Dean&Francis
ISSN 2959-6157

# Proof and Application of Law of Large Number and Central Limit Theorem

**Yiqiang Zhang**

Shenzhen Foreign Language School, Shenzhen, China

*Corresponding author: charlies2026@outlook.com

**Abstract:**

The law of large number (LLN) and central limit theorem (CLT) have been important in probability, statistics, finance, and other fields for several centuries. They are also core theorems in probability. Through the work of Laplace, Chebyshev, Markov, Lyapunov, Bernoulli, and other famous mathematicians, the proof and forms of LLN and CLT have been greatly improved and expanded. They enable people to extract useful information from a large amount of data for statistical inference and prediction. Nevertheless, more scholars nowadays propose slightly improved proof for them, which may inspire people to use them better and understand their core ideas inside their elegant form. This study unravels certain proofs and applications of these two theorems to try to inspire scholars some further study of these theorems. This article uses a combination of proof of LLN and CLT and their application in different fields. Using different lemmas, such as Markov's inequality and Chebyshev's inequality, it discusses the process of different proofs of LLN and CLT. Based on many scholars' research, it summarizes the application of LLN and CLT in many different fields. This article can provide learners a fundamental idea of LLN and CTL. These can help readers quickly catch the important points of LLN and CLT, which can help many different investigations in different fields, including Math, Physics, Economics, and various fields with data analysis.

**Keywords:** Chebyshev's inequality; Markov's inequality; Law of large number; Central limit theorem.

## 1. Introduction

The history of law of large number (LLN) and central limit theorem (CLT) begins with the use of $e^{-x^2}$ by Jacob Bernoulli as an approximation tool. It continues with the theory of errors of observation, mathematical astronomy problems, the emergence of the hypothesis of elementary errors, and Laplace's foundational work. The development of an abstract CLT goes through the work of Chebyshev, Markov, and Lyapunov [1]. This causes the important and profound period during the development of the CLT in

ISSN 2959-6157

1713-1901 [1]. With the development of these two theorems, their application also continues to become wider. The LLN can be used for sampling estimation of mean value when the sample space is very large, as they provide the value of mean for a sample based on definition. It can also be used to solve problems of some companies to crack financing problems. Even in geometry or under mixing conditions, CLT can be used to solve problems if the fundamental preconditions of CTL are satisfied. However, there are many different situations where people cannot ensure they can be used or not. The current work scholars should do with them is to find their wider application and build other theorems and deductions based on them, which can raise their importance and further help in other fields.

This article, by providing detailed proofs and application of LLN and CTL, helps scholars catch the core idea of CTL and LLN and further research in many other different fields. This paper first provides a fundamental introduction to LLN and CTL and their several methods of proofs, including using the Borel-Cantelli lemma and Poisson limit theorem. Based on these, the paper states several applications of CTL and LLN in diverse fields, including their use in economics, math, business, and so on, especially in data analysis. This paper provides an important basis for understanding the origin of CTL and LLN, their role in probability theorems, and their important application in different fields, which can help readers look at them in a whole picture.

## 2. Lemmas and Theorems.

### 2.1 Lemmas

Before proving CTL and LLN, the author will give several lemmas.

Lemma 1. The Markov inequality states that

$$P\left(x \geqslant a\right) \leqslant \frac{EX}{a} \tag{1}$$

Proof. Let X be any positive continuous random variable (RV), one can write that the expectation $EX = \int_{-\infty}^{\infty} xf\left(x\right)dx = \int_{0}^{\infty} xf\left(x\right)$ $dx \geqslant \int_{a}^{\infty} xf\left(x\right)dx \geqslant \int_{a}^{\infty} af\left(x\right)dx = a\int_{a}^{\infty} f\left(x\right)dx = aP(x \geqslant a)$.

So, $P(x \geqslant a) \leqslant \frac{EX}{a}$.

Lemma 2. The Chebyshev's inequality states that

$$P\left(Y-\mu \geqslant a\right) \leqslant \frac{\sigma^2}{a^2} \tag{2}$$

For continuous RV, assume $E\left(Y\right)=\mu$ and $Var\left(Y\right)=\sigma^2$. Based on Markov inequalities, $P(x \geqslant a) \leqslant \frac{EX}{a}$. Let X= $(Y-EY)^2$, so $P((Y-EY)^2 \geqslant a^2) \leqslant \frac{E(Y-EY)^2}{a^2}$. Notice that P(( $Y-\mu)^2 \geqslant a^2$ )=P( $Y-\mu \geqslant$ α), so P( $Y-\mu \geqslant$ α)$\leqslant \frac{\sigma^2}{a^2}$. This is called Chebyshev's inequalities.

