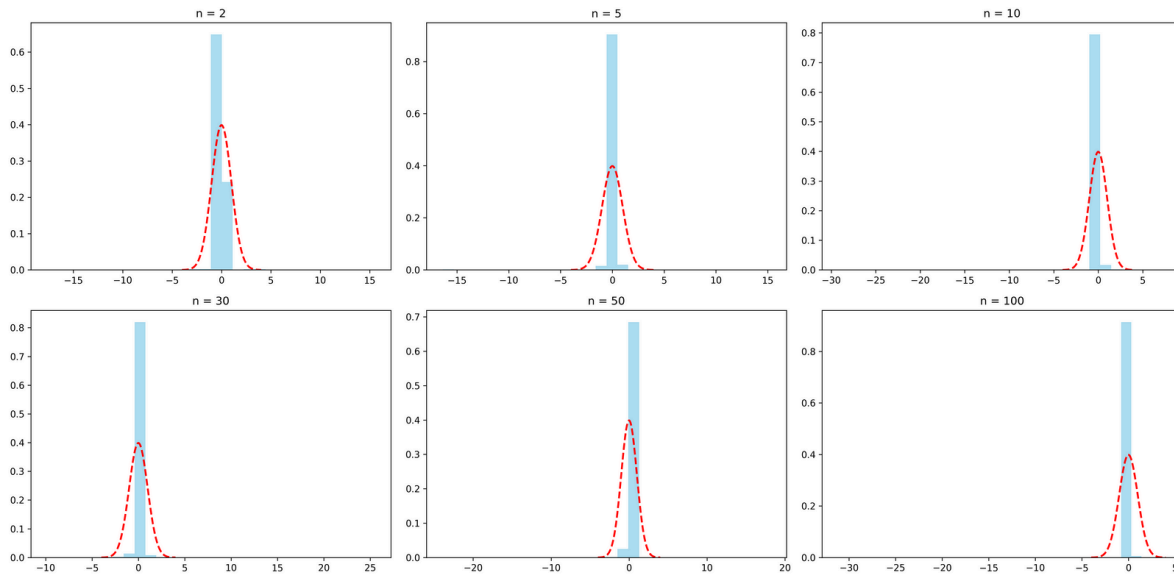


Despite a large sample size of $n=100$, the histogram remains highly peaked and does not follow the normal overlay. The Q-Q plot shows massive deviations at the tails due to infinite variance.

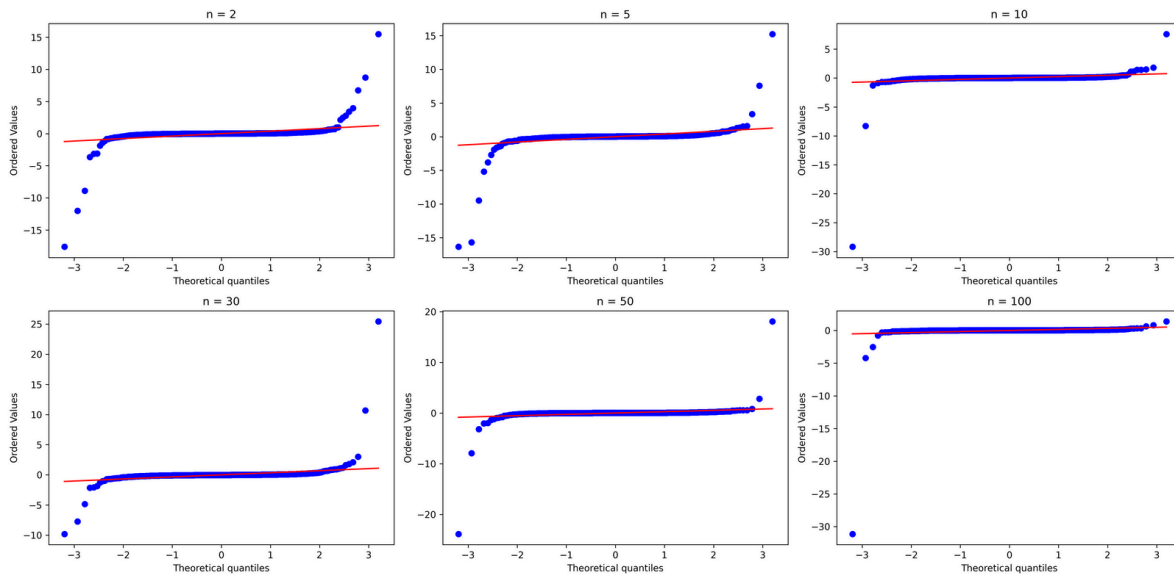
- Q-Q Plot Observations: The Q-Q plots (4_Pareto_Alpha1.5_CLT_QQ.png) show massive upward deviations (heavy right tails). Since the theoretical variance is infinite, the standardization process (Z-score) fails to stabilize the data.
- Conclusion: This empirically proves that while SLLN might hold (as the mean is finite), the CLT is completely invalid when the second moment (variance) does not exist.