Lemma 3 (Borel-Cantelli lemma). Assume $\{A_n\}$ is a set of sequence of events, if $\mathrm{P}\left(A_n\right) < \infty$, then Near-certain event $A_n$ only happens a limited number of times.

Proof. Define event B= $lim\,sup_{n\to\infty} A_n$, it is found that

$$B = \bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_k \tag{3}$$

Consider event $B_n = \bigcup_{k=n}^{\infty} A_k$, then $B_n \supset B_{n+1}$, and B $\bigcap_{n=1}^{\infty} B_n$. Based on monotonicity of probability, it is found that $P(B_n) \geqslant P\left(B_{n+1}\right) \geqslant \cdots \geqslant \mathrm{P}\left(B_n\right)$ and $P\left(B_n\right) \to P\left(B\right)$.

Since $B_n$ is the union of $A_k$, one has

$$P\left(B_n\right) \leqslant \sum_{k=n}^{\infty} P(A_k). \tag{4}$$

If $\sum_{n=1}^{\infty} P(A_k)$ converges, then $\sum_{k=n}^{\infty} P(A_k)$ approaches 0, so $P\left(B_n\right) \to 0$, and thus P(B)=0. This implies that event $A_n$ will almost surely occur only finitely many times, so P( $lim\,sup_{n\to\infty} A_n$) =0.

Lemma 4 (Poisson limit theorem). Assume $X_n\, B\,(\mathrm{n}, p_n)$, and $\mathrm{n} \to \infty$, $p_n \to 0$, then $X_n$ tends to approach Poisson distribution.

Proof. Based on definition, it is found that $E\left[X_n\right] = np_n = \lambda$ and $Var\left(x_n\right) = np_n(1-p_n) \approx \lambda$. Thus, $P(x_n = k) = \left(\genfrac0ptnk\right) p_n^k (1-p_n)^{n-k}$. Using Stirling approximation, people can get $P(x_n = k) \approx \frac{\lambda^k e^{-k}}{k!}$.

### 2.2 Proof of Law of Large Numbers

The first is about the weak LLN. It states that

$$\lim_{n\to\infty} P(|x_n - \mu| \geqslant a) = 0. \tag{5}$$

On the one hand, one can prove it by using Chebyshev's inequalities. Assume $x_1, x_2, \cdots x_n$ are independent identically distributed (iid) RV, and assume $Var\left(x_i\right) = \sigma^2$. Based on Chebyshev's inequalities (lemma 2), $P(|x_n-\mu|\geqslant a)\leqslant\frac{Var(\mu)}{a^2}$. Because variables are i.i.d, $Var(\mu=Var(\frac{\sum_{i=1}^{n}x_i}{n})=\frac{\sum_{i=1}^{n}Var(x_i)}{n^2}=\frac{\sigma^2}{n}$. So, $P(|x_n-\mu|\geqslant a)\leqslant\frac{\sigma^2}{na^2}$. Notice $\lim_{n\to\infty}\frac{\sigma^2}{na^2}$=0, So $\lim_{n\to\infty}P(|x_n-\mu|\geqslant a)$=0.

On the other hand, it can also be proved by using Borel-Cantelli lemma. Assume $A_n=\{\overline{x_n}-\mu\geqslant a\}$, based on Chebyshev's inequalities, it is found that $P(A_n)\leqslant\frac{Var(\overline{X_n})}{a^2}$. So, it is obvious that $\sum_{n=1}^{\infty}P(A_n)$ converges, thus $A_n$ only happens in limit times, as $\overline{X_n}$ near surely converges at $\mu$.

The second is about the strong LLN. The author notes it can be verified by Glivenko-Cantelli theorem [2]. In addition, fourth order moments of random variables are analyzed and the strong number theorem is proved by proof by contradiction.

Assume random variables $X_i$ has fourth order moments $E[(x_i-\mu)^4]$ exist and are limited. The fourth moment of the sample mean is calculated and its asymptotic behavior is analyzed. Based on Markov inequalities, $P(\overline{X_n}-\mu\geqslant a)\leqslant\frac{E[(\overline{x_n}-\mu)^4]}{a^4}$. Combined with the analysis results of fourth moment, it can be confirmed that $\overline{X_n}$ surely converges at $\mu$.

### 2.3 Proof of Central Limit Theorem

The author starts by using characteristic function. Assume $X_1, X_2, X_3, X_4\cdots, X_n$ are i.i.d RV, the mean value is $\mu$, varience is defined as $\sigma^2$. In addition, define Normalized sum $\mathbb{Z}_n=\frac{1}{\sqrt{n\sigma^2}}\sum_{i=1}^{n}(X_i-\mu)$ and define characteristic function of random variables as $\varphi Z_n$(t)=E$\left[e^{it\frac{1}{\sqrt{n\sigma^2}}\sum_{i=1}^{n}(X_i-\mu)}\right]$. Because $X_i$ are i.i.d, then $\varphi Z_n(\mathrm{t})=\left[E(e^{it\frac{X_1-\mu}{\sqrt{n\sigma^2}}}\right]^n$. Based on Taylor expansion,

$\varphi Z_n(t)=\left[1+it\frac{E[X_1-\mu]}{\sqrt{n\sigma^2}}-\frac{t^2E[(X_1-\mu)^2]}{2n\sigma^2}+o\frac{1}{n}\right]^n$.

Based on definition, $E[X_1-\mu]=0$ and $\mathrm{E}[(X_1-\mu)^2]=$. $\sigma^2$ So, $\varphi Z_n(t)=\left[1-\frac{t^2}{2n}+o\frac{1}{n}\right]^n$. It is thus analyzed that the characteristic function behaves asymptotically $\varphi Z_n$ (t) $\approx$ exp(-$\frac{t^2}{2}$). This is exactly the characteristic function of the standard normal distribution $N(0,1)$.

In addition, by using Poisson limit theorem, one can assume $x_1, x_2, x_3\cdots x_n$ are i.i.d discrete RV and define normalized sum $\mathbb{Z}_n=\frac{S_n-n\mu}{\sqrt{n\sigma^2}}$. Based on Poisson limit theorem, when $n\to\infty$, $\mathbb{Z}_n$ tends to approach normal distribution. Assume $X_i\,B(1,p)$, then $S_n\,B(n,p)$. Based on Poisson limit theorem, one finds that $\lim_{n\to\infty}P(\frac{S_n-np}{\sqrt{np(1-p)}}\leqslant x)=\Phi(x)$. Here, $\Phi(x)$ is normal distribution function.

## 3. Application

### 3.1 Law of Large Number

People can utilize the law of large number for sampling estimation. In statistics, the large number theorem is widely applied to the problem of estimating the mean value of population from the sample mean. As long as the sample size is sufficient, the sample mean can be used to reliably predict the population mean regardless of the population distribution.

Also, it can be used in insurance and finance. The Theorem of large numbers is the theoretical basis for insurance companies and financial institutions to calculate risk and price. With reams of historical data, these institutions can make reasonable estimates of future risks. Some scholars have given a certain way to use the LLN to control credit risk, solve the lack of credit for small and micro enterprises, help financial lending companies control credit risks, and achieve crack financing problem goal [3]. Based on the law of large number, some scholars also prove that in the case of many economic agents, the individual risk will gradually disappear in the population, and the average outcome will converge to the expected value. This result suggests that large economies are inherently stable due to

ISSN 2959-6157

the averaging of individual uncertainties [4].

Additionally, LLN can be used in analyzing the neural network parameters. The empirical distribution of the neural network parameters converges to the solution of a nonlinear partial differential equation, according to some researchers' analysis of one-layer neural networks in the asymptotic regime of concurrently large network sizes and large numbers of stochastic gradient descent training iterations, which could be seen as LLN [5].

Also, LLN has application under wider conditions, including when the RV are dependent on and have different distribution. That is, the sample mean can converge almost deterministically to the expected value even though RV are dependent on and have different distribution, if certain requirements depending on the structure and distribution conditions are satisfied. Some scholars used Markov's inequality, Martingale theory, and other tools to get this result, which is an important extension of traditional LLN [6]. In this case, LLN can be used in more complex and wider conditions in different fields.