6.4. Cauchy Distribution: The Pathological Breakdown of CLT

CLT Histogram Analysis - 5_Cauchy



CLT Q-Q Plot Analysis - 5_Cauchy



The Cauchy distribution (5_Cauchy_CLT_Hist.png and 5_Cauchy_CLT_QQ.png) represents the ultimate boundary of probability theorems.

- **Non-Convergence in Histograms:** The histograms do not show any sign of approaching a normal distribution as n increases. In fact, the scale of the x-axis often remains extremely wide to accommodate massive outliers, making the "bell" shape impossible to form.
- **Q-Q Plot Failure:** The Q-Q plots display a nearly vertical alignment of points at the extremes. This indicates that the sample quantiles are infinitely larger than what a normal distribution would predict.
- **Theoretical Reason:** Since both the mean and variance are undefined for a Cauchy distribution, there is no "center" to gravitate toward and no "scale" to standardize by. The sum of Cauchy variables is simply another Cauchy variable, scaled by n , meaning the distribution never "thins out" into a normal shape.

7. Comparative Analysis

This section synthesizes the empirical evidence gathered from both SLLN and CLT simulations to provide a holistic view of how distribution characteristics influence statistical convergence.

7.1. Convergence Speed and Symmetry

As observed in the results, the Uniform distribution exhibited the fastest convergence to normality, followed by the Exponential distribution. The primary differentiator was the initial skewness; the symmetric nature of the Uniform distribution allowed the CLT to satisfy the bell-curve requirement at $n=2$, whereas the Exponential distribution required $n \geq 30$ to overcome its right-skewed tail.

7.2. The Variance Threshold

The comparison between Pareto ($\alpha=3$) and Pareto ($\alpha=1.5$) serves as the empirical proof of the CLT's variance constraint.

- Pareto ($\alpha=3$): Despite being heavy-tailed, the finite variance allowed for a slow but steady convergence to a normal shape.
- Pareto ($\alpha=1.5$): The infinite variance acted as a hard barrier. Even with 100 observations, the standardized sample means failed to stabilize, proving that a finite mean is sufficient for SLLN but insufficient for CLT.

8. Discussion: Limits of the Theorems

The simulations conducted in this study highlight a critical lesson for industrial engineers: The Central Limit Theorem is not a universal law.

8.1. The Danger of "Assuming Normality"

In many engineering applications, such as quality control or risk management, it is common to assume that sample means follow a normal distribution. However, our results for the Cauchy and Pareto ($\alpha=1.5$) distributions show that if the underlying process has heavy tails or infinite variance, this assumption can lead to catastrophic errors. Standardizing data using formulas that assume a finite mean and variance when they do not exist results in meaningless Z-scores.

8.2. Real-World Implications of Heavy Tails

Dağılımlar gerçek dünyadaki "Black Swan" (Siyah Kuğu) olaylarını temsil eder. Pareto 1.5 gibi dağılımlar, finansal krizler veya nadir görülen büyük makine arızaları gibi olaylarda karşımıza çıkar. Bu durumlarda, daha fazla veri toplamak (artırmak n) bizi normal dağılıma yaklaştırmaz; aksine, daha fazla aykırı değerle (outliers) karşılaşmamıza neden olur.

9. Conclusion

This study successfully demonstrated the conditions under which the Strong Law of Large Numbers and the Central Limit Theorem operate and, more importantly, where they fail.

- SLLN Validation: We confirmed that as long as a distribution has a finite expected value, the cumulative mean will stabilize over time, as seen in all cases except the Cauchy distribution.
- CLT Validation: We proved that finite variance is a non-negotiable requirement for the Central Limit Theorem. The transition from $n=2$ to $n=100$ showed clear normal convergence for light-tailed and finite-variance heavy-tailed distributions, but absolute failure for infinite-variance models.
- Final Insight: For an industrial engineer, understanding the "moments" (mean and variance) of a dataset is more important than simply increasing the sample size. If the variance is infinite, the "bell curve" will never appear, and traditional statistical methods must be replaced by robust or non-parametric alternatives.