### 3.2 Central Limit Theorem

The CLT can be used for confidence interval estimation. The central limit theorem provides a theoretical basis for constructing confidence intervals. In practical applications, confidence intervals can be generated to estimate the population parameters by assuming that the sample mean distribution is nearly normal.

The CLT can also be used for SMC method and Bayesian inference. The CLT provides insight into the convergence properties of the estimators derived by the SMC method, describing how the pattern of distribution of the estimators approaches a normal distribution as the number of particles increases [7]. In this paper, scholars derive a CLT specific to the SMC method and show that under certain conditions, particle-based estimators will asymptotically follow a normal distribution.[7] The application of CLT ensures that the posterior distribution of particle approximations becomes more accurate as the number of particles increases. This result has important implications for Bayesian inference because it validates the effectiveness of using particle filters to approximate posterior distributions and provides a theoretical basis for evaluating the accuracy of estimators [7].

The central limit theorem is also applicable in time series analysis and econometrics. Some scholars have demonstrated that for a sequence of stationary RV that satisfy certain strong mixing conditions, the normalized sum of the sequence converges to a normal distribution, which means the CLT can be used in a wider range including under conditions of strong mixing [8]. Under strict conditions, the CLT can be used in geometry too. Some scientists have extended the use of the CLT to geometric probability, explaining that in random geometric structures, despite complex dependencies, normal limit behavior can still be observed under appropriate conditions and methods [9].

Besides the original form of CLT, variations of the CLT (such as the functional CLT) are useful for analyzing the long-term behavior of random processes, such as the long-term prediction of stock prices in financial models [10]. For random graphs and network structures, the CLT also provides a theoretical basis for distribution convergence properties of key quantities such as medium distribution and path length in large-scale networks, which is very important for understanding network behavior in the context of big data.

## 4. Conclusion

The LLN and CLT are fundamental in probability theory. They can be proved in various ways based on different lemmas, including Chebyshev's inequality and so on. They have different applications in different fields. With the current study, it is known that they can be used in insurance, analyzing neural network parameters, under more complex and wider conditions and long-term behavior of random processes. So based on the paper, the origin of LLN and CLT is evolutionary and meaningful, as they are very important parts of the evolution of probability theory. The proofs and application of LLN and CTL give readers a quick view of LLN and CLT, which can help the development of other theories and further investigation into their application in many fields. However, now the application and derived products from LLN and CLT are still limited. This paper also has some limits on describing some applications and some proofs of LLN and CLT in detail. Further investigation should focus on their further application in more fields that seem irrelevant. Overall, the total review of LLN and CLT on their proofs and application is consequential. Further study on their application and derived theories also has wide prospects.

## References

[1] Adams, William J. The Life and Times of the Central Limit Theorem. American Mathematical Soc., 2009.

[2] Etemadi, N. An Elementary school Proof of the Strong Law of Large Numbers. Zeitschrift Fur Wahrscheinlichkeitstheorie Und Verwandte Gebiete, 1981, (55):119-122.

[3] Huang Zi-jian, Wang Yan, et al. Law of large numbers, Internet finance: Solve the lack of credit for small and micro enterprises. Reform of the economic system, 2015, (01):163-168.

[4] Uhlig, H. A. Law of large numbers for large economies. Econ Theory, 1996, (8):41–50.
[5] Sirignano, Justin, Konstantinos Spiliopoulos. Mean Field Analysis of Neural Networks: A Law of Large Numbers. SIAM Journal on Applied Mathematics, 2020, (80):725–752.
[6] Andrews DWK. Laws of Large Numbers for Dependent Non-Identically Distributed Random Variables. Econometric Theory. 1988, 4(3):458-467.
[7] Nicolas Chopin. Central limit theorem for sequential Monte Carlo methods and its application to Bayesian inference. Ann. Statist. 2004, 32(6):2385 - 2411,
[8] Kwak, Sang Gyu and Jong Hae Kim. Central limit theorem: the cornerstone of modern statistics. Korean Journal of Anesthesiology, 2017, 70(2): 144-156.
[9] Florin Avram. Dimitris Bertsimas. “On Central Limit Theorems in Geometrical Probability.” Ann. Appl. Probab. 1993, 3 (4): 1033 - 1046.
[10] Brussolo, Monica E. Understanding the Central Limit Theorem the Easy Way: A Simulation Experiment. Proceedings, 2018, 2(21): 1322.